

A PRACTITIONER'S GUIDE TO ALTERYX

Alteryx 10

ji

A Practitioner's Guide to Alteryx®

Published by USEReady 469 7th Ave Suite #745, New York, NY - 10018 <u>www.useready.com</u> Copyright © 2015 by USEReady New York ISBN: 978-0-692-44795-6 Manufactured in the United States of America No part of this publication may be reproduced, stored in a retrieval system or

transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the USEReady 469 7th Avenue, Suite #745, New York, NY 10018 or online at http://www.useready.com/contact

Limit of Liability / Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a specific purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies provided here in might not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting professional services. Neither the publisher nor the author is liable for damages arising as a result of information published in this book. The fact that an organization or a website is referred to in this book as a citation and/or potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that internet sites listed in this work may have changed or disappeared between when this book was written and when it is read.

About USEReady

USEReady is a leading provider of advanced business intelligence, data management, and network security solutions. Our mission is to help businesses succeed by fast-tracking their business performance - through our digital world ready solutions, industry and technological expertise, agile global delivery practices, and our customer fanaticism.

Alteryx Solutions

Being their long-term partner, we understand Alteryx like no one else. Staying ahead of the rapidly evolving data-scape, USEReady has developed a specific, unified and democratized approach to data management to ensure quality, accelerated decision-making and collaboration. From profiling to optimization, we leverage Alteryx's powerful capabilities to address the biggest challenge faced by analysts today – how to access the right data quickly.

Our Alteryx expertise includes:



Domain specific data preparation



Reusable data preparation applications



Analyticsoriented prepared answer sets

API-driven Extract-Transform-Load



Cloud and on-premise integration



BI-ready modern data warehousing

Alteryx Training

We aim to give you the best chance of understanding the capabilities of Alteryx and make analytics interactive and explorable. You can choose from: virtual class-room or instructor-led, on-site format; as well as from Beginners, Intermediate or Advanced Training in Alteryx. We also offer customized Alteryx courses.

Our training will enable you to:

- Access, clean and blend data from multiple sources
- ✓ Join data sources of different types

- ⊘ Perform advanced predictive and spatial analytics









18.4

For more information, visit us at: **www.useready.com**, or write to us at: **contact@useready.com**

About Alteryx, Inc.

Alteryx, Inc. is the leader in data blending and advanced analytics software. Alteryx Analytics provides analysts with an intuitive workflow for data blending and advanced analytics that leads to deeper insights in hours, not weeks, which is typical of traditional approaches. Analysts love the Alteryx Analytics Platform because they can deliver deeper insights by seamlessly blending internal, third party and cloud data; and then analyze it using spatial and predictive drag-and-drop tools. This is all done in a single workflow, with no programming required. More than 1,000 customers and thousands of data analysts worldwide rely on Alteryx daily.

Visit www.alteryx.com or call 1-888-836-4274.

Alteryx is a registered trademark of Alteryx, Inc.

Acknowledgements

When we started writing this edition of the book, we decided to go with a team of authors instead of a single author. The team went through a rigorous iterative process of reading, updating and reviews to get the book where it is today.

I would like to express my gratitude to the team of authors and other colleagues who helped in making this book a reality. Without them the book truly would not be what it is today.

Uday Hegde: For being the driving force and a constant motivator.

Anand Vadul: For sharing the experiences and challenges faced in the last edition and being the technical guiding force.

Vijai Narasimha: For being one of main drivers of the project and putting in those long hours updating and reviewing the content.

Abhilash Ramanathan: For being a leader and helping the team with a very fine attention to detail.

Vijay Gowtham: For putting up with meetings at odd hours and being that much needed funny bone during our meetings.

Rahul Upadhye: For putting in endless hours of work in editing the book and always being committed.

Ashish Tergaonkar: For his enthusiasm in editing the book and his constructive suggestions during the meetings.

Rima Upadhye: For being involved despite a very busy schedule.

Prashant Singh: For creating marketing, a web presents, and an excellent cover design.

· Jagdeesh Narayanan

Letter from the CEO

First of all, we would like to thank our readers for making the first edition a grand success that we are super encouraged to attempt this edition. This edition is also a testament to growing popularity of Alteryx and appetite for self reliant solutions in the industry today.

We took a little bit of gamble with this edition. We invited a team of Alteryx practitioners within USEReady to contribute while enhancing the content with lot of experience and gotchas we learnt from the first edition. Jagdeesh Narayanan, Abhilash Ramanathan, Vijai Narasimha, Rahul Upadhye, Rima Upadhye, Anand Vadul, Vijay Gowtham, Ashish Tergaonkar and Prashant Singh have worked countless hours in shaping this book. The result has been phenomenal as you glean from the chapters.

Team has gone great heights to ensure the content is refined to the most recent version of Alteryx and exercises are useful to a fellow practitioner. With your valuable feedback we plan to augment the content to facilitate class-room like trainings.

We are grateful to our readers of the first edition and their constructive feedback has helped us improve this edition. We hope our efforts are well worth it and you are going to find this book useful.

Uday Hegde Chief Executive Officer USEReady

Foreword

As Vice President of product management at Alteryx, Inc. working closely with our product team, our customers and our partners, a training manual from USEReady is testament to the growing demand for easy-to-use data blending and advanced analytics solutions. USEReady knows the analytics industry and has created this book as both an independent guide and as a classroom aid to help its customers and others not only quickly learn Alteryx products, but more importantly grow in their ability to help gather deeper insights from their data.

The book has created a great launching point for beginners who want a manual, in addition to the real world use cases, so you can easily learn how best to use Alteryx Designer. By the time you finish working through this manual, someone who has never opened the Alteryx Designer before, will be able to create workflows, design reports, develop applications, and write macros to solve any of their data needs.

These are exciting times as the Alteryx community continues to grow globally, and we continue to witness an unprecedented demand for data analytics with actionable information. The Alteryx Designer and materials like this manual help the community of self-service data analysts make the most of their data.

We appreciate the partnership with companies like USEReady that know how to help these self-service analysts and reduce the time to insight with Alteryx.

Laura Sellers Vice President, Product Management of Alteryx

Preface

Every day we are faced with options, questions, and choices. These decisions, as we all know, are much easier to make when when we are well informed. Let's say that we want to eat. We litterally have an entire world of possibilites, given the proper resources. However, practically, there are real limitations. Are we at home without transportation? Are we backpacking in the mountians? Are we in the middle of Times Square? Do we have food restrictions for health reasons? Do we have \$5 or \$5000? What are we in the mood for?

This task that we all solve day in and day out depends on a considerable amount of information that we know about our world, and often take for granted. This information is all based on data about our world.

What Is Data?

Data is stored information. It comes in various forms ranging from the number and types of items on our desk, to the total mass of the universe, to the contents of this book, to the information in digital files and systems, which will be our focus.

What Does Data Do?

Data does nothing. It simply exists. It is what we do with data that is important. When we look at data we interpret it to create meaningful information, which gives us the ability to make better informed decisions.

How Do We Consume Data?

Data can be consumed in many forms. We can look at all of the raw data and read every piece individually. We can use aggregation methods to create summary data so that we can easily see high level trends. We can visualize the data because a picture truly is worth a thousand words. Since we often do not want to look directly at the original data source and read each individual piece of data, we need to perform data preparation.

What Is Data Preparation?

Data Preparation is the process by which raw data is converted into a clean, usable source for later consumption.

The three core components of data preparation are data retrieval, data manipulation, and data exportation. In more traditional analytic terms, data preparation refers to the extracttransform-load process referred to as ETL. However, in order to ease communication, we are going to avoid these technical terms and discuss the aspects of the processes as follows:

Data retrieval refers to the process of going to a data source, asking for data, and returning with the desired data.

Data manipulation refers to anything we decide to do to the data between the time we retrieve it and the time we export it.

Data exportation refers to what we do with data after we have extracted and manipulated it even if we haven't finished transforming it.

What is Data Manipulation?

Data retrieval and exportation are fairly straightforward; respectively, they can be likened to drawing water from a well and putting an ice cube in someone's drink. However, data manipulation is that tricky process of running the water through the pipes into our house, then filling the ice cube tray, then putting the tray in the freezer, and letting the water have enough time to freeze so that we have ice to consume. Going forward, we will be using an allegory to a river to explain the entire process of the data preparation and specifically the data manipulation portion.

Data manipulation can come in many forms which typically fall into three buckets as follows:

Combination

One of the most common problems with data is that it comes from multiple sources. It is generally possible to perform the analysis separately, or through a significant amount of manual effort, but these methods often leave something to be desired or are too slow for effective use. In order to solve this problem, we will be designing data streams that come together.

If we think about data streams as actual rivers, original data locations can be thought of as glacial streams, smaller rivers, or lakes. Bringing data together is like the tributaries that bring these different water sources together to form a river. Along the course of this river, way we can perform calculations.

Calculation

If the data is to be used, it is generally advisable to have as much data pre-calculated as possible. One reason for this is that it allows an organization to create a standardized formula for everyone's use. Another is that when we can run calculations before data is provided to a front end user or system, the consumer will be experience a much faster process.

If the data is to be used in a report, then the calculations are often fundamental aspects of that report.

Returning to the river analogy, we can think of calculations as hydroelectric dams along a river, we are using the resources that already exist in order to generate something new.

We may change the landscape, because we are changing the flow of water, and we are also slowing down the river (introducing calculations will slow down the data preparation process).

We also have the ability to transform the data stream into a more useable format.

Transformation

It is often the case that data is not in the format that we need. We may have been given access to a database that has data stored in a very machine readable format, and we need to pivot the table to make it human readable, or we may have been working with an Excel file which has data extremely normalized that makes it hard to use in a front end system. Either way we need to transform the structure of the data so that it can be effectively consumed. In thinking about the river, we can imagine this as the process of cutting a channel into the riverbed so that the river is deep enough to move barges up and down. In doing this, we are fundamentally changing the structure of the river in order to make it more useable.

To the reader

In the following chapters, we are going to cover many topics, but the format of the chapters will all be the same.

You will assume the role of a new consultant at a company that works with Alteryx. We introduce a business scenario, discuss the tools that we will use to solve the problem(s), walk through the initial problem(s), and then provide you a selfguided exercise. We conclude this book with a capstone assignment in *NYC*.

The exercises will use data that can be download from <u>http://www.useready.com/resources/a-practitioners-guide-to-alteryx</u> by following the instructions on the website to unpackage the file.

Additional data will also be needed to install the *US* 2010 *Census SF1* and *USGS North America Map* packages from <u>http://downloads.alteryx.com/data.html</u> which we will start using in *Cultural Musings*.

Let us know what you think by emailing us at <u>AlteryxBook@USEReady.com</u> and we will try to incorporate reader requests going forward.

If interested in Alteryx training sessions or Alteryx consulting, visit <u>http://www.useready.com/</u>.

Best of Luck,

USEReady

Table of Contents

About U	JSEReady	iii
About A	llteryx, Inc	v
Acknow	ledgements	vii
Letter f	rom the CEO	ix
Forewo	rd	xi
Preface		xiii
To the r	eader	xvii
CHAPTI	ER 1	
An Intro	oduction to Alteryx	
1.1	What Is Alteryx Analytics Platform?	
1.2	Alteryx Gallery	4
1.3	The Alteryx Interface	4
1.4	The Interface	9
1.5	File Menu	
1.6	Edit Menu	
1.7	View Menu	
1.8	Options Menu	
1.9	Help Menu	
1.10	Samples	
1.11	User Settings	
1.12	Toolbar	
1.13	Tool Palette	
1.14	Overview	23
1.15	Results Window	24
1.16	Properties Window	25
1.17	Interface Designer	
1.18	Canvas	
1.19	Using Tools	
1.20	Insert Tool Menu	
CHAPTI	ER 2	49

The Gar	nes	
2.1	Tools & Concepts	52
2.2	Improved Features	53
2.3	Browse	58
2.4	CrossTab	
2.5	Comment	64
2.6	Filter	
2.7	Formula	69
2.8	Input Data	71
2.9	Join	75
2.10	Output Data	
2.11	Running Total	80
2.12	Sample	81
2.13	Select	
2.14	Sort	89
2.15	Summarize	
2.16	Tool Container	
2.17	Transpose	
2.18	Union	96
2.19	Freestyle	
2.20	Let's Tidy Things Up	
2.21	Modern History	
2.22	Brains vs Brawns	
CHAPTI	ER 3	155
Unisex	Baby Names	
3.1	Tools & Concepts	
3.2	Imputation	
3.3	Multi-Field Formula	
3.4	Multi-Row Formula	
3.5	Text Input	
3.6	Gender Swapped	
3.7	What About Me?	
3.8	What's In A Name?	
CHAPTI	ER 4	
The Dir	ect Approach	

4.1	Tools	197
4.2	Append Fields	
4.3	Date Time Now	
4.4	DateTime	
4.5	Directory	
4.6	What's The Policy On That?	
4.7	Where, Oh Where Have The Three Files Gone?	
CHAPT	ER 5	217
Cultura	l Musings	
5.1	Tools & Concepts	
5.2	Allocate Input	
5.3	Find Replace	
5.4	Join Multiple	
5.5	Text To Columns	
5.6	Culturally Rich	
5.7	Culturally Divided	
Chapte	r 6	251
The Spo	ort Report	
The Spo 6.1	ort Report Tools & Concepts	251 254
The Spo 6.1 6.2	ort Report Tools & Concepts Charting	251 254 255
The Spo 6.1 6.2 6.3	ort Report Tools & Concepts Charting Create Points	251 254 255
The Spo 6.1 6.2 6.3 6.4	ort Report Tools & Concepts Charting Create Points Layout	251 254 255
The Spo 6.1 6.2 6.3 6.4 6.5	ort Report Tools & Concepts Charting Create Points Layout Render	251 254 255 264 266
The Spo 6.1 6.2 6.3 6.4 6.5 6.6	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer	251 254 255 264 266
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer Report Map	251 254 255 264 266 268 268 271 272
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer Report Map Report Header	251 254 255 264 266 268 268 271 271 272 272
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer Report Map Report Header Table	251 254 255 264 268 268 271 272 272 275 277
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer Report Footer Report Map Report Header Table Major Sports Teams	251 254 255 264 266 268 268 271 272 275 275 277 280
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10 6.11	ort Report Tools & Concepts Charting Create Points Layout Render Report Footer Report Map Report Header Table Major Sports Teams Football!	251 254 255 264 268 268 271 272 275 277 280 310
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10 6.11 CHAPT	Drt Report Tools & Concepts. Charting. Create Points. Layout. Render Report Footer. Report Map. Report Header Table. Major Sports Teams. Football!	251 254 255 264 268 268 271 272 275 277 280 310 313
The Spot 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10 6.11 CHAPT Expense	ort Report Tools & Concepts. Charting. Create Points Layout. Render Report Footer Report Map. Report Header Table. Major Sports Teams Football! ER 7 ive Beauty Products	251 254 255 264 266 268 271 272 272 275 277 280 310 313
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10 6.11 CHAPT Expens 7.1	ort Report Tools & Concepts. Charting. Create Points. Layout. Render Report Footer. Report Map. Report Header Table. Major Sports Teams. Football! ER 7 Tools & Concepts.	251 254 255 264 266 268 271 272 275 277 280 310 313 313 313
The Spo 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 6.10 6.11 CHAPT Expens 7.1 7.2	Drt Report Tools & Concepts. Charting Create Points Layout Render Report Footer Report Map Report Header Table Major Sports Teams Football! ER 7 Tools & Concepts Fuzzy Match	251 254 255 264 266 268 271 272 275 275 277 280 310 313 313 313 316 317

7.4	RegEx	322
7.5	Unique	
7.6	Tool Container	
7.7	Expensive Beauty Products	
7.8	More Flags	
Chapter	· 8	
Applica	tions Wanted	
8.1	Tools & Concepts	
8.2	Action	
8.3	Condition	
8.4	Date	
8.5	Date Filter	
8.6	Drop Down	
8.7	File Browse	
8.8	Radio Button	
8.9	Text Box	
8.10	Let's Build an App For That!	
8.11	To Summarize or Not to Summarize: That is the Question	on 404
Chapter	9	
Chapter Where's	• 9 s the loe?	
Chapter Where's 9.1	• 9 s the Joe? Tools & Concepts	405 405 408
Chapter Where's 9.1 9.2	9 s the Joe? Tools & Concepts Macro Input	405 405 408
Chapter Where's 9.1 9.2 9.3	9 s the Joe? Tools & Concepts Macro Input Macro Output	405 405 408
Chapter Where's 9.1 9.2 9.3 9.4	9 s the Joe? Tools & Concepts Macro Input Macro Output Map	405 405 408 409 411 413
Chapter Where's 9.1 9.2 9.3 9.4 9.5	9 5 the Joe? Tools & Concepts Macro Input Macro Output Map Map Input	405 405 408 409 411 413 416
Chapter Where's 9.1 9.2 9.3 9.4 9.5 9.6	9 s the Joe? Tools & Concepts Macro Input Map Map Input Numeric Up Down	405 408 408 409 411 413 413 416 419
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7	9 s the Joe? Tools & Concepts Macro Input Macro Output Map Map Input Numeric Up Down Spatial Match	405 405 408 409 411 413 416 419 421
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8	9 s the Joe? Tools & Concepts Macro Input Map Map Map Input Numeric Up Down Spatial Match Trade Area	405 408 408 409 411 411 413 416 419 421 424
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9	9 s the Joe? Tools & Concepts Macro Input Map Map Input Numeric Up Down Spatial Match Trade Area Where's the Joe?	405 408 408 409 411 413 413 416 419 421 424 424 426
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9 9.10	9 s the Joe? Tools & Concepts Macro Input Map Map Map Input Numeric Up Down Spatial Match Trade Area Where's the Joe? Who Wants Coffee?	405 408 409 411 413 413 416 419 421 424 426 448
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9 9.10 Chapter	9 s the Joe? Tools & Concepts Macro Input Map Map Map Input Numeric Up Down Spatial Match Trade Area Where's the Joe? Who Wants Coffee?	405 408 409 411 413 413 416 419 421 424 424 426 448 449
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9 9.10 Chapter Meta-M	9 s the Joe? Tools & Concepts Macro Input Map Map Map Input Numeric Up Down Spatial Match Trade Area Where's the Joe? Who Wants Coffee? t0	405 408 408 409 411 413 413 416 419 421 421 424 424 424 426 448 449
Chapter 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9 9.10 Chapter Meta-Ma 10.1	9 s the Joe? Tools & Concepts Macro Input Map Map Map Input Numeric Up Down Spatial Match Trade Area Where's the Joe? Where's the Joe? Who Wants Coffee? Tools & Concepts	405 408 409 411 413 411 413 416 419 421 424 424 426 448 449 449 452

10.3	Detour	455
10.4	Detour End	
10.5	Dynamic Rename	
10.6	JSON Parse	
10.7	Message	
10.8	All the Tools Along the Way	
10.9	Narrowing Down the Search	
CHAPTI	ER 11	
NYC		
СНАРТИ	ER 12	
Self-Gui	ded Solutions	
12.1	The Games: The Almighty Dollar	
12.2	Unisex Baby Names: What's In a Name?	
12.3	The Direct Approach: Where, Oh Where Have My Three I	Files
Gone?	493	
12.4	Cultural Musing: Culturally Divided	
12.5	The Sport Report: FOOTBALL!	
12.6	Expensive Beauty Products: More Flags	
12.7	Applications Wanted: To Summarize or Not to Summari	ze: That
is the	Question	
12.8	<i>Where's The Joe? Zip 5 Total Population Macro and Wor 501</i>	kflow
12.9	Metamorphosis: Narrowing Down the Search	502
12.10	Capstone: NYC	
Append	ices	
Append	ix A – File Types	509
Append	ix B – Hot Keys	
Append	ix C – Downloads/Content	
Append	ix D – Field Types	
Append	ix E- Properties Window	
Append	ix F – Boolean Expressions	
Append	ix G – Data Components	

Appendix H – Date/Time Units	517
Appendix I – RegEx Cheat Sheet	
Appendix J – Action Tool Sets	
Appendix K – Directory Tool Data	
Credited Original Data Sources	

CHAPTER 1 An Introduction to Alteryx

1.1 What Is Alteryx Analytics Platform?

The Alteryx Analytics Platform from Alteryx Inc. consists of two products: Alteryx Designer and Alteryx Server.

Alteryx Designer allows us to build a data stream (or module) and run it locally. In addition to a base package, Alteryx offers datasets and tool packages to expand the product to enhance computing capabilities. These packages permit additional features like drive-time analysis when TomTom data is available.

Alteryx Server has two main functions. The first allows users who have Alteryx Designer to publish applications for consumption. The second allows scheduled workflows to run without external intervention.

The Alteryx Analytics Platform has solved two major issues affecting the data preparation process: time to run the process and high specialization of labor. Typically, the preparation process is highly time-intensive because it is largely manual. While automation solves the time required to run, it still requires personnel with skills in computer programming. This means we need additional resources allocated to every project, eventually raising the project costs and often preventing projects from being taken on. Alteryx, Inc. has elegantly addressed both of these problems. It has designed an easy-to-use drag-and-drop interface where all we need to do is give each tool the appropriate settings. This simple concept has allowed business users to develop workflows that meet their exact needs much faster than has ever been possible because it provides business people with the tools to design or modify a fully or semi-automated data preparation.

1.2 Alteryx Gallery

The *Alteryx Gallery* is a public version of *Alteryx Server* where the community can upload interesting data streams, share ideas, and get inspired. The *Alteryx Gallery* can be accessed at gallery.alteryx.com.

1.3 The Alteryx Interface

By the end of the chapter, the reader should be able to understand all aspects of the Alteryx Designer Interface and how to work with the product.



Figure 1-1 - Alteryx Intro



Figure 1-2 - Workflow Configuration

Above are the two things we may see when we open Alteryx Designer. The first image is the default for when the program is first opened. Alteryx Designer asks us what we want to do with the session. These options, fairly clearly, are to open up the basic tutorials, open an existing module, or open a new module. The second image is what comes up when we have previously selected the option *Don't show this again and always open a blank canvas*. When using this option, Alteryx Designer always opens a new module. (For those who see the first image, please select "New Workflow.")

<u>F</u> ile	<u>E</u> dit	<u>V</u> iew	<u>Options</u>	<u>H</u> elp
🗹 <u>N</u> ev	v Workflo	w		Ctrl+N
Ope	en Recent			>
D pe	en Workfl		>	
Ope	en Autosa	ved Files		
Sav	e			Ctrl+S
Sav	e <u>A</u> s			>
Prin	t			
Prin	t Setup			
Exit				

Figure 1-3 - File Dropdown

The first thing we will do is save our module. Open the file drop-down at the top left of the screen, and select the *Save As* option as indicated above.

save in:	🕌 Ateryx		-	G 🦻 📂 🛄 🗸				
las .	Name 🔺		-	Date modified	• Туре	- Size	-	
S	New Workfi	ow1		3/31/2016 12:22 PM	Alteryx Workflow	2	7 KB	
nt Places	New Workfl	ow2		3/31/2016 2:34 PM	Alteryx Workflow	4	5 KB	
esktop braries mputer								
twork	File name:	New Workflow1					•	Sav
	Save as type:	Designer x64 Workfl	ows (".yomd)				-	Canc

Figure 1-4 - Save As

We can save our data stream as one of three file types. At a high level, we have the ability to save our data streams as:

• .yxmd

Files saved in this format are data streams in a standard *Alteryx Workflow*.

• .yxwz

Files saved in this format are data streams that have been designed as *Analytic Applications*, which means they have user inputs and can be run on *Alteryx Server*.

• .yxmc

Files saved in this format are data streams that have been designed as *Macros* that allow the reuse of the data stream.

Please save this workflow as *The Basics.yxmd*.

1.4 The Interface



Figure 1-5 - Alteryx Designer Interface

This is one potential view of the *Alteryx Designer Interface*, if every *View* option is on. The numbers represent each of the seven primary aspects of the interface. One through six can be disabled and will be discussed under the corresponding sections in *View*. The seventh is the canvas, where we build workflows.

1.5 File Menu

<u>F</u> ile	<u>E</u> dit	<u>V</u> iew	<u>O</u> ptions	<u>H</u> elp
	v Workflo	w		Ctrl+N
Ope	n Recent	t		>
🔊 <u>О</u> ре	en Workfl	ow		>
Ope	en Autosa	aved Files		
Save	e			Ctrl+S
Save	e <u>A</u> s		>	
Prin	t			
Prin	t Setup			
Exit				

Figure 1-6 - File Menu

The *File Menu* functions in the same way as most programs, with a few specialized options:

New Workflow: Opens a blank canvas to work in. (*Ctrl+N* means that there is a *Hot Key* that allows us to use this action without opening the drop-down. Pressing the *Ctrl* and *N* buttons at the same time does the same thing as selecting *New Workflow*.) (Hot Key: Ctrl+N)

Open Recent: Opens a list of the most recently opened workflows.

Open Workflow: Opens a file browser that allows us to open an existing *Workflow*, *Analytic App*, or *Module* in a tab of the canvas. (Hot Key: Ctrl+O)

Save <*File Name*>: Saves the workflow that is in the active canvas. (We will see the name of the current workflow instead of <*File Name*>) (Hot Key: Ctrl+S)

Print: Prints a copy of the workflow.

Print Setup: Opens a window that allows us to configure how the workflow will print if we choose to print out a copy.

Exit: Closes Alteryx Designer entirely. This differs from *Close* because *Close* will close the active canvas but *Exit* will close every open data stream.

1.6 Edit Menu

<u>E</u> dit	<u>V</u> iew	<u>O</u> pti
Undo		Ctrl+Z
C Redo		Ctrl+Y
<mark>⊁</mark> Cuṯ		Ctrl+X
Copy		Ctrl+C
🛱 <u>P</u> aste		Ctrl+V
Delet	e	Del
Element 1	7 Ed	A Manua

Figure 1-7 - Edit Menu

Undo: Allows us to undo the steps just performed, in reverse order. (Hot Key: Ctrl+Z)

Redo: Allows us to restore a step from what was undone, as long as we have not changed anything else. (Hot Key: Ctrl+Y)

Cut: Allows us to remove the selected items on the canvas while keeping a copy on the clipboard. (Hot Key: Ctrl+X)

Copy: Allows us to copy the selected items on the canvas to the clipboard. (Hot Key: Ctrl+C)

Paste: Allows us to take the most recent thing copied or cut into the clipboard and put it on the canvas. (Hot Key: Ctrl+V)

Delete: Removes the selected items on the canvas. (Hot Key: Del or Delete)



1.7 View Menu

Figure 1-8 - View Menu

Toolbar: Turns the toolbar on and off. When it is checked, the buttons on the toolbar are visible (1 in Figure 1-5). (Hot Key: Ctrl+Alt+B)

Tool Palette: Turns the tool palette on and off. When checked, we will see the two rows of drag-and-drop tools (2 in Figure 1-5). (Hot Key: Ctrl+Alt+T)
Overview: Turns the overview window on and off. When checked, we can see a high-level overview of our canvas (3 in Figure 1-5). (Hot Key: Ctrl+Alt+V)

Interface Designer: Turns the interface designer window on and off. When it is checked, we will see a window that allows us to modify the user interface (6 in Figure 1-5). (Hot Key: Ctrl+Alt+D)

Find Tool: Opens a pop-up window that allows us to quickly find particular tools on the canvas. (Hot Key: Ctrl+F)

1.8 Options Menu



Figure 1-9 - Options Menu

Schedule Workflow: Allows us to set up a schedule to run our workflow as long as we have Alteryx Server.

View Schedules: Allows us to look at all of the scheduled jobs that we have access to on Alteryx Server.

Run Analytic Apps: Allows us to run Alteryx Analytical Apps that we have created.

Export Workflow: Allows us to export the workflow along with the packages that have been created.

Activate License Key: Opens a window that asks for an activation key so we can use Alteryx Desktop.

Manage Licenses: Opens a window that allows us to look at historical license keys and see what we currently have available.

User Settings: Opens a window that allows us to customize the way the canvas looks and the way some of the default settings behave.

Advacnced Options: Allows us to manage aliases, workflow dependencies, and encrypt workflows.

Download Predictive Tools: Allows us to download tools for the purpose of predictive analytics from the Alteryx repository.

```
1.9 Help Menu
```



Figure 1-10 - Help Menu

Alteryx Help: Opens the Alteryx Help home page.

What's New: Opens the Alteryx Help release notes for the version of Alteryx that we are currently using.

Getting Started: Opens the Alteryx Getting Started page that is loaded in the beginning.

Sample Workflows: Allows us to access the Sample Workflows that are shipped with Alteryx.

Community: Opens the Alteryx Community web page.

Check for Updates: Checks our current version of Alteryx Designer and verifies if it is the most recent version.

Alteryx Downloads: Opens the Alteryx Download page where we can find additionally downloadable content. One example is the U.S. Census data.

About: Opens a window that tells us version, publisher, and serial number information about Alteryx Designer.

1.10 Samples

Help			
🕜 Alteryx Help	Tutorials	>	*
What's New	One-Tool Examples	>	Preparing Data
Getting Started	Basics	>	Filtering Data
Sample Workflows >	Data Artistry	>	Blending Data
Community	Presentation Samples	>	Analyzing Data
<u>Community</u>	Analytic Apps	>	Location Analysis
Check for Updates	Macro Samples	>	Build an App
Alteryx Downloads			Build a Macro 1
About			· · · · · · · · · · · · · · · · · · ·

Figure 1-11 - Samples

Under the *Help Menu*, there is the option to open Sample Workflows. Depending on what we have downloaded from Alteryx, we will see different samples.

These samples are good starting points to get used to Alteryx and to learn how to use tools that we do not know.

Walking through the samples is beyond the scope of this book, as all of the samples are well documented and show how to build the workflow step by step. However, below is a picture of what the *Preparing Data* sample looks like before you begin to work with it.

Darg an ingut Data tool and glace it on the image the workflow. In the Configuration Window, clack it on window, clack it window, clack it win	Contract a Select tool to contract a Select tool contract to the ADDR lease the Select CONANE and click the Select CONANE and clic	Cick the Run Workflow Dation on the toolbar. Select at ool in the Modular Select at ool in the workflow to display the dust areauts in the fiscaling tool to tool to see how the data changes. The data in new suitable to be used as a mailing bit.	How to Mew Yourkflow Results with 12-bit Installations ③		
---	---	---	--	--	--

Figure 1-12 - Sample Tutorial

1.11 User Settings

Under the *Options menu,* there is an option called *User Settings*. This sub-menu has four options.

<u>Options</u> <u>H</u> elp		
C Schedule Workflow ⊻iew Schedules Run Analytic Apps Export Workflow Activate License Key Manage Licenses		Edit <u>U</u> ser Settings Save Layout & Settings on E <u>x</u> it Save Layout & Settings <u>N</u> ow <u>R</u> estore Defaults
User Settings	>	
Advanced Options Download Predictive Tools	>	

Figure 1-13 - User Settings

Edit User Settings: Opens a pop-up window for the user to edit the settings affecting the way Alteryx Designer looks and behaves.

Save Layout & Settings on Exit: Allows Alteryx to copy the settings we had when Alteryx was last closed as the default when we next open the program. Unchecking the option can turn off this functionality.

Save Layout & Settings Now: Allows us to tell Alteryx to use the settings we currently have as the default behavior for the program. This is most effectively used when we are not using *Save Layout & Settings on Exit*.

Restore Defaults: Allows us to revert Alteryx Designer to the default settings.

1.12 Toolbar

Figure 1-14 - Toolbar

The *Toolbar* has buttons we can use to do the following tasks. It is the same as if we were to navigate to the parallel options in the menus described above. The buttons simply make it easier to perform the most common actions.

Allows us to create a new workflow. (Same as File > New Workflow) (Hot Key: Ctrl+N)

Allows us to open a workflow. (Same as File > Open Workflow) (Hot Key: Ctrl+O)

Allows us to save our current workflow. (Same as File > Save <File Name>) (Hot Key: Ctrl+S)

Allows us to cut our current selection. It is greyed out and inaccessible when we have nothing selected and dark when we do. (Same as Edit > Cut) (Hot Key: Ctrl+X)

Allows us to copy our current selection. It is greyed out and inaccessible when we have nothing selected and dark when we do. (Same as Edit > Copy) (Hot Key: Ctrl+C)

Allows us to paste what we have on the clipboard. If we do not have something on the clipboard that can be pasted into Alteryx, we will get a pop-up that says *Nothing to paste*. (Same as Edit > Paste) (Hot Key: Ctrl+V)

Allows us to undo what we just did. It is greyed out when we have nothing that we can undo. (Same as Edit > Undo) (Hot Key: Ctrl+Z)

Allows us to redo what we just undid. It is greyed out when we have nothing that we can redo. (Same as Edit > Redo) (Hot Key: Ctrl+Y)

C Allows us to schedule the workflow on Alteryx Server. (Same as Tools > Schedule Workflow)

Allows us to zoom in on our workflow. (Same as View > Zoom In)

Allows us to zoom out on our workflow. (Same as View > Zoom Out)

Allows us to run the workflow. (Same as Tools > Run Workflow) (Hot Key: Ctrl+R)

1.13 Tool Palette



Figure 1-15 - Tool palette

The tool palette is one of the most important aspects of the Alteryx Designer interface. It is the easiest way to bring the tools onto our canvas. In the later section below called *Canvas*, we will discuss the alternative method.

The three main elements of the tool palette are:

- *Search*: Located in the top-left corner of the tool palette, this element allows us to type something into the bar, and every tool associated with that word is displayed in the Tool Selection.
- *Tool Categories*: Located to the right of *Search*, this element allows us to access all of the tools grouped into categories that Alteryx has defined. We have the ability to add to this list by developing macros, which we will discuss toward the end of the book.
- *Tool Selection*: Located at the bottom of the tool palette and shows all tools meeting either the criteria of the *Search* or *Tool Category* that we have selected.

There is one special category in the *Tool Categories* called *Favorites*. It is special because it does not have tools of its own. Instead, we have the ability to add and remove tools from this list by clicking the star at the top-right corner of each tool. When it is yellow, it means that it will show up in the *Favorites Category*. When we cannot see the star unless we point our mouse at the tool, and it shows up grey, that means that it is not in the *Favorites Category*. *Category*.



Figure 1-16 - Favorites

The arrows on either side of the *Tool Categories* offer the ability to scroll across to see each of the categories. We see that, by default, *Favorites* doesn't follow this behavior. That is because it is pinned outside the scroll arrows. We can do this with any of the categories by right-clicking on them and selecting *Pin* <*Category Name*>. See the example below.

Right click on In/Out:



Figure 1-18 - Selecting Pin In/Out

In/Out is pinned:



Select Unpin In/Out:



Figure 1-20 - Tool Categories

We can see that certain icons show up on the *Tool Category* names. That is because each of the categories has a *Default Tool*. If we were to click on the *Tool Category* and drag it to the canvas, the tool in the icon is what we bring onto the canvas.



Figure 1-21 - Tool Category

We can see in Figure 1-21 an option called *Show Deprecated Tools*. Selecting this option shows tools that belong to outdated

versions of Alteryx. If we need a former functionality, we can access it here.

Figure 1-22 - Additional Tool Categories

All of these categories may not be available to all users because Alteryx has many additional packages that we can download in addition to the basic package. One example is if we had not chosen to download the Predictive package when we downloaded Alteryx Designer, we would not see the *Predictive Category*.

We will only cover the basic package in this book, plus some downloaded content.

1.14 Overview

Below is a picture of the canvas and overview window with a simple data stream that is too wide to see what is happening in the view of the canvas. The right side of the image is the *Overview* window, and it shows the view of the canvas in relation to the entire workflow.



Figure 1-23 - Overview Window

In addition to showing where we are, the overview window makes it easy to find the part of our program that we are interested in, because we can move the view of the canvas by clicking and dragging the white box in the overview window.

1.15 Results Window

Every time we run a workflow, we will get a series of returns that lets us know things about the run. One example of this is below.



Figure 1-24 - Results Window

In this run, we can see that we had 0 Errors, 0 Conversion Errors, 0 Warnings, 1 Message, and 2 Files. The counts here reflect the eight notes we see about the run in the text portion of the Output window. We have eight notes instead of the six we would expect from the counts of the messages. This is because the first note is that the workflow started to run, and the last note is that the module had finished the run. Even if we have no issues or messages during the run, we will still have these two notes because they are always written. The last message changes depending on whether or not we have any errors during our execution. The output window is important in diagnosing issues with the data stream.

- *Errors*: These are critical failures that will terminate the data stream they are in.
- *Conversion (Conv) Errors*: These are potentially problematic issues that stem from a data mismatch when converted from one format to another.
- *Warnings*: These are things we should be aware of but could be expected, so the results are printed in the Output for us to know about.
- *Messages*: These are text in the program that are designed to print when certain stages of the data stream have been finished or certain conditions were met.
- *Files*: These are links that will allow us to open any file created during the running of the workflow
- 1.16 Properties Window

The properties window is where we will spend the vast majority of our time. Every tool has a different properties window, and we will discuss each of them as we discuss the individual tools. Here, we will discuss the properties window of the workflow and connections. There is a pane on the left side of every properties window that has a universal set of symbols. We will discuss each of these as they first appear in this book. A list of all of these symbols is in Appendix E.

Workflow - Configuration X Configuration Canvas Workflow Runtime Events Meta Info Window 9 **Canvas** Options Layout Direction: Horizontal • Show -Annotations: Connection Progress: Show Only When Running -Alteryx Help Page Apply Button Figure 1-25 - Workflow Configuration -

Canvas Tab

Workflow configuration

Notice the button that looks like a wrench at the top-left of each of these properties windows. Selecting the wrench shows the *configuration window* so that we can modify the settings of the tool. As we discuss other tools, the importance here will become apparent.

The "?" in the circle opens the Workflow Alteryx Help page.

The " \checkmark " at the bottom-left corner is an apply button. We can click this to apply the changes we have made. We do not need to click this button; we can just click somewhere on the canvas.

The first tab of the *Workflow configuration* window is Canvas. We have the ability to change three things in this window.

- *Layout Direction* is a drop-down list that allows us to rotate our data stream from a horizontal layout, which is the default, to a vertical layout, which is the layout for older versions of Alteryx. We must set either *Horizontal* or *Vertical* and cannot use a combination of the two.
- Annotations is a drop-down list that allows us to change what the text boxes under Tools say. When this option is set to *Hide*, the text does not show. When this option is set to *Show*, the default text is shown for each tool. When this option is set to *Show w/ Tool Names*, the annotations tell us what the tool names are as well as the default text.
- *Connection Progress* is a drop-down list that allows us to change when we see the progress labels on each of the tools. By default, this is set to *Show Only When Running*, but we can also set it to *Hide* and *Show*, which will never or always show the last run's progress respectively.

Canva	s Workflo	W Runtime Events	Meta Info		
Path	N:\Altery	x\12.yxmd			1
Тур	pe			-	1
•	Standard V	Vorkflow			
C	Analytic Ap	p		_	
0	Macros	Standard M	3000	-	
	Macro.	Jotanuaru M	ucro .	_	
Con	stants:	Name	Value	#] ⊕
Cons	stants: Type Engine	Name TempFilePath	Value C:\Users\jnarayanai	# n' []	•
Cons	Type Engine Engine	Name TempFilePath Version	Value C:\Users\jnarayanar 10.1.7.12188	#	•
Cons	Type Engine Engine Engine	Name TempFilePath Version WorkflowDirectory	Value C:\Users\jnarayanai 10.1.7.12188 N:\Alteryx	#	• •

Figure 1-26 - Workflow Configuration -Workflow Tab

The next tab is *Workflow*.

- The *Path* (file location) shows up in this box when the file has been saved.
- The *Type* allows us to choose which format the file should be. This is also where we select the type of *Macro* we want to create. (This will be discussed in *The Mermaid Coffee Company* chapter.)
- The *Constants* section allows us to define variables that we can use throughout our Alteryx Module.

TUT	
•	Canvas Workflow Runtime Events Meta Info
0	Dedicated Sort/Join Memory Usage
	Use Global Default
	C Use Specific Amount: 4081 🛨 Megabytes
	Temporary files
	Use Global Default
	C Use Specific Folder:
	C:\Users\jnarayanan\AppData\Local\Temp\
	Conversion Errors
	Limit Conversion Errors
	Maximum Errors per Location: 10 +
	☐ Stop Processing When Limit is Reached
	Record Limit for All Inputs: No Limit
	Cancel Running Workflow on Error
	Disable All Browse Tools
	Show All Macro Messages
	Disable All Tools that Write Output
	Enable Performance Profiling

Figure 1-27 - Workflow Configuration – Runtime Tab

The *Runtime* tab allows us to configure settings that impact the way that Alteryx runs.

- The *Dedicated Sort/Join Memory Usage* allows us to set the maximum amount of memory that Alteryx is allowed to use for the *Sort* and *Join* functions, which can be highly memory-intensive.
- The *Temporary Files* setting allows us to define the folder that our temporary files are written to.
- The *Conversion Errors* setting allows us to change the way the workflow behaves if it experiences conversion errors, including limiting them from showing up for each tool and stopping the workflow if they occur.
- The *Record for All Inputs* setting allows us to set a maximum number of records read from any of our input files.

- The *Cancel Running Workflow on Error* is useful if we are trying to diagnose a specific issue in a workflow.
- The *Disable All Browse Tools* should be set after we finish testing and the process is in production. We want to do this because browse tools are slow to create and unnecessary unless testing. It is better to use this than to delete all of the tools because it will make modifying the workflow easier later.
- *Show All Macro Messages* allows us to see the messages we are getting in the macro tools so that when we are running a workflow, we can see everything that may be an issue.
- *Disable All Tools that Write Output* is useful for testing because we can turn off only the outputs so we are not rewriting the data every time we run the workflow.



Figure 1-28 - Workflow Configuration -Events Tab

The *Events* tab allows us to define events that will allow secondary actions to take place.

When we add an event, we can choose to define a command line code or send an email when a condition is met. This is useful for security or maintenance because we can set up our workflows to let us know anytime the workflow ran or any time it ran with issues. The *Meta info* tab allows us to define metainformation about our workflow. This is important when we are publishing a workflow to Alteryx Server or the public Alteryx Gallery.

now - Configuration	
Canvas Workflow Runtime Events Meta Info	
Workflow Name	
G Use File Name	
C Custom: 12	
Description	
URL	
Display Text (optional):	
r Tool Settings	
Root Tool Name:	
Tool Version	
[Tool In-Databaser 1	1
Catagony Name	
Category Name.	
Search Tags:	
Author	
Mamar	
Name.	
Company:	
Company:	

Figure 1-29 - Workflow Configuration – Meta Info Tab

- *Workflow Name* allows us to customize what we want the workflow to be called (independent of the file name) so that versioned workflows can have the same name, as far as the end user is concerned.
- The *Description* is a place for us to describe what the workflow does.
- The *URL and Display Test* fields allow us to identify where the module is published.
- The *Author* section allows us to publish information about us and our company along with the workbook.

Connection configuration

When using the default settings, connections are the curved black lines that connect one tool to the next.

٦v	Vireless			
Field	ds	-		-
	Name	Туре	Size	Source
1	F1	V_WString	255	File: N:\TLM Check\Data Sources\Book1.xl

Figure 1-30 - Connection Configuration

The *Connection configuration window* has three components.

- *Name* allows us to rename our data connection. This is most useful when we have multiple data connections going into the same tool connector.
- The *Wireless* check box allows us to make the connection wireless (invisible) unless one of the tools it connects is selected.
- The *Fields* box allows us to see some of the metadata about each field in the data stream at this connection.

1.17 Interface Designer

The Interface Designer is a window used in developing apps and macros. This is where we get to design the user interface the end user will see.



Figure 1-31 - Layout View

The default tab view of the Interface Designer window is of the Layout View, which allows us to design the user interface in a graphical format.



Figure 1-31a - Interface designer settings

The Interface Designer, like the Properties window, has a side bar that tells you what you are looking at. From top to bottom, the icons are:

- *Layout View,* where we design the interface graphically.
- *Test View*, where we see what the end user would see.
- *Tree View,* where we look at the structure of our questions in a tree structure
- *Properties,* where we set properties associated with the interface (this is independent of the Properties window itself).



Figure 1-31b - Interface designer drop down

In the *Layout View*, we design the interface and thus have a significant amount of flexibility in what we can do.

The first and most intuitive thing to do is move questions up and down according to the order we want them asked to the end user. We can do this by clicking the up and down arrows on the right-hand side. We may notice that while we are moving questions, they sometimes move within a box that frames the question we were trying to pass. This is because we have the ability to make one question determine if the other is also asked. By leaving a question inside another, we make that question dependent.

The next thing we can do is select the Add drop-down menu and see that we can add any of four things. Those things are:

- *Group Box*: a text element that we can put questions in.
- *Label*: a text element that we cannot put questions in.
- *Link*: a hyperlink in the interface.
- *Tab*: a tab in the interface.

When we are working with the *Interface Designer*, we will be opening up the Properties windows associated with whichever tool we have selected at the time. The following are the properties windows of the four objects described above.



Figure 1-32 Label Group Configuration

ΞĻ	abel (24) - Configuration	•	p	\times
3	Enter the text or question to be displayed:			
3	Label			
0				
Q				
0				
0				

Figure 1-33 Label Configuration



Figure 1-34 Link Configuration



Figure 1-34 Tab Configuration

As we can see, the four windows are similar. All four ask what we want the element in the interface to be called, and *Link* also asks us what link address we want to add to the interface. In addition to these questions, we see for the first time that we have three new icons in the side bar of the Properties window.



Figure 1-34a Side Bar Icons

- *Navigation*: Allows us to move directly to the incoming or outgoing tools.
- *Annotation*: Allows us to change how this tool is named and how its annotation behaves.
- *Messages*: Allows us to look at the output messages specific to this tool.

Navigation, annotation, and *messages* are part of every tool properties window, and they always behave the same.

Below are examples of a *Check Box* Navigation, Annotation, and Messages window when connected to an *Action* tool.



We can see that *Navigation* lists the tool that this *Check Box* is connected to, Action (20), and tells us that it is an outgoing connection. That means that the Action is downstream of the Check Box. If the Wireless box were checked, the wire would be invisible. We will discuss both of these tools, as well as wireless connections, in chapters *Applications Wanted* and *Meta-morphosis*.



Figure 1-36 Checkbox Annotation

The Annotation window tells us that the tool is a Check Box and that it has the ID number 3 (which means it will try executing this tool third if there are not outside influences). We then have the ability to change the name of the tool. We can change the annotation setting specific to this tool as well as provide a custom annotation (this will show up under the tool in the data stream).



Figure 1-37 Checkbox Messages

The Messages window allows us to look at this tool's specific output messages generated during either the last configuration or last run, so we can investigate any issues while we are looking at the tool.

ao 1 Tao 2 Tao 3	App Values
folder Brokse	Bosse, Com
File Branse	san San
Drop Down Forward 1	
Numerc Up Down 30 (2) Group Question Together Radio Button	Gpen Det
© Radio Button Text Box	

Figure 1-38 Test View

Looking back at the Interface Designer, *Test View* allows us to answer the questions in the same way that the end user would so we can test the app.



Figure 1-39 TreeView

TreeView allows us to look at our user interface as a tree structure so it's easier to make sure our logic is properly grouped.

	On Success - Run Another Analytic App	_
7	On Success - Show Results to User Open Files From These Tools	
	Output Message	
	No Output Files Message	
G	Graphic (Desktop only)	

Figure 1-40 Interface Designer Properties

Interface Designer Properties allows us to customize many aspects of our app or macro.

1.18 Canvas

This is what the default canvas looks like before we have brought any tools onto it.

12.yxmd*	01 Preparing Data yomd	New Workflow3	New Workflow4* ×
~			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			
1			Drop tools here
1			
1			
1			
1			
i l			
1			
i i			
1			1
1			
1			

Figure 1-41 - Default Canvas

We can see three elements: the tab name, a drop-down list of all of the tabs, and a white area that reads "Drop tools here." Let's see what happens if we follow the following steps.

- 1. Open three new windows.
- 2. Click on the third tab.
- 3. Click on Window > New Horizontal Tab Group.
- 4. Click and drag the fourth tab (now third on the top) from the top group to the bottom group.
- 5. Click on Window > New Vertical Tab Group.

We should see that our canvas has split into three sections and looks similar to the image below.



Figure 1-42 - Canvas with elements

The ability to display workflows next to one another makes it much easier to test parts of a program. We can copy and paste portions between canvases to test, compare potential changes, or just work with multiple workflows at the same time.

Moving the locations of window elements



Figure 1-43 - Movement of Canvas Elements

1.19 Using Tools

Data streams start with some sort of data source, and in most cases, this will be the input tool. From there, we may do a series of operations before bringing the data back together to look at.

We will cover each of these methods in detail in the next chapter, but for now, we will look at the structure of a data stream. This Alteryx Workflow takes input data from a file called Test.txt, splits the data stream (doubling the data), adds a new field to one side, and brings the two streams together to be viewed.

Notice that there are two types of arrows on the incoming (left) side of the tools. A green arrow indicates that a single input can be connected there, while a grey double arrow indicates that multiple connections can be made. This is important because it allows us to know how to connect items to each of these tools.

We can see that there is no parallel tool on the outgoing (right) side of the tools, despite the data stream having multiple outgoing connections. This is because every tool allows us to branch the data stream off into many directions. We also see lightning bolt and question anchors, which we will discuss when we first use them in *Applications Wanted*.

There are two ways to bring tools onto the canvas: We can either drag the tool from the Tool Palette or right-click on the canvas and navigate the menu on the next page.

1.20 Insert Tool Menu



Figure 1-44 - Insert Tool Menu

As an alternative to using the tool palette, you can rightclick on the canvas to add tools using the Insert Tool Menu. Simply navigate this menu to find the tool you are looking for
CHAPTER 2 The Games

We all love playing make believe, don't we? As we discussed in the last chapter, assume that you, dear reader, are the new consultant at a company that works with the Alteryx Analytical Platform.

This book is your mentor and here is our first problem to solve!

То
Cc
Subject Welcome - Let's Get Started
Неу,
Welcome Aboard!
We try to get all of our new hires a basic understanding of Alteryx as quickly as possible.
We will center the basic training around the most important sporting event in the world, which should need no introduction.
I will be asking you a few questions and walking you through examples until I feel like you are ready to handle it yourself.
The first question we are going to explore: Which country has produced the best Freestyle Skiing results overall in the 2002 and 2006 Winter Games?
Assume each Gold is worth 3 points, Silver is worth 1.5 points and Bronze is worth 1 point.
Something important to recognize is that I am asking you for the answer to a very specific question. Once you have some of the basics down, we will talk about making a generalized tool for you or your end user to ask related questions. For now, just understand that when you are asked about a specific answer, they are going to only want the result.
I'll show you how this works.
Thanks,

2.1 Tools & Concepts

In this chapter, we will cover the improved features for Alteryx designer 10.x and the various tools and concepts as mentioned below:

Tools	Concepts
Browse	Importing Data
Comment	Viewing Data
Filter	Outputting Data
Formula	Identifying Desired Results
Input Data	Answers to Questions
Join	Tidy Datasets
Output Data	Normalized Datasets
Running Total	Creating Calculations
Sample	Combining Data
Select	Creating Data Subsets
Sort	Summarizing Data
Summarize	Organizing Data
Tool Container	Organizing Workflows
Transpose	Documenting Your Work
Union	

2.2 Improved Features

Several updates were implemented since Alteryx 10.0 was released. Most of the tool icons got a new look, and some tools got new features. New features for each tool will be covered in their respective sections. Here, we will cover the new UI and Alteryx Designer's updated features.

- Reading and writing Excel files (.xlsx) is done using a new Alteryx driver for Microsoft Excel.
- The Predictive Tools Installer has been separated from the Alteryx Designer install to allow for future updates to the Predictive Tools and packages without the need for an update to the Designer.
- After running a workflow, clicking on the Input or Output tools will automatically populate results (up to 1MB).
- Tool palette is customizable to show only those categories and tools you want to use.
- The File menu is reorganized to be simpler and more intuitive to use.
- The enhanced integration between the Designer and Gallery makes it easier to collaborate on workflows hosted in a Gallery. From within the Designer, one can add and maintain connections to a Gallery, open Gallery workflows and edit them, and then save a new version of the workflow back to the Gallery. Previous versions of the workflow are maintained and can be retrieved via a version history window in both the Gallery and the Designer. Additionally, any workflow version can be made the "published" version that users will see by default.
- In-Database support is added for these Data Platforms: Amazon Redshift, Impala, Teradata, and Spark.

- Changes to the Connect In-DB tool include: new streamlined UI, new option to allow password decryption, and new option for creating file-based In-Database connections in order to simplify obtaining connection information from IT Admin/DBA for server users.
- Changes to the Data Stream In tool include: new option to allow password decryption, users can write out to a permanent table when streaming data from an external data source.
- New Macro Input In-DB and Macro Output In-DB tools have been added for In-Database processing so users can build macros with In-Database.
- New Dynamic Input In-DB and Dynamic Output In-DB tools have been added for In-Database processing so users can retrieve the underlying SQL query and other metadata info being sent to the database or to the In-Database tools in a workflow.
- These connectors are now supported: Amazon Redshift bulk load (write), Netsuite (read/write), Qlik (read), PostgreSQL 9.4 (read/write), SAP Hana (read/write), Spark SQL (read).
- Users can now read and write to JSON from the Input and Output Data tools.
- Users can now browse to their Hadoop distributed file system and read/write to HDFS via the new connections in the Input Data tool. (HDFS tools released in 9.5 have been deprecated).
- Salesforce and Marketo connector tools built on the REST API provide enhanced functionality. The existing SOAPbased connectors are now deprecated.

- MongoDB Input and Output tools now support version 3.0. Mongo input will also read from replica set members including primary and secondary.
- Users now have the ability to retrieve a list of sheet names from an Excel (.xlsx) file and read the list via the Alteryx XLSX driver.
- Apache Avro support is no longer listed as being in beta. Support for SQLite has now been upgraded to 3.8.9.
- The Download tool now supports multithreading to increase the speed at which data is retrieved. The maximum number of connections that can be used in the same tool is 32.
- Authenticated proxy can now be enabled from User Settings. The Download Tool, Amazon Redshift Bulk loader, Amazon S3 Upload, and Amazon S3 Download tools will function through a proxy server.
- Alias Repository has been renamed to Alias Manager and supports both standard as well as In-Database connections.
- A new option that provides information about the performance of tools in a workflow is available. Select the "Enable Performance Profiling" checkbox on the Runtime tab of the workflow, and run the workflow to view the percentage of time spent processing each tool in the Results window. This option should only be used when debugging a specific workflow, as it may decrease the performance of the workflow slightly. Sample workflows that demonstrate different configuration options for one tool at a time are one example. Twenty new examples have been added, which can be accessed via Help > Sample Workflows > One Tool Examples. In/Out: Directory; Map

Input Preparation: Auto Field; Generate Rows; Imputation; RecordID; Tile; Unique Join: Append Fields; Find/Replace; Fuzzy Match; Make Group Parse: Date/Time; RegEx; Text To Columns; XML Parse Transform: Count Records; Cross Tab; Running Total; Transpose.

- The Block Until Done tool now has numbered outputs that output "in order."
- The Map tool now has a zoom/bounds option that makes the map zoom/pan to the reference file.
- In the Alteryx Designer version 10.x, the Configure Workflow and Results windows are split into two distinct views.

Wo	kflow - Configuration		→ ₽ X	Results - Mes	sages		▼ ₽	×
8	Canvas Workflow Ru Canvas Options	ntime Events Meta Info		0 Errors	0 Conv Errors	🔒 0 Warnings	0 Messages	**
-	Layout Direction:	Horizontal	•					Â
	Annotations:	Show	-					Ξ
	Connection Progress:	Show Only When Running	•					
0				<			÷	*

Figure 2-1 - Workflow and Results windows

• Configuration and Results windows can be dragged and displaced to different possible positions.



Figure 2-2 – Possible positions

• Every tool in Alteryx now has a Preview feature to show a snapshot of the data it contains.

C	hapter 2 The Gan	nes\Me	dals\All Medal	s.xlsx		
Dp	otions					
	Name		Value			
1	Record Limit					
2	File Format		Microsoft Ex	cel (* xl	sx)	,
3	Table or Query		'Athletes\$'			
4	Search SubDirs					
5	Output File Name	as Fiel	d No			
6	First Row Contains	s Data				_
					11-4-1-	
					Opdate	: 30
Pre	eview		-			
Pre	Athlete	Age	Country	Year	Closing Ceremony Date	S
Pre	Athlete Michael Phelps	Age 23	Country United States	Year 2008	Closing Ceremony Date 2008-08-24	Sp Sh
Pre	Athlete Michael Phelps Michael Phelps	Age 23 19	Country United States United States	Year 2008 2004	Closing Ceremony Date 2008-08-24 2004-08-29	SI SI SI
Pre	Athlete Michael Phelps Michael Phelps Michael Phelps	Age 23 19 27	Country United States United States United States	Year 2008 2004 2012	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12	S S S S
Pre	Athlete Michael Phelps Michael Phelps Michael Phelps Nichael Phelps Natalie Coughlin	Age 23 19 27 25	Country United States United States United States United States	Year 2008 2004 2012 2008	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12 2008-08-24	5 5 5 5 5
Pre 1 2 3 4 5	Athlete Michael Phelps Michael Phelps Michael Phelps Nichael Phelps Natalie Coughlin Aleksey Nemov	Age 23 19 27 25 24	Country United States United States United States United States Russia	Year 2008 2004 2012 2008 2000	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12 2008-08-24 2000-10-01	5 5 5 5 5 5 6

Figure 2-3 - Preview Data

2.3 Browse

F	The <i>Browse</i> of the data it is connect	tool gives us a in a data strean ted.	tabular view n at the point
	Group	Input	Output
Figure 2-4 - Browse	In/Out	Data stream	None

Note: It is very important for the development of workflows, applications, and macros but should be disabled when development is completed to improve speed.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Results Window:

The *Browse Results* Window allows us to view the data that was in the data stream during the last run.

The message option shows the number of records processed, processing time, and any errors during processing.

of 9 Field	ls 🔻 🐦 🛛 Cell Vie	ewer •	1 1 8.	618 recon	ds displayed	, 304 KB				
Record #	Athlete	Age	Country	Year	Closing	Ceremony Date	Sport	Gol	Silver	Bronze
1	Michael Phelps	23	United States	2008	2008-08-2	14	Swimming	0	0	0
2	Michael Phelps	19	United States	2004	2004-08-			6	0	2
3	Michael Phelps	27	United States	2012	2012-08-	Course	Carro	4	2	0
4	Natalie Cooghlin	25	United States	2008	2008-08-	Copy &	Save	1	2	3
5	Aleksey Nemov	- 24	Duratio	2000	2000-10-		cs.	2	1	3
5	Alicia Coutts		f		2012-08-1	12	Swimming	1	3	1
7	Missy Franklin	r	viessages	5	2012-08-1	12	Swimming	4	0	1
в	Ryan Lochte				2012-08-1	2	Swimming	2	2	1
9	Allison Schmitt	22	United States	2012	2012-08-1	2	Swimming	3	1	1
10	Natalie Coughlin	21	United States	2004	2004-08-2	29	Swimming	2	2	1

Figure 2-5 – Browse, Results

The two icons on the right side (shown above) are *Copy* and *Save* functions. They allow us to directly copy the data out of the browse tool or save it to a file.

Athlete Age Country Year Closing Sport	Ceremony Date	Click field r	to see names		
Gold Silver Bronze		Select/D	eselect f	ields	ļ,
Gold Silver Bronze	Athlete	Select/D	eselect f	ields Year	Closing Cerei
Gold Gold Silver Bronze Record #	Athlete Michael Phelps	Select/D	eselect f	ields Year	Closing Cerei 2008-08-24
Gold Silver Bronze Record # 1 2	Athlete Michael Phelps Michael Phelps	Select/D	eselect f	ields Year 2008 2004	Closing Cerear 2008-08-24 2004-08-29

Figure 2-6 - Browse, Field Names

Clicking on the down arrow shown above gets us a list of all field names so we can select only the relevant fields to be displayed. The checkboxes \checkmark in this list allow us to select or deselect every field.

The text shown after clicking on *Cell Viewer* depends on what is selected in the records.

9 of 9 Field Schema I	ds • • Cell Viewer •		8,618 records dis	splayed, 304	KB 📑 🚽
Field #1 Name: At Type: V_ Size: 28	hlete WString 55				
Source: Descript	File: C:\Users\Desk	top\Alteryx	Book 11\Data	a and wor	kflow\
			C 11 1		
Field #2 Name: Aq Type: V_ Size: 28 Source: Descript	ge String 55 File: C:\Users\Desk :ion: 	top\Alteryx	Cell Book II\Data	V1eWe	er :kflow\
Field #2 Name: Aq Type: V_ Size: 28 Source: Descript (Record #	fe String S5 File: C:\Users\Desk :ion: 	atop\Alteryx Age	Cell Book II\Data	V1eW6 a and Woz Year	closing Cerei
Field #2 Name: Aq Type: V_ Size: 25 Source: Descript Record # 1	se String 55 File: C:\Users\Desk tion: ////////////////////////////////////	Alteryx Age 23	Cell Book II\Data	V 1ewe	Closing Cerei
Field #2 Name: Ag Type: V_ Size: 25 Source: Descript Record # 1 2	st St St File: C:\Users\Desk ion:	Alteryx Age 23 19	Cell Book II\Data Country United States United States	V 1 e W e a and Wox Year 2008 2004	2 r :kflow\ Closing Cerei 2008-08-24 2004-08-29
Field #2 Name: Ag Type: V_ Size: 25 Source: Descript Record # 1 2 3	File: C:\Users\Desk File: C:\Users\Desk ion: MichaelPhelps MichaelPhelps MichaelPhelps	Atop\Alteryx Age 23 19 27	Cell Book II\Data Country United States United States United States	V 1ewe a and Wox Year 2008 2004 2012	Closing Ceres 2008-08-24 2004-08-29 2012-08-12

Figure 2-7 – Browse, Cell Viewer

Clicking *Record* # selects everything, allowing us to see the information about each column. If a single column is selected, we see metadata about that column. If a row is selected, we see how much data the record has in it. If a cell is selected, we can see the contents of that cell formatted with all line breaks.

	8,618 records (displa	yed, 9 fields, , 3	804 KB		Table					1	.
9 of 9 Field	🗸 🗸 🗸 🕹 Cell V	iewer	• † ↓	Ľ	• 🖬 •	9 of 9 Fiel	ds 🔻 🖌	Cell V	iewer	▼ ↑ ↓	Ľ	
Record #	Athlete	Age	Country	Year	Closi ^	Record #	Athle	te	Age	Country	Year	Clo: 4
1	Michael Phelps	23	United States	2008	2008-	1	Michael Pl	helps	23	United States	2008	2008
2	Michael Pholps	19	United States	2004	2004-	2	Michael Pl	helps	19	United States	2004	2004
3	Michael Phelps	27	United States	2012	2012-	3	Michael Pl	helps	27	United States	2012	2012
4	Natalie Coughlin	25	United States	2008	2008-	4	Natalie Co	ughlin	25	United States	2008	2008
5	Aleksey Nemov	2-	Dunnin	2000	2000	r	AlakaanAl	omov	24	Russia	2000	2000
6	Alicia Coutts	24						S	24	Australia	2012	2012
7	Missy Franklin	17	Side -b	v-S	ide c	ompai	rison	din	17	United States	2012	2012
8	Ryan Lochte	27		5		1		e	27	United States	2012	2012
9	Allison Schmitt	22	United States	2012	2012-	9	Allison Sc	hmitt	22	United States	2012	2012
10	Natalie Coughlin	21	United States	2004	2004- 🖵	10	Natalie Co	oughlin	21	United States	2004	2004
•					+	•	m					*

Figure 2-8 - Browse, Side by Side

We have the ability to compare the dataset either Side-by-Side (vertical compare) or Top-and-Bottom (horizontal compare) by clicking on the new window button located at the top-right of the results window.

Table	X	ispiaye	d, 9 fields, , 30	4 KB				×
9 of 9 Field	ds 🔻 🗸 Cell Vi	ewer •	• † ↓				•	ŀ
Record #	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Gold	1
1	Michael Phelps	23	United States	2008	2008-08-24	Swimming	8	L
2	Michael Phelps	19	United States	2004	2004-08-29	Swimming	6	
3	Michael Phelps	27	United States	2012	2012-08-12	Swimming	4	
4	Natalie Coughlin	25	United Stat			wimming	1	
c	Alalia au Mamau	24	Duesia	lon	and Bottom	umpactica	2	
٠ 	Aleksev Nemov	24	III	co C	omparison	viillastics	•	
Table	ds 👻 🖌 Cell Vi	ewer 1		c	omparison	Viniastics]
Table 9 of 9 Field Record #	ds 🔹 🗸 🖌 Cell Vi Athlete	ewer •	Country	CC Year	Closing Ceremony Date	Sport	Gold	3
Table 9 of 9 Field Record #	ds 🔹 🖌 📔 Cell Vi Athlete Michael Phelps	ewer • Age 23	Country United States	C(Year 2008	Closing Ceremony Date	Sport Swimming	Gold	3.
Table 9 of 9 Field Record # 1 2	ds 🔹 🖌 Cell Vi Athlete Michael Phelps Michael Phelps	ewer • • • • • • • • • • • • • • • • • • •	Country United States United States	Year 2008 2004	Closing Ceremony Date 2008-08-24 2004-08-29	Sport Swimming Swimming	Gold 8 6]
Table 9 of 9 Field Record # 1 2 3	ds 👻 🖌 Cell Vi Athlete Michael Phelps Michael Phelps Michael Phelps	ewer • • • • • • • • • • • • • • • • • • •	Country United States United States United States	Year 2008 2004 2012	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12	Sport Swimming Swimming Swimming	Gold 8 6 4]
Table 9 of 9 Field Record # 1 2 3 4	ds 👻 Velilov Athlete Michael Phelps Michael Phelps Natalie Coughlin	ewer • Age 23 19 27 25	United States United States United States	Year 2008 2004 2012 2008	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12 2008-08-24	Sport Swimming Swimming Swimming Swimming	Gold 8 6 4]
Table 9 of 9 Field 9 cf 9 Field Record # 1 2 3 4 5	ds Venitov Cell Vi Athlete Michael Phelps Michael Phelps Michael Phelps Michael Phelps Mathie Coughlin Aleksey Nemov	ewer • • • • • • • • • • • • • • • • • • •	Country United States United States United States Russia	Year 2008 2004 2012 2008 2000	Closing Ceremony Date 2008-08-24 2004-08-29 2012-08-12 2008-08-24 2008-08-24 2000-10-01	Sport Swimming Swimming Swimming Swimming Gymnastics	Gold 8 6 4 1 2	

Figure 2-9 – Browse, Top and Bottom

You will also notice that we have an Input button on the Properties window side pane.

Fie	lds				
	Name	Туре	Size	Source	Description
1	Athlete	V_WString	255	File: C:\Users\	
Z	Age	V_String	255	File: C:\Users\	
3	Country	V_String	255	File: C:\Users\	
4	Year	V String			
5	Closing Ceremony Date	Date	Ι	nput wind	ow
6	Sport	V_String		1	
7	Gold	Double	8	File: C:\Users\	
8	Silver	Double	8	File: C:\Users\	
9	Bronze	Double	8	File: C:\Users\	

Figure 2-10 - Browse, Input

This is the *Input* window. It tells us information about each field that comes into the tool: the field name, field type, total size of each cell, the original source of the data, and a description of the field.

2.4 CrossTab

	The <i>Crosstab</i> (more huma creating colu data.	tool creates a n-readable) d umns out of th	normalized ataset by ne rows of
	Group	Input	Output
Figure 2-11 - Crosstab	Transform	Any data stream	Data stream wider than input
<i>Note:</i> The <i>Crosstab</i> characters to under	tool will conv scores in the c	vert all spaces column heade	and special rs.
An <i>Action</i> tool can modify how this to	connect to th ol works in ap	e <i>Lightning B</i> pps and macro	olt Anchor to os.

Properties Window:

The Cross Tab Properties window has four components, as shown in the following figure.



Figure 2-12 - Crosstab, configuration

We notice the grey boxes that have arrows at the top-left and bottom-right corners. These show us what the input and output fields are. These boxes are standard for tools that have input and output.

- *Grouping Fields* allows us to select which fields we want to group by in the resulting dataset. If nothing is selected, we will only get a single record as output.
- *Header Field* is the field we are splitting into multiple columns.
- *Data Field* is the field we want to put in each record for the columns created by the Header Field.

• *Methodologies* (Numeric Data Field) allows us to select the type of aggregation method used if we have multiple data entries that fit into the same cell of the resulting dataset.

2.5 Comment

	The <i>Comment</i> too to write notes on add additional in data stream.	l gives us th our workfl formation o	ne ability ows to on the
	Group	Input	Output
Figure 2-13 - Comment	Documentation	None	None
<i>Note: Comment</i> is an an	notative tool to h	elp give n	neaning to

Properties Window:

The *Comment* Properties window allows us to customize the comment field that appears on the canvas.

Ċ.	Text		
	Notes are important for readability		
	* Shape		
	Rounded Rectangle		
	Font		
	Book Antiqua, 12pt, Style =Bold, Italic 🛄		
	Text Color		
	Black		
	Background Color		
	White		
	Text Alignment	1	Notes are important for
	TopCenter 👻 🗕		readability
	Background Image		

Figure 2-14 – Comment, configuration

The configuration settings in the figure above create the comment to the right of the window. We can edit text, shape, font and color of the comment background, adjust alignment of the text, or select an image to write over.

Using the settings in the *Comment* tool allows us to create easily recognizable, distinct comments throughout our data stream.

2.6 Filter



The *Filter* tool gives us the ability to create a function that will split the data row by row into either the true or false outputs.

Group	Input	Output
Preparation	Any data	T & F
	stream	section
		below

Note: The formula must evaluate to a True or False Boolean value (null evaluates to False unless the formula is looking for nulls). The formulas we create here can be arbitrarily complex, and thus can significantly slow our data stream. "//" is the comment character.

Application questions can be connected to the *Top Black Question Anchor* to use those answers in this tool.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output T: This is the set of records where the formula evaluated to true.

Output F: This is the set of records where the formula evaluated to false.

A basic filter was added to the Filter tool. One can use this basic filter to quickly construct a simple query on a single field in the incoming data stream.

Properties Window:

We see that there are two different ways to create a filter: *Basic Filter* and *Custom Filter*. The regions for each are as shown below.



Figure 2-16 – Filter, basic filter

• *Basic Filter*: Allows us to pick a field and operator, and type in the value that the field should be compared to.

The options for the *Basic Filter* change depending on the field type. The *Basic Filter* option allows us to easily create simple filters as well as start building formulas before we know the syntax.

As we enter values into the *Basic Filter* options, it populates the Expression box at the bottom of the window. This helps us learn the associated syntax.

• *Custom Filter:* allows us to click on variables, function, and saved expressions to populate the expressions window or

type the formula directly. A sample custom filter is shown below.



Figure 2-17 – Filter, custom filter

We will discuss creating formulas in the *Formula* tool and throughout the exercises in this book. On an error, the red Error symbol replaces the message symbol, as shown in the figure below.





2.7 Formula

0 3	The <i>Formula</i> t create a funct a new column	tool gives us ion that will n in our data	the ability to be written to a.
	Group	Input	Output
	Preparation	Any data	Augmented
		stream	original
Figure 2-19 - Formula			data stream

Note: The formulas created here can be arbitrarily complex, and thus can significantly slow the data stream. Ensure the created output field has a file type that's compatible with the result being created. We can use the formulas created higher in the list in calculations lower in that list. "//" is the comment character.

Application questions can be connected to the *Top Black Question Anchor* to use those answers in this tool.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: The original data stream with one additional field for each formula we create.

A new function is added to the Formula tool: Starts With, Ends With, and Contains.

Properties Window:

The formula window looks similar to the filter window; in fact, the expression building section here is identical to the *Custom Filter*.

The top of the *Formula Properties* window has an option to define multiple calculations, change the order the fields are created, and define the major metadata for the field that we are creating.



Figure 2-20 – Formula, options

The Gar	nes
---------	-----

The arrows on the side move the created formula up and down the list while the circle with the line in it removes the highlighted formula.

Each of these output fields will have an associated formula and will add a column to our data stream.

Under the *Variables Tab*, we can see *Fields* and *Constants*. The incoming data determines what list of Fields, and the *Environmental Variables* determine what the Constants are. (*Environmental Variables* can be defined in the workflow properties window; see Properties Window in Chapter 1 for more information.)

Under the *Functions tab*, we see a tree structure. This allows us to look for the functions needed by double-clicking on and moving them into the expressions window to work with.

The *Saved Expressions Tab* allows us to access recent and saved expressions, as well as save our current expression for later use.

2.8 Input Data



The *Input Data* tool gives us the ability to import data from specific databases.

Group	Input	Output
In/Out	None	Data stream
		in initial
		data format

Note: This is the most common start to a data stream. We can use full or relative file paths to files as well as database connections.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

A new option is available for relational database connections. When enabled, the data is stored in a yxdb file on disk so that data sources aren't hit repeatedly during workflow development.

Properties Window:

The *Input Data Property* window has three main components. The first is the field that shows our data connection. When connected to a data source, we can see the address of the file or database that we are connected to.

Alte	ryx Book\Workflow\Ch	apte	er 2 The Game	s\Meda	als\All Medals.xlsx	-	
Opti	ions						
	Name		Value			^ E	
1	Record Limit						Dropdown Opti
2	File Format	1	Microsoft Excel ("xdsx)	-		1 1
3	Table or Query	1	'Athletes\$'				File Browse
4	Search SubDirs			[11]		_	Database Connection
5	Output File Name as Fie	d	No		•		32 Bit Database Connection
6	First Row Contains Data			1		-	HDFS Connection
_		_				_	Alias
Prev	/iew				Update Sam	ple	Pallos
	Athlete	Age	Country	Year	Closing Ceremony	*	
	Michael Phelps	23	United States	2008	2008-08-24		
1		19	United States	2004	2004-08-29		
1 2	Michael Phelps			2012	2012-08-12		
1 2 3	Michael Phelps Michael Phelps	27	United States	LUIL			
1 2 3 4	Michael Phelps Michael Phelps Natalie Coughlin	27 25	United States United States	2008	2008-08-24		
1 2 3 4 5	Michael Phelps Michael Phelps Natalie Coughlin Aleksey Nemov	27 25 24	United States United States Russia	2008	2008-08-24 2000-10-01		
1 2 3 4 5 6	Michael Phelps Michael Phelps Natalie Coughlin Aleksey Nemov Alicia Coutts	27 25 24 24	United States United States Russia Australia	2008 2000 2012	2008-08-24 2000-10-01 2012-08-12		
1 2 3 4 5 6 7	Michael Phelps Michael Phelps Natalie Coughlin Aleksey Nemov Alicia Coutts Missy Franklin	27 25 24 24 17	United States United States Russia Australia United States	2008 2000 2012 2012	2008-08-24 2000-10-01 2012-08-12 2012-08-12		

Figure 2-22 – Input Data, options

When we click on the drop-down arrow on the right-hand side of the field, we see this menu and can use this option to connect to data sources that are accessible. If present, a list of recent connections is shown, allowing us to save connections. This would appear under the Alias Link shown above. The second main component of the window is the Options section where we can change the setting associated with the data connection to modify exactly what we are connecting to. This will allow us to modify many of the options that define the connection.

The third component is the Preview, which gives us a view of the first 100 records to help ensure that we are connected to the correct data source.

Table or Query: This option needs to be called out separately because it allows us to open a new connection window by clicking on the "…" button. That menu looks similar to this, depending on what we are connecting to.

The image below describes each of the tabs available when clicking on the " \dots " button.



Figure 2-23 - Input Data, table and query options

2.9	Join
	,



The <i>Join</i> tool gives us the ability to
combine two data streams by lining
up records based on matching fields.

Input	Output
See Input Left	See Output
and Input	Left, Output
Right	Join, and
	Output Right
	See Input Left and Input Right

Note: Join does not work like a join in SQL; the tool creates three groups in order to perform a left, right, or full outer join We need a *Union* tool after the join to identify which outputs should be brought together. (See example "Brains vs. Brawns.") If we have multiple fields that match each other, records are replicated from the original data stream.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input Left: A data stream with at least one common field to Input R (the fields do not need to share a name).

Input Right: A data stream with at least one common field to Input L (the fields do not need to share a name).

Output Left: Data stream containing records that did not match anything from the right input.

Output Join: Data stream containing records that match both left and right inputs. Records may be replicated as a result of this operation. *Output Right:* Data stream containing records that did not match anything from the left input.

Properties Window:

The Join Properties window has two major components.

The top asks how we want to join the two data sets. We can join by position or by specific fields. Most of the time, we would be joining on specific fields because it allows greater control over the join.



Figure 2-25 – Join Properties

The bottom allows us to define which fields will be in the output as well as define the metadata for those fields (more on this when we talk about the *Select* tool). We can see that there are three images separating the top and bottom of the window. These Venn diagrams show us what will be in each of the three outputs. More succinctly, we can consider the image below. The two inputs are the pink and blue circles, and the three outputs are the pink, purple, and blue shaded regions.



Figure 2-26 – Join Properties Venn Representation

2.10 Output Data



The Output Data tool allows us to				
write the data stream out to a file or				
database				
Group	Input	Output		
-	-	-		

Any data

stream

File Or

Database

Figure 2-27- Output Data

Note: The output window has the ability to write to files or to databases using SQL.

In/Out

An Action tool can connect to the Lightning Bolt Anchor to modify how this tool works in apps and macros.

Properties Window:

The top half of the Output Data Property window is very similar to the Input Data window. Both allow us to navigate to a file or database and set options related to the dataset.

u	t Data (7) - Configuration
	Write to File or Database
1	Options
	Take File/Table Name From Field
	Append Suffix to File/Table Name
	Field Containing File Name or Part of File Name
	Athlete

Figure 2-28 – Output Data Properties

The difference is that at the bottom, there are some special options that allow us to modify the way the metadata is written based on the incoming data stream.

We can see here that there is a file format called an "Alteryx database (*.yxdb)." Alteryx allows us to store data in files specifically designed to work well with Alteryx. These file types are native to Input and Output Data tools, so they do not need to perform any conversion to connect to these files. We can also use this file type to store both spatial and non-spatial data together.

2.11 Running Total

5	The <i>Running Total</i> tool allows us to create a running sum for a numeric field in the incoming data stream		
	Group	Input	Output
	Transform	See	See
		details	details
Figure 2-29 – Running Total		below	below

Note: Running Total produces the running sum of the data from the top of the column down, so it is important to make sure the data is properly ordered (See the Sort Tool).

Input: Any data stream with at least one numeric field.

Output: The original data stream with additional columns called *RunTot_<Original Field Name>* for each of the selected "Create Running Total" fields.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Properties Window:

The Running Total Properties window has two components.

Runn	ing Total (8) - Configuration	×
	ing Total (8) - Configuration Group By (Optional) Athlete Age Country Year Closing Ceremony Date Create Running Total Gold Silver Bronze Right_Gold Right_Silver Right_Bronze	All Clear All Clear
0		

Figure 2-30 – Running Total Properties

The first is *Group By*, which allows us to define a field or set of fields that the running sum will be unique to the set of elements in our group by fields.

The second is the selection of which fields we want to create a running total on.

2.12 Sample

The <i>Sample</i> tool allows us to work with a subset of data		
Group	Input	Output^

5	Preparation	Any data	See details
		stream	below
Figure 2-31 - Sample			

Note: This is useful for limiting the amount of data we've run through our data stream when we are testing, creating different samples of our dataset, and skipping header or footer information that may exist in our data.

Output: The original data stream with potentially modified metadata and truncated fields.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Properties Window:

The Sample Properties window has three different settings.

Samp	le (9) - Configuration	×
~~	First N Records	
3	Last N Records	
Ð	Skip 1st N Records	
	1 of every N Records	
0	Random 1 in N Chance for each Record	
Å	First N% of Records	
0	N = 100	
	Grouping Fields (Optional)	
	Sport	All
	V Age	Clear
		Ciedi
Ø		

Figure 2-32 – Running Total Properties

The first of these settings is the type of sampling that is needed.

First N Records: the first N records in our data stream.

Last N Records: the last N records in our data stream.

Skip 1st *N Records*: all but the first N records in our data stream.

1 of every N Records: create groups of records based on the order, and take one of each of those records.

Random 1 in N Chance for each Record: Every record has a 1/N chance of being kept.

First N% of Records: the first N percent of Records in our data stream.

The second settings is 'N' to be used in the sampling.

The third is the ability to select the fields we want to group the sampling by. In the scenario pictured above, this filtering will keep the first 100 records for each date in the data.

2.13 Select

Figure 2-33 - Select	The <i>Select</i> tool allows us to modify metadata associated with the data stream, including the order of columns		
	Group	Input	Output
	Preparation	Any data stream	See details below

Note: Select is used after every data connection and periodically throughout the data stream to ensure everything in the data stream is in the right format, named appropriately, and necessary. Use this tool to drop fields that are no longer needed to save space.

Output: The original data stream with potentially fewer fields with modified metadata.

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.
Properties Window:

The *Select Properties* window has multiple important tasks associated with the maintenance of data and metadata. The window shown below shows us a list of every field name coming in so we can reset the information, if needed.



Figure 2-34 – Select Configuration

In this case, the *Date* field has been renamed to *Sales Date*, and the *Sales* field was converted to a float type and given a description saying that it is the *Total Cash Sales*. The red cells above indicate that the metadata in the cell has been modified.

There are two other things we can do with the fields without going into the options menu: The first is checking and unchecking the boxes at the left of each row in order to drop that field from the data stream. The second is reordering the fields so that the columns are in a different order downstream. This can be accomplished by selecting a field and clicking the up and down arrows or by right-clicking in the space to the left of the checkmark and dragging the field up or down in the list. In addition to the two fields that are part of the incoming data stream, there is a special field called **Unknown* that acts like a placeholder for all new fields that come into this tool.

Select Options Menu:

The options menu allows us to systematically modify the fields.



Figure 2-35 - Select Options 1

Save/Load allows us to create or load a Field Type File (.yxft). Field Type Files are a metadata file that can be used to appropriately define or redefine the columns in our data stream.

Save Field Configuration creates a new yxft file.

Load Field Names imports field names from an existing yxft file.

Load Field Names & Types imports field names and the type of field that should be allocated.

Other options are as shown in the figure below.

Save/Load Select Change Field Type of Highligh	hted Fiel	bescription Sele ds Desc	ct All elect All	
Sort	•	Sort on Field Name Sort on New Field Name Sort on Field Type Revert To Incoming Field) I Order	Ascending Descending
Sort on Field Name Sort on New Field Name	•			
Sort on Field Type Revert To Incoming Field Order		Ascending Descending		
Move	•	Move Highlighted Fields	to Top	1

Figure 2-36 - Select Options 2

Select allows us to select or deselect all fields.

Sort has four primary methods of ordering fields.

Sort on Original Field Name will alphabetically sort our fields in either ascending or descending order based on the field names that came in from the data stream.

Sort on New Field Name will alphabetically sort our fields in either ascending or descending order based on the field names that leave the tool.

Sort on Field Type will group all fields that have the same data type together.

Revert To Incoming Field Order will clear the ordering of fields.

If we have selected fields, Move allows us to group them all at the top of the data field list.

Add Prefix to Field Names will allow us to add a prefix to all fields or all selected fields.

Add Suffix to Field Names will allow us to add a suffix to all fields or all selected fields.

Remove Prefix will allow us to remove a common prefix between selected fields.

Remove Suffix will allow us to remove a common suffix between selected fields.

Clear All Renames will remove all renaming that has been defined for this select.

Clear Highlighted Renames will remove all renaming in the highlighted (selected) fields.

Revert All To Original Type & Size will remove all changes to the field types or allocated data sizes.

Revert Highlighted To Original Type & Size will remove all changes to the field types or allocated data sizes for highlighted fields.

Forget all Missing Fields will remove this tool's metadata about fields that are no longer coming into the tool from the data stream.

Forget Highlighted Missing Fields will remove this tool's metadata about fields that are no longer coming into the tool from the data stream that are highlighted.

Type: Type is an important thing to know because each have different attributes and mean different things to other tools. Each of the types are described in Appendix D.

2.14 Sort

5	The <i>Sort</i> tool records.	allows us	to reorder
	Group	Input	Output
Figure 2-37 - Sort	Preparation	Any data stream	See below

Note: Sorting is most important when we are working with calculations that consider multiple rows or ordering data for normalized consumption.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: The original data stream with records sorted in a different order.

Properties Window:

The *Sort Properties Configuration* window allows us to select one or more fields by name, and Ascending or Descending for each to determine an order to our records. We can change the order of these sorts by moving them up and down the list.

	Name	Order		
•	Year	▼ Ascendi	ng 🔻	
*		-	-	

Figure 2-38 – Sort Configuration

By checking the *Use Dictionary Order* option, we can select the dictionary order that should be applied to sort the data when appropriate.

2.15 Summarize

5	The <i>Summarize</i> tool allows us to summarize data in our data stream.				
	Group	Input	Output		
	Transform	See below	See below		
Figure 2-39 - Summarize					

Note: When summarizing data, it may be necessary to reexamine the underlying calculations because aggregating those calculations may not make sense. Summarizing a single field using Group By is a good way to get a unique list of the data. Running complex analysis like geocoding is often more efficient to do on a summarized list and then join back onto the full dataset.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Any data stream with too granular a level of detail.

Output: A summarized data stream with a less granular level of detail.

Properties Window:

The Summarize Configuration window has two basic components.

	Field	Туре			
	Athlete	V_WString			
	Age	V_String			
	Country	V_String			
	Year	V_String			
	Closing Ceremo	Date			
	Sport	V_String			
Þ	Gold	Double			
	Silver	Double			
	Bronze	Double			
Action	is: Field	Add ·	•	Output Field Name	
Action	15: Field Closing Ceremo	Add Action GroupBy	•	Output Field Name Date	
Action	15: Field Closing Ceremo Gold	Add Action GroupBy Sum	•	Output Field Name Date Sum_Gold	

Figure 2-40 – Summarize Configuration

The *Fields* list shows each of the incoming fields, and the *Actions* list shows the fields created in this tool.

The select button at the top-right allows us to select tools in a systematic way so that if we want to take the sum of all of our Numeric fields, we can select them all and add them in one step.

When we have something in the fields list selected, we can click on the *Add* button, which shows the drop-down menu shown above. It lists every operation that can be used to aggregate the data using *Summarize*.

We will not be going through this menu in detail, as it is a list of aggregation methods; however, *Group By* needs to be given special attention.

When we use *Group By* on fields in the summary, we will end up with one line item for each combination of field elements that we grouped by. If this is unclear, it should make more sense as we go through exercises.

2.16 Tool Container

	The <i>Tool Container</i> tool allows u group tools together for clarity a allows the tools to be disabled when unnecessary.		ows us to arity and bled
	Group	Input	Output
Figure 2-41 – Tool Container	Documentation	None	None

Note: Click and drag tools onto the box to put them into the tool container. Tool containers make it much easier to navigate our data stream because they allow us to consider a series of tools as a single unit. If we click on the arrow at the top-right corner, it will collapse the box without disabling it.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Properties Window:

The *Tool Container Configuration* window allows us to customize the text and formatting of the container as well as disable the tools inside it.

*	Caption Tool Container make	s your data stream easy to read and test	ID 4		
(V) (I)		Res	et to Dercults		
0	Text Color	R=49, G=76, B=74			
-	Fill Color	R=236, G=242, B=242			
	Border Color	R=49, G=76, B=74		Tool Container makes your data stream easy to read and test	۲
	Transparency	25	×.		
	Margin	Medium	•	•	

Figure 2-42 – Tool Container Configuration

The *Disabled* option allows us to turn off sections of our data stream. This is typically used in testing data and application building.

2.17 Transpose

F	The <i>Transpose</i> tool allows us to de normalize data		us to de-
	Group	Input	Output
Figure 2-43- Transpose	Transform	See below	See below

Note: This tool converts all spaces and special characters in the titles into underscores in each record.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Any data stream with multiple fields that need to be combined into rows.

Output: A taller data stream because records have been duplicated in order to consolidate the columns. The columns consolidate into two columns; *Name*, which is a column with each of the former column names in it, and *Value*, which is a column with each of the data values.

Properties Window:

The *Transpose Configuration* window has three elements, as shown in the figure below.



Figure 2-44 – Transpose Configuration

Key Fields allows us to select the fields that will be maintained after the transposition.

Data Fields allows us to select the fields that will be combined following the transposition.

The drop-down at the bottom allows us to change the message behavior if there are missing fields in the incoming data stream.

2.18 Union



The *Union* tool allows us to append records together one after another from multiple data sources.

Group	Input	Output
Join	See below	See below

Note: The order that we connect data stream to this tool's input will determine the default order that they are combined. In this tool, naming the incoming connections is often helpful.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Multiple data streams that should be combined by adding the records from one set to the end of the others.

Output: A data stream that has the records from multiple data streams.

Properties Window:

The Union Configuration window has three core elements.

The first drop-down allows us to change the method that is used to align the columns from the different inputs.

Auto Config by Name	
Properties	
When Fields Differ	
Warning - Continue Processing Records	
Output All Fields	
Output Order	
Set a Specific Output Order	
#1	

Figure 2-46 - Union Configuration by Name

Auto Config by Name makes the union align the fields that have the same name. This is best used if we know that our data will always be named the same way.

ц.	Union (7) - Configuration	
5	Auto Config by Position	•
Ð	Properties	
	When Fields Differ	
-	Warning - Continue Processing Records	▼
0	Output All Fields	-
Ą		
0		

Figure 2-47 - Union Configuration by Position

Auto Config by Position makes the union align the fields by the column number. This is best used if we know our data will always be in the same order but may have different (or no) field names. In both of these options, we see the same Properties section asking what should happen when the fields differ. The first drop-down allows us to change the behavior between an error, a warning, and nothing. The second drop-down allows us to decide if we want all fields or only the fields matched from all outputs to be in the output.



Figure 2-48 – Union Configuration Manual Configuration

Manually Configure Fields is the third option in the first drop-down. It allows us to select exactly which fields to be

brought together by manipulating the Output Columns portion of the configuration window. This is best used if we know that our fields may be named differently or may be in different order.

The final component is the *Output Order*. It allows us to set the order of the records by choosing the order that the data streams are combined.

2.19 Freestyle

То
Сс
Subject Welcome - Let's Get Started
Неу,
Welcome Aboard!
We try to get all of our new hires a basic understanding of Alteryx as quickly as possible.
We will center the basic training around the most important sporting event in the world, which should need no introduction.
I will be asking you a few questions and walking you through examples until I feel like you are ready to handle it yourself.
The first question we are going to explore: Which country has produced the best Freestyle Skiing results overall in the 2002 and 2006 Winter Games?
Assume each Gold is worth 3 points, Silver is worth 1.5 points and Bronze is worth 1 point.
Something important to recognize is that I am asking you for the answer to a very specific question. Once you have some of the basics down, we will talk about making a generalized tool for you or your end user to ask related questions. For now, just understand that when you are asked about a specific answer, they are going to only want the result.
I'll show you how this works.
Thanks,

Let's start building a workflow that will answer our question. We are going to start with a blank canvas and save it as Freestyle Skiing. Next, bring an Input Data tool so that we can connect to data.



Figure 2-49 – Freestyle Skiing, Data Input

Now navigate to where data files are unpackaged, and connect to the file *All Medals.xlsx* in *Chapter 2 -- The Games> Medals*. For downloading the data associated with this book, please refer to the letter to the reader on this chapter's first page.

When connected, we see this window pop up. Click on *Athletes*\$ then OK to connect to the Athletes sheet in the All Medals Excel file. This is shown in the figure above.

Best practices are to put a select and a browse after every input.

• Browse helps us check the data at the time of import. This ensures that the data we are getting is correct.

• Select allows us to make sure that the fields are in the right format from the beginning.



Figure 2-50 - Freestyle Skiing, Select configuration

If we click on Select, we should see that our fields are in different types than the above image. Change them to match what is shown.

Now that we have the data and the fields are the right type, the first thing we should do is filter the data. We always want to *limit the data as soon as possible*, since this will speed up our data stream and prevent memory errors by limiting the information.

Best practice is to remove data as soon as it is no longer needed.

It makes sense that the first step in filtering would be to bring in the Filter tool; however, if we are not familiar with the data set and we have not run it, we may not have enough information to filter properly. In this case, we want to run the module so that there's data in the Browse tool for us to work with.

a A	Itery	k Des	signer x64 - Freestyle.y	xmd*					Alternative Statistical
File	9	Ec	lit <u>V</u> iew <u>O</u> p	tions I	Help	2			
C i	7		» 同 合 ら で	C Q	e	0)			
Searc	h All	Tool	s	Q- +	Favo	Intes	In/Ou	t Preparation	n 📕 Join 🍈 Parse 📕 Transform 📲 In-Datab
Bro	owse	*	Input Data Output D	Data Text	Inpu	🖈 i t	Filter	Formula Samp	De Select Sort Join U
Anev	ver ve	ersio	n of Alteryx Designer x6	4 is available	e: [Click	here for option	▼	LICK KUN to
0	oloci	+ (2)	- Configuration					po	pulate Browse
	00	tion		TID. T.				at sinks all als and d	
*	Op	uons	•• I ↓	TIP: TO	reor	der mul	tiple rows: see	ect, right-click and d	
()			Field	Туре		Size	Rename	Description	
•	Þ	V	Athlete	V_WString	-	255			
-		V	Age	V_String	•	255			
0		V	Country	V_String	-	255			
		V	Year	V_String	•	255			
0		1	Closing Ceremony Date	Date	-	10			
~		V	Sport	V_String	-	255			
		V	Gold	Double	-	8			
			Silver	Double	-	8			All Medals.xlsx
			Bronze	Double	•	8			Table=`Athletes\$`
		V	*Unknown	Unknown	-	0		Dynamic or Unkn	
0			1,					A S. III III	

Figure 2-51 – Freestyle Skiing, populate Browse

We can see that when the module finishes running, we get a pop-up window that lets us know how long it took to run and if there were errors.



Figure 2-52 – Freestyle Skiing, message after Running

Feel free to click on the *Don't show this message again* check box before closing if the pop-up window is distracting.

Now, we can start thinking about the filter. We know that we are only interested in freestyle skiing results for the 2002 and 2006 games. So the first thing we are going to filter is the sport to "freestyle skiing." If we look at the browse tool, we first see freestyle skiing at row 5818, identified by the string *Freestyle Skiing*.

	9 of 9 Field	ls 🔻 🖋 🛛 Cell View	ver 🔻	1 4 8,618 re	cords displa	ayed, 304 KB				• 🔒 • 🛛
1	Record #	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Gold	Silver	Bronze
-	5815	Irina Slutskaya	23	Russia	2002	2002-02-24	Figure Skating	0	1	0
	5816	Aleksey Yagudin	21	Russia	2002	2002-02-24	Figure Skating	1	0	0
	5817	Zhao Hongbo	28	China	2002	2002-02-24	Figure Skating	0	0	1
	5818	Shannon Bahrke	29	United States	2010	2010-02-28	Freestyle Skiing	D	0	1
	5819	Dale Begg-Smith	25	Australia	2010	2010-02-28	Freestyly Skiing	0	1	0
	5820	Hedda Berntsen	33	Norway	2010	2010-02-28	Freestyle Skiing	0	1	0
	5821	Alexandre Bilodeau	22	Canada	2010	2010-02-28	Freestyle Skiing	1	0	0
	5822	Aleksey Grishin	30	Belarus	2010	2010.02.29	Era abila Shiina	1	0	0
	5823	Audun Grønvold	33	Norway	Freestyle Skiing starts from here					
	5824	Guo Xinxin	26	China	1100	by it braing b		i nei		1
	5825	Jennifer Heil	26	Canada	2010	2010-02-28	Freestyle Skiing	0	1	0
	5826	Lydia Ierodiacono	28	Australia	2010	2010-02-28	Freestyle Skiing	1	0	0
	5827	Marion Josserand	23	France	2010	2010-02-28	Freestyle Skiing	0	0	1
	5828	Hannah Kearney	23	United States	2010	2010-02-28	Freestyle Skiing	1	0	0

Figure 2-53 – Freestyle Skiing, Browse configurations

This is the crucial piece of information we didn't have before. Now that we know exactly what we need to look for in our data so we can create the filter.

Drag a Filter tool after the Select, and make sure there is a connection between the Select output and the Filter input. This time, we will use the Basic Filter builder. Set the field drop-down to *Sport*, and type *Freestyle Skiing* into the text box like we see below.



Figure 54 - Freestyle Skiing, Browse configurations

Notice that the Expression says **[Sport] = "Freestyle Skiing."** This is because field names are in brackets, and string values are in quotes. What is happening here is that for each record, we test to see if the value in Sport in exactly *Freestyle Skiing*. If it is, then True; if it isn't, then False.

The next thing we want to do is create a filter to keep only 2002 and 2006. If we look at the Select tool on the previous page, we will see that the *Year* field is a string. This is fine; we need to remember that when we are writing the filter formula. Drag a new Filter tool onto the canvas, and make sure that the first Filter's true (T) output is connected to the new Filter's input.



Figure 2-55 - Freestyle Skiing, Filter configurations - Functions

This time, we are going to create the filter logic on our own using the Custom Filter option. We know from the previous filter that [**Field Name>] = "<Value>"** is the syntax for filtering a string field, so creating the first half of the filter is not hard – it is [**Year**] = "2002" – but what we need to do now is make sure 2006 is also kept.

One way we could do this is by using logical operators. *Logical operators* are terms that allow us to combine two or more Boolean (true or false) values to create a single Boolean from the two. The three Boolean operators that we will be discussing are *and*, *or*, and *not*.

AND: if **both** the value to the **left and** the value to the **right are true, then true**.

OR: if **either** the value to the **left or** the values to the **right or both** are **true**, **then true**.

NOT: **if** a value is **true** then **false**; **if** the value is **false**, then **true**.

We can see in the *Functions Tab* that we have the option to use the *Boolean OR - Keyword* or the *Boolean OR* ||. There is no computational difference between using the keyword or the double vertical bar symbol. Both options are available for our convenience. For those not used to programming, the keyword *OR* is much easier to remember and use, but for those who program, double vertical bars (||) is a common standard they may be used to.



Figure 2-56 - Freestyle Skiing, Custom filter

See Appendix F for examples of Boolean logic.

Now that we know about logical operators, we can finally finish putting the filter together. We can use the formula **[Year] =** "2002" OR **[Year] =** "2006" in order to filter out this data.

Note: We could have combined both of these filters together by using the following: **[Sport] = "Freestyle Skiing" AND ([Year] = "2002" OR [Year] = "2006")**

The parentheses allow us to change the order of operations so that this formula reads "Freestyle Skiing in the years 2002 or 2006" instead of "Freestyle Skiing in 2002 or anything in 2006."

We now have removed all of the information we don't need in order to answer the question. But we have the problem that the data is too granular. We know who the athlete was and in which year they won their medal(s). We should bring in a Summarize Tool in order to bring the data up to the country level. If we add Country using Group By, and Gold, Silver, and Bronze using Sum, we will get a list of countries and their total medal count for Freestyle skiing for 2002 and 2006. Place a Browse Tool, and run the module to see what we have so far.



Figure 2-57 – Freestyle Skiing, Summary configuration

Best practice is to place a browse tool after every tool that transforms data into a significantly different shape. Summarize is one of the tools.

4 of 4 Field	ds 🔻 🖌 🛛 Cell V	iewer 🔻 🗍		• 🛨 •	
Record #	Country	Sum_Gold	Sum_Silver	Sum_Bronze	
1	Australia	2	0	1	
2	Belarus	0	1	1	
3	Canada	1	1	1	
4	China	1	1	0	
5	Czech Republic	1	0	0	
6	Finland	1	1	0	
7	France	0	0	2	
8	Japan	0	0	1	
9	Norway	1	1	0	
10	Russia	0	0	1	
11	Switzerland	1	0	0	
12	United States	0	3	1	

Figure 2-58 – Freestyle Skiing, Browse configuration after summarize

We can see that we have four columns with the total counts of gold, silver, and bronze medals listed for each of the 12 countries that won freestyle skiing medals during 2002 and 2006.

Notice the fields are titled *Sum_* followed by the original field name. Alteryx is making sure we know the method used to summarize the data.

The next thing we need to do is determine which country was the *best*. If we look back at the email, we can see that *best* is defined as a function of the medals won; 3 points for gold, 1.5 points for silver and 1 point for bronze.

Bring a Formula Tool onto the canvas following the Summarize tool, and we are going to create a calculation called Score that has the Type Double, with the formula [Score] = 3*[Sum_Gold] + 1.5*[Sum-Silver] + [Sum_Bronze]



Figure 2-59 – Freestyle Skiing, Formula configuration

We can now add another Browse after the Formula Tool to see what the data looks like.

5 of 5 Field	ds 🔻 🖌 🛛 Cell	Viewer • 1	1		• 🕣 •
Record #	Country	Sum_Gold	Sum_Silver	Sum_Bronze	Score
1	Australia	2	0	1	7
2	Belarus	0	1	1	2.5
3	Canada	1	1	1	5.5
4	China	1	1	0	4.5
5	Czech Republic	1	0	0	3
6	Finland	1	1	0	4.5
7	France	0	0	2	2
8	Japan	0	0	1	1
9	Norway	1	1	0	4.5
10	Russia	0	0	1	1
11	Switzerland	1	0	0	3
12	United States	0	3	1	5.5

Figure 2-60 - Freestyle Skiing, Browse configuration after formula

We see there is a new field called Score that is an unordered dataset and with multiple unnecessary values. We can also see that Australia has the highest score and therefore is the answer to the original question. But for good practice, we are going to continue to build this workflow so that no interpretation is needed.

This process is going to take four steps:

- 1. Reorder the data based on the score field.
- 2. Select only the top-scoring country.
- 3. Remove all data other than the name of the best country.
- 4. Browse that data.

Like we discussed, the Sort Tool is how we reorder the data and will be our first step. We will set up our data in a descending order based on Score, like we see here.



Figure 2-61 – Freestyle Skiing, Sort configuration

Next, we need only the first record, so we are going to use the Sample Tool to keep only the Top 1 Record coming out of the Sort.

G Sample (9) - Configuration	
 First N Records 	_
C Last N Records	
Skip 1st N Records	2 <u>2 </u>
I of every N Records	
Random 1 in N Chance for each Record	
First N% of Records	
0 N = 1	
Grouping Fields (Optional)	Score - First 1
Country	All Descending
Sum_Gold	Class
Sum_Bronze	Clear
Score	
0	

Figure 2-62 – Freestyle Skiing, Sample configuration

We know we have fields we no longer need, so we can use a Select Tool to eliminate everything that is not the country name.

	Held	Type	-	Size	Rename	Description	
	Country Sum Cold	V_String	•	255			ITEN .
	Sum_Gold	Double	+	8			
	Sum_Bronze	Double	-	8			
	Score	Double	-	8			STI
[*Uoknown	Upknown	-	0		Dynamic or Unk	1

Figure 2-63 – Freestyle Skiing, Selection after Sample

Finally, we can put a browse tool at the end and run the workflow to see the results.

Res	ults - Browse (3)	- ₽ >	Freestyle.yxmd* ×	
	1 of 1 Fields 🔻 🖌 🛛 Cell Viewer 👻 🗍 🗍		>	
Ä	Record # Country	*		·····
	1 Australia		-1000	-1
	*			

Figure 2-64 - Freestyle Skiing, Browse after Select

We could have stopped when we first saw Australia had the highest score in the previous browse tool. The reason we did not, is that when we are performing an analysis, we want our results to be perfectly repeatable. If we had interpreted the previous browse tool incorrectly then there would be no way of finding out why an error occurred. This is a problem because it makes the individual analyst entirely responsible for the answer, and anyone who checked the results could easily find the correct answer where we mistakenly picked the wrong one. Finishing the workflow in this way affords us two benefits:

- We would have a second verification that the answer was what we expected.
- Repeatability of the result so we can point to a single issue in the data preparation process that needs fixing instead of not being able to fix it at all.

Here is how the workflow would look like when complete:



Figure 2-65 – Freestyle Skiing, Data stream after completion

2.20 Let S Huy Hungs Op

-	From +	
Send	To	Alteryx Consultants
Sena	Сс	
	Subject	Let's Tidy Things Up
Неу,		
That wa	s great!	
So the n just an a	ext thing we a answer.	re going to cover is a question that requires you to produce a dataset instead of
The goa formats The first "human are repl names t sets are think of	I of most data that are most t and more rel friendly". The icated because that we will be e categorized b it as "compute	manipulation is to get the data in a more useable format. Typically, there are two t appropriate. Which you create is going to depend on what you are trying to do. atable is to have a wide <i>normalized</i> data structure, which you can think of as use datasets tend to have multiple columns that have the same metrics in them but e you have a variable you want to compare across. The other has two common e using interchangeably; "tidy" is one and "denormalized" is the other. These data by having a single field for each variable and are often very "tall" (long). You can er friendly".
Let's co we have convert downst	nsider the dat a mostly den ing the three o ream systems	a source on the historical medal counts that we were just working on. In this case, ormalized dataset. Let's take that last step in creating a truly tidy dataset by columns gold, silver, and bronze into "Medal Type" and "Medal Count" so that can process the data better.
Thanks,		

This process must include at least four steps:

- 1) Import the dataset.
- 2) Transpose the dataset.
- 3) Make sure the fields are named correctly.
- 4) Export the dataset.

However, we are going to make the data cleaner and employ best practices. So our process is:

- 1) Import the data.
- 2) Browse the data.
- 3) Make sure the data has the right type.
- 4) Transpose the data.
- 5) Browse the restructured data.
- 6) Make sure the fields are named correctly.
- 7) Remove records that say there was no medals won.
- 8) Browse the data that will be exported.
- 9) Export the dataset.

Let's create a new workflow and save it as *Let's Tidy Things Up*.

We need to import the same data that we used in the last example. Bring an input tool onto the canvas, navigate to where we saved this book's data, and connect to the file *All Medals.xlsx* in *Chapter 2 -- The Games> Medals*.



Figure 2-66 - Let's Tidy Things Up, Data Input

Now we will put a Browse and select statement following the Input Tool.



Figure 2-67 – Medals data

All of the fields are in appropriate types for what we are trying to do, so we can move directly to the transposition.

Let's run the Module to see how the data is structured.

	9 of 9 Field	is 🔻 🖌 🛛 Cell Vi	ewer	• † ↓	8,618 re	cords displayed, 304 KB			L	- 🛛 - 🖸
	Record #	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Gold	Silver	Bronze
1	1	Michael Phelps	23	United States	2008	2008-08-24	Swimming	8	0	0
1	2	Michael Phelps	19	United States	2004	2004-08-29	Swimming	6	0	2
	3	Michael Phelps	27	United States	2012	2012-08-12	Swimming	4	2	0
4	4	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	1	2	3
1	5	Aleksey Nemov	24	Russia	2000	2000-10-01	Gymnastics	2	1	3
(5	Alicia Coutts	24	Australia	2012	2012-08-12	Swimming	1	3	1
1	7	Missy Franklin	17	United States	2012	2012-08-12	Swimming	4	0	1
8	B	Ryan Lochte	27	United States	2012	2012-08-12	Swimming	2	2	1
9	9	Allison Schmitt	22	United States	2012	2012-08-12	Swimming	3	1	1
1	10	Natalie Coughlin	21	United States	2004	2004-08-29	Swimming	2	2	1
1	11	Ian Thorpe	17	Australia	2000	2000-10-01	Swimming	3	2	0
1	12	Dara Torres	33	United States	2000	2000-10-01	Swimming	2	0	3
		ar. 1. 14	i			·				۰.

Figure 2-68 - Let's Tidy Things Up, Browse configuration

Th	ie	Ga	m	es
I	ie	Ga	m	es

The Transpose tool takes normalized data and denormalizes it. If we take the data stream coming out of *Select* and pass it into a Transpose, we can make the data tidier.

Adhlete Age Gountry Year Country Year Closing Ceremony Date Sport Gold	All Clear	All Medals.xlsx
Athlete Age Country Year Closing Ceremony Date Sport Ø Gold Ø Silver Ø Bronze Ø Dynamic or Unknown Fields	All	Table = Athletes

Figure 2-69 - Let's Tidy Things Up, Transpose configuration

We want to keep all of the fields as they are except for gold, silver and bronze. So we select all but those three fields under Key Fields and we select gold, silver and bronze under the Data Fields. If we had wanted to drop a field entirely – say, Closing Ceremony Date – we could leave it unchecked in both lists.

Best practice is to always include a Browse after a tool that modifies the structure of a data stream. Transpose is one of these tools. Let's add a Browse to the end of the data stream and run it to see what we have.

Record # Athlete Age Country Year Closing Ceremony Date Sport Name Value 1 Michael Phelps 23 United States 2008 2008-08-24 Swimming Gold 8 2 Michael Phelps 23 United States 2008 2008-08-24 Swimming Silver 0 3 Michael Phelps 23 United States 2008 2008-08-24 Swimming Bronze 0 4 Michael Phelps 19 United States 2004 2004-08-29 Swimming Sold 6 5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 6 Michael Phelps 19 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 20		8 of 8 Field	ds 🔻 🖌 🛛 Cell Vi	ewer 🔻	↑ ↓ 25.	854 reco	ords displayed, 478 KB			B • B	+
1 Michael Phelps 23 United States 2008 2008-08-24 Swimming Gold 8 2 Michael Phelps 23 United States 2008 2008-08-24 Swimming Silver 0 3 Michael Phelps 23 United States 2008 2008-08-24 Swimming Bronze 0 4 Michael Phelps 19 United States 2004 2004-08-29 Swimming Gold 6 5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 0 6 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 7 Michael Phelps 19 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 9 Michael Phelps 27 United States 2012 <th></th> <th>Record #</th> <th>Athlete</th> <th>Age</th> <th>Country</th> <th>Year</th> <th>Closing Ceremony Date</th> <th>Sport</th> <th>Name</th> <th>Value</th> <th></th>		Record #	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Name	Value	
2 Michael Phelps 23 United States 2008 2008-08-24 Swimming Silver 0 3 Michael Phelps 23 United States 2008 2008-08-24 Swimming Bronze 0 4 Michael Phelps 19 United States 2004 2004-08-29 Swimming Gold 6 5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Slover 0 6 Michael Phelps 19 United States 2012 2012-08-12 Swimming Bronze 2 7 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 201	1	1	Michael Phelps	23	United States	2008	2008-08-24	Swimming	Gold	8	
3 Michael Phelps 23 United States 2008 2008-08-24 Swimming Bronze 0 4 Michael Phelps 19 United States 2004 2004-08-29 Swimming Gold 6 5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Silver 0 6 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 7 Michael Phelps 19 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 4 11 Natalie Coughlin 25 United States 2		2	Michael Phelps	23	United States	2008	2008-08-24	Swimming	Silver	0	
4 Michael Phelps 19 United States 2004 2004-08-29 Swimming Gold 6 5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Silver 0 6 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 7 Michael Phelps 27 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 21 Natalie Coughlin 25 United States <t< td=""><td></td><td>3</td><td>Michael Phelps</td><td>23</td><td>United States</td><td>2008</td><td>2008-08-24</td><td>Swimming</td><td>Bronze</td><td>0</td><td></td></t<>		3	Michael Phelps	23	United States	2008	2008-08-24	Swimming	Bronze	0	
5 Michael Phelps 19 United States 2004 2004-08-29 Swimming Silver 0 6 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 7 Michael Phelps 27 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3 12 Natalie Coughlin 25 United States		4	Michael Phelps	19	United States	2004	2004-08-29	Swimming	Gold	6	
6 Michael Phelps 19 United States 2004 2004-08-29 Swimming Bronze 2 7 Michael Phelps 27 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		5	Michael Phelps	19	United States	2004	2004-08-29	Swimming	Silver	0	
7 Michael Phelps 27 United States 2012 2012-08-12 Swimming Gold 4 8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		6	Michael Phelps	19	United States	2004	2004-08-29	Swimming	Bronze	2	
8 Michael Phelps 27 United States 2012 2012-08-12 Swimming Silver 2 9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2		7	Michael Phelps	27	United States	2012	2012-08-12	Swimming	Gold	4	
9 Michael Phelps 27 United States 2012 2012-08-12 Swimming Bronze 0 10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		8	Michael Phelps	27	United States	2012	2012-08-12	Swimming	Silver	2	
10 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Gold 1 11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		9	Michael Phelps	27	United States	2012	2012-08-12	Swimming	Bronze	0	
11 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Silver 2 12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		10	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Gold	1	
12 Natalie Coughlin 25 United States 2008 2008-08-24 Swimming Bronze 3		11	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Silver	2	
		12	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Bronze	3	

Figure 2-70 - Let's Tidy Things Up, Browse configuration

If we compare the top three records from the new Browse to the one that came out of the Input in figure 2-66, we see that we have two fields called *Name* and *Value* and no longer have the fields *Gold*, *Silver*, and *Bronze*. We also notice from *Athlete* to *Sport*, all fields are identical to the first three records in the original dataset. This is because we replicated them for each column we created.

This is one of the reasons that tidy data is not particularly human readable but is highly computer readable. Since all of the information is displayed in each record and there is only a single column to work on, interactive front-end software can work very fast with the data.

		Field	Туре		Size	Rename	Description
Þ		Athlete	V_WString	•	255		
	V	Age	V_String	•	255		
	V	Country	V_String	•	255		
		Year	V_String	•	255		
		Closing Ceremony Date	Date	•	10		
		Sport	V_String	•	255		
	V	Name	String	•	6	Metal Type	
		Value	Double	•	8	Metal Count	
		*Unknown	Unknown	-	0		Dynamic

Figure 2-71 - Let's Tidy Things Up, Select configuration

Making this data truly tidy would mean we need to rename *Name* and *Value* to names that give better context to the field. Add a Select statement, and rename the Name and Value fields *Medal Type* and *Medal Count*, respectively.

We know we have rows that say zero medals were won by looking at the values in the last Browse tool we created. We are going to filter those data points out by adding a Filter tool after the Select.

Our goal is to filter out any records that have zero medals. We are filtering on a numeric field for the first time, which means we should use the Basic Filter to learn about the syntax. The configuration is as shown in the following figure.

[Pick Field] ~ [=	•	
Custom	Filter		
Variables	Functions	Saved Expressions	
Expression			

Figure 2-72 - Let's Tidy Things Up, Filter configuration

We can see how if we select Medal count, we have different options in the operator drop-down. This is because numeric fields allow different comparison methods than string fields.

We want to select greater than - ">" - and type "0" in the text box. When we look at the Expression below, we see that it says **[Medal Count]** > 0. This is because we do not put numeric values in quotes. Alteryx recognized that when we selected a numeric field in the basic filter drop-down, the "0" we typed in meant the number 0 and not the string 0, so it put the numeric value into the formula.

The last step involves two tools: the *Browse* tool and the *Output Data* tool.

Best practice dictates that we put a Browse before every data output so that we do not need to open the file to make sure we created it correctly.
We now add a Browse tool to the end of the data stream and also add an Output Data tool. We are going to write the file to the same folder we have saved the *Let's Tidy Things Up.yxmd*.

	11.1		
lame	Value		Specify
lax Records Per File			Location and
ile Format	Comma-Delimited Text F	les (* 🔻	File name
elimiters	•		The name
irst Row Contains Field Names	v		
Juote Output Fields	Auto	-	
ode Page	ISO 8859-1 Latin I	•	
ine Ending Style	Windows	-	
	lax Records Per File ile Format lelimiters irst Row Contains Field Names tuote Output Fields Code Page ine Ending Style	Iax Records Per File ile Format Comma-Delimited Text Fi lelimiters . irst Row Contains Field Names Nuote Output Fields Auto Code Page ISO 8859-1 Latin I ine Ending Style Windows	Iax Records Per File ile Format Comma-Delimited Text. Files (* ▼ lelimiters . irst Row Contains Field Names Nuote Output Fields Auto ▼ Code Page ISO 8859-1 Latin I ▼ ine Ending Style Windows ▼

Figure 2-73 - Let's Tidy Things Up, Output Data configuration

To do this, we are going to type .*Tidy Medal Data.csv* in the text box labeled *Write to File or Database*.

We just used a relative file path. Which allows us to reference files in relation to where we currently are. Some basics of relative paths are ".\", which means the current folder. "..\" means the parent folder (the folder that our current folder is in). ".\Folder Name\" will move our file into a folder below where we have the workflow.

We do not necessarily need to use relative paths, but if we are sharing Alteryx files, it is very beneficial to do so. We can use absolute paths (full file locations) by pasting them into this box or navigating to them in the File Browse option.

If we run the module, we can see what the transformed dataset looks like. This ensures that the information written into the .csv was correct.

8 of 8 Field	ds 🔻 🖌 🛛 Cell Vi	ewer 🔻	1 1 9.1	07 record	ds displayed, 287 KB			D• 🗗 🖸
Record #	Athlete	Age	Country	Year	Closing Ceremony Date	Sport	Metal Type	Metal Count
1	Michael Phelps	23	United States	2008	2008-08-24	Swimming	Gold	8
2	Michael Phelps	19	United States	2004	2004-08-29	Swimming	Gold	6
3	Michael Phelps	19	United States	2004	2004-08-29	Swimming	Bronze	2
4	Michael Phelps	27	United States	2012	2012-08-12	Swimming	Gold	4
5	Michael Phelps	27	United States	2012	2012-08-12	Swimming	Silver	2
6	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Gold	1
7	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Silver	2
8	Natalie Coughlin	25	United States	2008	2008-08-24	Swimming	Bronze	3
9	Aleksey Nemov	24	Russia	2000	2000-10-01	Gymnastics	Gold	2
10	Aleksey Nemov	24	Russia	2000	2000-10-01	Gymnastics	Silver	1
11	Aleksey Nemov	24	Russia	2000	2000-10-01	Gymnastics	Bronze	3
12	Alicia Coutts	24	Australia	2012	2012-08-12	Swimming	Gold	1

Figure 2-74 – Let's Tidy Things Up, Final Browse configuration

Here is how the *Let's Tidy Things Up* data stream looks on completion.



Figure 2-75 - Let's Tidy Things Up, Data stream after completion

То	Alteryx Consultants							
Subject	Modern History							
Great!								
Now that y create a no	you are getting the sense of tidy data, let's go in the opposite direction and prmalized dataset.							
Now that y reate a no low abour each year i nedal cou	you are getting the sense of tidy data, let's go in the opposite direction and ormalized dataset. It we create a nice table with countries alphabetically on the left, a column for in the dataset ordered from longest ago to most recent, and a historical total in the cross section?							

Notice that there is considerably less context built into this email. We often get very sparse information from people, and they will assume we know the context. In this case, it was assumed we were talking about the medal data that we have been working with during the training so far.

This is a much more complicated process than the last exercise, but that is only because the data was set up very well for what we were doing last time, and it isn't here.

We are going to be connecting to the same data source that we have been using, but we are going to use a shortcut in the connection process. Open a new workflow and save it as *Modern History.yxmd*, but make sure that *Let's Tidy Things Up.yxmd* is still open.

Click on the data input in *Let's Tidy Things Up* and copy it. Move over to the Modern History canvas, and paste what we have copied. We see that the input has been copied over and we do not need to recreate the connection.



Figure 2-76 Modern history, Input Data

Best practice will once again bring in Browse and Select tools. But since we know from past experience what the data looks like and how it is read in, we will move directly into the next step.



Figure 2-77 - Modern history, Summarize configuration

We know that this data is too granular for our desired result. So we can summarize it. Based on the email, we know the only information we will need in the end is the country, year, and something to do with the medals. So when using the Summary tool, we can group by the country and year fields and take the sum of each of the medal counts to take our first step down this path.

Output Field Type	Size	e Expression	1		
1 Total Metals - Double	▼ 8	[Sum_Gold]			
▪ Double	- 8		+		
Variables Functions Saved Exp		ł	Θ		
				All Medals.xlsx Table=`Athletes\$`	Total Metals= [Sum_Gold] +
Expression: [Sum_Gold] + [Sum_Silver]	+ [Sum_Br	ronze]		All Medals.xlsx Table=`Athletes\$`	Total Metals= [Sum_Gold] + [Sum_Silver] + [Sum_Bronze]

Figure 2-78- Modern history, Formula configuration

We now add a formula that creates a *Total Medal* count by adding the gold, silver and bronze fields for each record. (Remember that we used a Summarize tool so we should have a Browse tool.)



Figure 2-79 – Modern History Select Configuration

We can now add a select statement that will allow us to keep only the Country, Year, and Total Medals fields, which we will use to create the table.

We know we need a historical medal count, which means we are going to need to take the running total along with the country and year. But because Running Total is a tool where order matters, we need to sort the data.

We can sort the Country and Year in ascending order to help us in two places: Initially, this will help because we are creating the order for the Running Total, but it will also help us with the order of records and columns when we normalize the data set.



Figure 2-80 – Modern History Sort Configuration

Now that we have ordered the data, we can create the Running Total for each country across years. To do this, we *Group By* Country and *Create Running Total* on Total Medals. What this will do is create the running sum of Total Medals down the data set (as time increases) and have that count restart every time a new country shows up.



Total Configuration

Let's take a look at what we have created so we can get a better sense of what the process so far has done.

The (Games
-------	-------

iii e	Browse (9)	- Configura	ation	▼ 🖓	X Modern History
Res	ults - Brov	vse (9)			
	4 of 4 Field	ds 🔻 🖌 🛛 C	ell Viewer	• 1 1 40 ⁻	l records displayed, 7764 by
Ä	Record #	Country	Year	Total Metals	RunTot_Total Metals
	1	Afghanistan	2008	1	1
	2	Afghanistan	2012	1	2
	3	Algeria	2000	5	5
	4	Algeria	2008	2	7
	5	Algeria	2012	1	8
	6	Argentina	2000	20	20
	7	Argentina	2004	49	69
	8	Argentina	2008	51	120
	9	Argentina	2012	21	141
	10	Armenia	2000	1	1
	11	Armenia	2008	6	7

Figure 2-82 – Modern History Browse Configuration after Total

If we add a Browse tool and run the workflow, we can see that we have an alphabetical list of countries with a record for every year they won a medal. We can also see the year is increasing as we move down the list within a country. We then see the Total Medal count for that year and the running total for medals that the country has won going from one year to the next in a field called RunTot_Total Medals

<u></u> c	ross Tab (6) - Configuration	- ₽	×	Modern History.yxmd* ×
	Grouping Fields			
	Country Year Total Metals RunTot_Total Metals			
Ø	Header Field			
	Year		-	
	Data Field			
	RunTot_Total Metals		-	
	Methodologies (Numeric Data Field)			
	Sum Average Count (without Nulls)		-	
	Count (with Nulls) Percent Row Recent Column		-	

Figure 2-83 - Modern History Cross Tab Configuration

The next step in this process is to convert the data into a Cross Tab. If we add the Cross Tab tool to the end of the data stream and apply the settings in the above image, we will be close to our goal.

(1) E	Browse (7)	- Configuratio	n	- ₽ X	Modern History.y	xmd* ×
Res	ults - Brov	vse (7)				
	8 of 8 Field	ds 🔻 🛹 🛛 Cell V	iewer 🕶 🕇	110 recor	rds displayed, 5181 byte	s
Ä	Record #	Country	2000	2002	2004	2006
	1	Afghanistan	[Null]	[Null]	[Null]	[Null]
	2	Algeria	5	[Null]	[Null]	[Null]
	3	Argentina	20	[Null]	69	[Null]
	4	Armenia	1	[Null]	[Null]	[Null]
	5	Australia	183	185	341	343
	6	Austria	4	24	32	62
	7	Azerbaijan	3	[Null]	8	[Null]
	8	Bahamas	11	[Null]	13	[Null]
	9	Bahrain	[Null]	[Null]	[Null]	[Null]
	10	Barbados	1	[Null]	[Null]	[Null]

Let's add a browse tool and see what we have so far.

Figure 2-84 – Modern History Browse Configuration Cross Tab

The result in the image is close but not exactly what we wanted. We get the correct running totals in the years that each country won medals, however we get nulls in the years that they did not.

What we need to do now is create a series of formulas that replace the nulls with zero or the previous value as appropriate. Since we need to create formulas. we are going to need the formula tool; but this time, we are going to need to create seven similar calculations because we need to replace the values in seven different fields.

Let's think through these formulas. We want to change the cell only if it is null. If the column we are fixing is 2000, then it should be replaced with 0, and if the column is not 2000, it should be replaced with whatever is in the previous fixed column.

For those familiar with conditional statements, the syntax for an if-then statement is:

IF b1 THEN x ELSEIF b2 THEN y ELSE z ENDIF

For those unfamiliar with conditional statements, the concept is: Given a true or false (Boolean) expression, the calculation should do one of two things. The logic is if something is true, then do that; else, if the previous is false and something else is true, do the second option; else, do the default.

The other thing we need to know in writing these formulas is the test to see if something is null. The function used is:

IsNull(x)

Both of these syntax are under the functions tab in the Formula tool if we need to reference them.

	Output Field	Ту	/pe	Size	Expression
1	2000 Fixed	Do	ouble	- 8	IF IsNull([2000]) THEN 0 ELSE [2000] ENDIF
2	2002 Fixed	Do	ouble	- 8	IF IsNull([2002]) THEN [2000 Fixed] ELSE [2002] ENDIF
3	2004 Fixed	Do	ouble	- 8	IF IsNull([2004]) THEN [2002 Fixed] ELSE [2004] ENDIF
4	2006 Fixed	- Do	ouble	- 8	IF IsNull([2006]) THEN [2004 Fixed] ELSE [2006] ENDIF
5	2008 Fixed	Do	ouble	- 8	IF IsNull([2008]) THEN [2006 Fixed] ELSE [2008] ENDIF
6	2010 Fixed	Do	ouble	- 8	IF IsNull([2010]) THEN [2008 Fixed] ELSE [2010] ENDIF
7	2012 Fixed	Do	ouble	- 8	IF IsNull([2012]) THEN [2010 Fixed] ELSE [2012] ENDIF
•		- Do	ouble	- 8	

The formulae that we need are:

Figure 2-85 - Modern History Formula List

Add a Formula tool to the end of data stream, and add the seven formulas we see here with corresponding field names. We can also add a browse tool after that to see what we have created. We can see that we have two sets of fields: those with the original sparse data, and those with the new dense data.

browse (1	s) - Configuratio	'n						• • • •	Mode	m History.yom	d" ×					
sults - Brow	wse (13)															• 4
15 of 15 Fi	ields 🔹 🕪 🛛 Cell	Viewer • 1	L 110 re	cords displa	ryed, 8246 bytes											Ø - 🖬 - 🖬
Record #	Country	2000	2002		2004	2006	2008	2010	2012	2000 Fixed	2002 Fixed	2004 Fixed	2006 Fixed	2008 Fixed	2010 Fixed	2012 Fixed
1	Afghanistan	[Null]	[[Null]	[Null]		[Null]	1	[Null]	2	0	0	0	0	1	1	2
2	Algeria	5	[Null]	[Null]		[Null]	7	[Null]	8	5	5	5	5	7	7	8
3	Argentina	20	[Null]	69		[Null]	120	[Null]	141	20	20	69	69	120	120	141
4	Armenia	1	[Null]	[Null]		[Null]	7	[Null]	10	1	1	1	1	7	7	10
5	Australia	183	185	341		343	492	495	609	183	185	341	343	492	495	609
6	Austria	4	24	32		62	65	91	[Null]	4	24	32	62	65	91	91
7	Azerbaijan	3	[Null]	8		[Null]	15	[Null]	25	3	3	8	8	15	15	25
8	Bahamas	11	[Null]	13		(Null)	20	(Null)	24	11	11	13	13	20	20	24
9	Bahrain	[Null]	[Null]	(Null)		(Null)	[Null]	(Null)	1	0	0	0	0	0	0	1
10	Bachados		Chi-113	(h)-113		Chi. 113	Children 11	(NoII)	Data (0.3	1	1	1	1	1	1	1

Figure 2-86 - Modern History Formula List

The next thing that we need to do is remove and rename the columns that we have, so add a Select tool to the end of the data stream.



Figure 2-87 – Modern Select Configuration after Formula

Now the data has finished being prepped. We need to write it out, which we know because we were asked for a data set and not a specific answer. We should add a Browse tool and an Output Data tool to end the data flow. Save the output as Historical Medal Count.csv.



Figure 2-88 - Modern History Output

After doing these steps, the final workflow is as shown in the below figure.



Figure 2-89 – Modern History Data Stream When Complete

2.22 Brains vs Brawns

	Altryx Consultants
Subject	Brains vs Brawns
Awesome	L. C.
We only h know.	ave one more basic skill to go over before we test to see how much you
Combining	g data.
I have bee it compare	n working with the medals dataset for a while and it is interesting to see how es to different metrics.
I think we countries.	should compare the medal counts to Nobel Laureates from each of the
Let's put t and Nobe to the cou	ogether the data to see what the relationship between the count of medals Laureates was since 2000. (We will map the country of Nobel Laureate birth intry that won the medal).

E.

The	Games
-----	-------

Since we are combining data, let's revisit the analogy presented in the preface. When we look at a river, we see there are tributaries all along its length. Each of these tributaries may have gone through different terrain and could have started as very different sources. When they come together, they add whatever they carried along with them into the river they form.

To relate it to the task at hand, tributaries are branches of our data stream that come together, and when they come together, we have a richer data stream because we have the information that comes from everything contributing to it.

We are going to start by prepping the medals data and preparing them to be joined. We'll create a table with two columns called Country and Medal Count.

In order to do this, we are going to follow the following steps:

- 1. Import data
- 2. Transpose and rename the columns so that the data is tidy.
- 3. Filter out the 0 medal records
- 4. Summarize the data so that we only have one record per country and the total medal count.
- 5. Rename the medal count column *Total Medal Count*.

Since we have covered the tools and the concepts used in this exercise in previous exercises, overall flow should look familiar. Please rebuild the following workflow with the following configurations.

The properties windows for each of these tools as well as the data stream that is produced are shown in the following figures.



Figure 2-90 – Brains vs Brawns data stream

		Field A	Туре		Size	Rename	Description
►		Age	V_String	-	255		
		Athlete	V_WString	-	255		
	V	Bronze	Double	-	8		
	V	Closing Ceremony Date	Date	-	10		
		Country	V_String	-	255		
		Gold	Double	-	8		
		Silver	Double	-	8		
		Sport	V_String	-	255		
	V	Year	V_String	-	255		
		*Unknown	Unknown	-	0		Dynamic o.



ranspose (3) - Configuration	▼ 무
Key Fields	
Age	▲ All
Athlete	
Bronze	_ Clear
Closing Ceremony Date	
Gold	
Silver	
Data Fields	
Age	All
Athlete	
V Bronze	Clear
Closing Ceremony Date	
Country	
V Gold	
Snot	
Sport Vear	

Figure 2-92 – Initial Steps - Transpose

			•			The Toriet	ider manuple rows. select, righ
		Field	Ту	pe	Size	Rename	Description
•		Age	V	-	255		
		Athlete	V	-	255		
		Closing	D	•	10		
		Country	٧	•	255		
	V	Sport	٧	-	255		
	V	Year	٧	-	255		
		Name	S	-	6	Metal Type	
		Value	D	-	8	Metal Count	
		*Unkno	U	-	0		Dynamic or Unknown Fields

Figure 2-93 – Initial Steps - Select

Racio Eil	tor		
O Basic Fil	ter		
[Pick Field]		▼ = ▼	
Custom F	Filter		
Variables	Functions	Saved Expressions	
11			
Expression:			

Figure 2-94 – Initial Steps - Filter

	Tielu	Туре	
Þ	Age	V_String	
	Athlete	V_WString	
	Closing Ceremo	Date	
	Country	V_String	
	Sport	V_String	
ction	Tield	Add	Output Field Name

Figure 2-95 – Initial Steps - Summarize

	Ор	tions	- + +	TI	IP: T	o reor	der multiple rows: sele	ect, right-click and drag.
>			Field	Туре		Size	Rename	Description
L			Country	V_String	-	255		
9			Sum_Metal Count	Double	•	8	Total Olympic Metals	
		1	*Unknown	Unknown	-	0		Dynamic or Unknown Fields

Figure 2-96 - Initial Steps - Sort

Se	elect	: (1	5) - Configu	Iration				- 무
1	Ор	tions	- 1 1		TI	P: To rea	order multiple rows: sele	ect, right-click and drag.
;			Field	Туре		Size	Rename	Description
	Þ		Birth Country	V_WString	•	254		
			Count	Int64	•	8	Total Nobel Laurates	
			*Unknown	Unknown		0		Dynamic or Unknown Fields

Figure 2-97 - Initial Steps - Select

Now that we have the data in the above stream prepared to be combined, we should prepare the other contributing data stream.

Let us open the file called *Nobel Laureates.csv* in the folder Chapter 2 -- The Games > Nobel Laureates. (Remember that we should always bring in a Browse and Select Tool with an input.)



Figure 2-98 - Running unrelated analysis simultaneously

Notice that we now have two completely separate workflows. This ability is often a useful feature because we can run unrelated analyses at the same time, which aids in testing and in conditional application development.

If we run the workflow, we can look at the structure of the Nobel Laureates dataset. Here, we want to make sure that the field we plan on joining (Birth Country) is in the same type as County in the medal data stream.

(1) E	Browse (29)) - Configurat	ion	1	• 7 ×	Brains vs Braw	ms.yxmd* ×	
Res	ults - Brov	vse (29)						
=	6 of 6 Field	is 🔻 🖌 🛛 Cell V	′iewer ▾│ † ↓	943 records	displayed, 35	KB		
Ä	Record #	Birth Country	Category	Country		Name	Nobel Laureats	Year
	1	Germany	Austria Chemistr	Germany	Richard K	uhn	1	1938
	2	Germany	Chemistr	Germany	Fritz Hab	er	1	1918
	3	Germany	Chemistr	Germany	Carl Boso	h	1	1931
	4	Germany	Chemistr	Germany	Otto Hah	n	1	1944
	5	Argentina	Chemistry	Argentina	Luis Fede	rico Leloir	1	1970
	6	Australia	Chemistry	Australia	John War	cup Cornforth	1	1975
	7	Austria	Chemistry	Austria	Friderik P	regl	1	1923
	8	Austria	Chemistry	Austria	Richard A	dolfZsigmondy*	1	1925
	9	Austria	Chemistry	Austria	Richard K	uhn*	1	1938
	10	Austria	Chemistry	Austria	Max F. Pe	rutz	1	1962

Figure 2-99 - Nobel Laureates - Browse Configuration

		Field	Туре		Size	Rename	Descrip
Þ		Birth Country	V_WString	-	254		
		Category	V_WString	•	254		
	V	Country	V_WString	•	254		
	V	Name	V_WString	•	254		
	V	Nobel Laureats	V_WString	•	254		
	V	Year	Double	•	8		
	V	*Unknown	Unknown	-	0		Dynamic

Figure 2-100 – Nobel Laureates – Select Configuration

Now that we know what the data structure is and that it parallels the medals file, we can start our preparation for the join.



Figure 2-101 – Nobel Laureates – Preparing to join Medals

We know we want to limit this data to years starting in 2000. One way we can do this is to convert Year to a Double Type and set up a filter to be **[Year]** >= 2000.

Since we only need to know what the total number of Nobel Laureates there were for each country of birth, we can summarize the data.

2 of 2 Field	ds 🕶 🖌 🛛 📘	• - • •	• >
Record #	Birth Country	Count	-
1	Australia	4	
2	Austria	4	
3	Bangladesh	2	
4	Canada	2	
5	China	2	
6	Egypt	1	
7	Finland	1	
8	France	5	
9	Germany	9	
10	Ghana	1	
11	Hong Kong	3	
12	Hungary	3	
13	India	2	
14	Iran	3	
15	Israel	4	
16	Italy	4	
17	Japan	10	
18	Kenya	1	
19	New Zealand	2	-

Figure 2-102 – Nobel Laureates – Browse after Summarize



Figure 2-103 - Nobel Laureates - vis a vis All Medals

The Games

We see that we have a list of countries and a count of the number of Nobel Laureates. However, it is unclear what the number is because the field is called *Count*. We should rename it *Total Nobel Laureates*.



Figure 2-104 - Nobel Laureates - Select Configuration

We now have two data streams ready to be merged. We want to align the two datasets so that matching countries from each of the data steams share the same record, which means we want to join the data. Because we don't want to lose any data points if we have countries in one dataset but not the other, we will want to unite the three outputs from the join into a single data stream.

	ι	.eft			Right				
	1 (Country	82 82	• • •	Birth C	Country	•		
1	*		2	• ••			v		
			1 1					(R) (
Op	ptior	∎ 15 • ↑	↓ ↓	1	IP: To r	reorder multi	ple rows: s	elect, right-cli	ck and drag
Op	ptior	ns • † Input	↓ Field	T Type	TIP: To r Size	reorder multi Rename	ple rows: s Descrip	lect, right-cli	ck and drag
Op •	ptior	Input	Field Country	T Type V •	IP: To r Size 255	reorder multi Rename	ple rows: s Descrip	■elect, right-cli	ck and drag
0p		Input Left Left	Field Country Total O	Type V • D •	TIP: To r Size 255 8	reorder multi Rename	ple rows: s Descrip	elect, right-cli	ck and drag
o⊧ ▶		Input Left Left Right	Field Country Total O Birth C	Type V • D • V •	IP: To r Size 255 8 254	reorder multi Rename	ple rows: s Descrip	LR> (elect, right-cli	ck and drag
0p	ptior	Input Left Left Right Right	Field Country Total O Birth C Total N	Type V • D • V • L •	IP: To r Size 255 8 254 8	reorder multi Rename	Descrip	R (ck and drag

Figure 2-105 - Nobel Laureates - Join Configuration

We want to join on Country field from the Left (Input L) with Birth Country field from the Right (Input R).

It is important in this instance that we keep both joining fields because we intend to combine all three outputs in the next step. However, if this was not our intention, we could have removed the joining field from one of the two inputs.

Best practice is to give useful names to every connection that enters a multiple connection anchor.

Thus, we can see in the following image that we have relabeled the connections from #1, #2, and #3 to *Left*, *Join*, and *Right*.



Figure 2-106 – Nobel Laureates – Union Configuration and Output Stream

Since we are doing a union of three output streams of a Join tool, we know that we will have matching column names. This allows us to use the *Auto Config by Name* setting for the Union tool and leave the rest of the defaults.

We need to add a browse tool again as we have just altered the structure of the data. This is to make sure the data looks the way we expect. Notice that we are doing this after the Union and not the Join. That is because when we are combining the three output streams of a Join tool using a Union, we are performing a single logical step called an outer join. Because this is a single step, we know that we should check both tools if an issue arises.

We are getting close to our goal; however; the data stream is also starting to get complex. So we should take a minute to annotate what we have so it will be easier to follow later. We are going to add *Tool Containers* and *Comments* to the two contributing data streams so we can easily identify different parts of this data stream. We can create the comments and containers like we see in the next image.



Figure 2-107 - Nobel Laureates - Comments and Containers

We can now drag the appropriate tools into the tool containers so the data stream is easier to understand.



Figure 2-108 - Nobel Laureates - Output Stream with Comments

Looking at the data stream this way is helpful, but if we click on the arrows at the top-right corner, we can condense what we are looking at.



Figure 2-109 - Nobel Laureates - Simplified Output Stream

Now we can easily see the medal count preparation and the Nobel Laureate count preparation as two single processes instead of a series of tools. Now that we have made the data stream easier to understand, we should finish building the workflow.

1 1 1	13 records displayed		
Record #	Countrys	Total Olympic Metals	Total Nobel Laurat
1	Afghanistan	2	[Null]
2	Algeria	8	[Null]
3	Argentina	141	[Null]
4	Armenia	10	[Null]
5	Australia	609	4
6	Austria	91	4
7	Azerbaijan	25	[Null]
8	Bahamas	24	[Null]
9	Bahrain	1	[Null]
10	Bangladesh	[Null]	2
11	Barbados	1	[Null]
12	Belarus	97	[Null]
13	Belgium	18	[Null]
14	Botswana	1	[Null]
15	Brazil	221	[Null]
16	Bulgaria	41	[Null]
17	Cameroon	20	[Null]
18	Canada	370	2
19	Chile	22	[Null]

Figure 2-110 - Nobel Laureates - Browse Configuration

We can observe from the Browse that the country names matched the names in both the Country and Birth Country fields. Let's create a conditional formula that allows us to convert the two columns with nulls into a single column that always has a Country name.

Add a formula tool to the end of the data stream with a formula called **Countries** with the formula: **IF IsNull ([Country]) THEN [Birth Country] ELSE [Country] ENDIF**. This will take the Country value unless it is null and the Birth Country if it is.

Now we only need to clean up the data and export it to a .csv file. Add a Select tool with the following configuration, and export the file to *Brains vs Brawns.csv*.



Figure 2-111 - Nobel Laureates - Select Configuration

The final workflow is as shown in the following figure.



Figure 2-112 - Nobel Laureates - Complete Workflow

CHAPTER 3 Unisex Baby Names
	From +						
	То	Alteryx Consultants					
	Сс						
	Subject	Gender Swapped					
	Attached	Unisex Names.txt					
He	eγ,						
A na na de	A major newspaper is considering writing an article on the change in assignment of names to different sexes, and they would like our help finding out if there is anything definitive that they can say.						
Th th	ey provid e count of	ed the attached file that has unisex baby names from 1880 to 2013 and children given each name by sex.					
Th	is is a pret	tty interesting dataset, so I am going to work on this with you.					
Th 18 be	The first things they want to know are: Are there any names that between 1880 and 1889 were completely assigned to one sex and then completely assigned to another between 2004 and 2013? If so, which were they?						
Th	anks,						

3.1 Tools & Concepts

Tools	Concepts
Imputation	Handling Nulls
Multi-Field Formula	Multiple Record Calculations
Multi-Row Formula	Repetitive Calculations
Text Input	Ad Hoc Data
	Problem Analysis

	The <i>Imputation</i> tool replaces a specific value in numeric fields. It is most commonly used to handle Null values in formulas.			
Figure 2.1. Importation	Group	Input	Output	
Figure 3-1 – Imputation	Preparation	See below	See below	

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Any data stream with a numeric field.

Output: A data stream with at least one imputed field.

The tool will only modify an individual value and cannot be used to impute a range of values, such as all negatives.

Properties Window:

The *Imputation Configuration* window has five components.

Fields to impute	ΔII	No
Vear	<u>AII</u>	140
Female		
V Male		
Incoming value to replace		
Null()		
User Specified Value		
0.00000		
Replace with value		
O Average		
Median		
Mode		
User Specified Value		
0.00000 k.		
Include imputed value indicator field		

Figure 3-2 – Imputation configuration

• *Fields to Impute* is a list of numerical fields from the incoming data stream that one can choose to impute.

- *Incoming value to replace* allows us to choose to replace nulls or a specific numeric value that exists in the incoming dataset.
- *Replace with value* allows us to replace the value selected in *Incoming value to replace* with either the average (mean), median (middle), mode (most common), or a custom value that we select.
- *Include imputed value indicator field* creates a field for each of the *Fields to impute* with ones and zeros, where 1 indicates that the data was imputed and 0 that it was not.
- *Output imputed values as a separate field* creates a new field for each of the Fields to impute instead of replacing the values in the original field.

3.3 Multi-Field Formula

	The <i>Multi-Fiel</i> provides the a formula that v multiple fields Group	<i>d Formula</i> to bility to crea vill be reuse 5. Input	ol ate a d across Output
Figure 3-3 - Multi-field formula	Preparation	See below	See below

The formulas created here are arbitrarily complex, and thus can significantly slow down the data stream. Care must be taken to make sure the output field created has a field type compatible with the result created. // is the comment character.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: A data stream with at least two fields of the same type.

Output: The original data stream with fields modified or original data stream with a new field for every selected field from the input.

Application questions can be connected to the *Black Question Anchor* for "Use Answers" in this tool.

Properties Window:

The *Multi-Field Formula Configuration* window has five main components.

W N	Aulti-Field Formula (6) - Configuration 🛛 👻 🕂 🗙
**	Select Numeric Fields:
() 	Vear Vear Vear Male Right_Female
0	Copy Output Fields and Add Percent as a Suffix -
Ŭ	Variables Functions Saved Expressions
	Current Field Original Fields Oristants
	Expression:
0	Go

Figure 3-4 – Multi-field formula configuration

- *Select ... Fields*: populates the list with the fields that match what we have selected in the drop-down. Calculation is performed on each of the fields selected from the generated list.
- *Copy Output and Add ... as a ...*: when selected, creates a new field as the output instead of replacing the ones that we are performing the calculation on. *Decimal*, in the example, will be appended to the field name as a Suffix or Prefix.
- *Change Output Type to ... Size ...*: when selected, we can change the metadata of the output that we create.

- The last two sections are similar to the formula window that we are familiar with, except that since we are working across multiple fields, we have the ability to use the _*CurrentField_* in our calculations.
- 3.4 Multi-Row Formula

	The <i>Multi-Row Formula</i> tool provides the ability to create a formula that will reference other records.			
	Group	Input	Output	
Figure 3-5 – Multi-row formula	Preparation	Any data stream	See below	

The values are updated by moving down the records in order to modify the current record based on the result of the formula in the previous record. The formulas we create here can be arbitrarily complex, and thus can significantly slow down the data stream. Care has to be taken to ensure the output field created has a file type that is compatible with the result being created. The // is the comment character.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Application questions can be connected to the top *Black Question Anchor* for "Use those Answers" in this tool.

Output: The original data stream with a field modified or the original data stream with a new field.

Properties Window:

The *Multi-Row Formula Configuration* window has six core components.

🗳 N	/lulti-Row Formula (7) -	Configuration	🚽 🕂 🗙			
 ∢ ⊗	 Update Existing Field Oreate New Field Potential Swap 	Type String 🗸	Size 254			
	Num Rows Values for Rows that don't Exist 1 Image: Construction of the second					
0	Group By (Optional) Name Year Bucket Sum_Female Sum_Male					
	Variables Functions Save 	ed Expressions				
	Expression: IF [Name] != [Row-1:N [Name] != [Row+1:Name ELSE "Check" ENDIF	ame] AND] THEN "Remove"				
0			Go			

Figure 3-6 - Multi-row formula configuration

- *Update Existing Field* allows us to replace the entries in the field selected in the list below.
- *Create New Field* is the other option for the output, which allows us to create a new field instead of updating the current one. With this selection, we need to name the field and set the metadata of the field.
- *Num Rows* changes the number of rows before and after the current row we have access to in the formula.
- *Values for Rows that Don't Exist* allows us to define the behavior for records that do not actually exist. (In this case, for the first record, there is no *[Row-1: ...]*. So the field is replaced by a 0 since the field is numeric.)
- *Group By* allows us to reset the calculation based on changes in values in the fields we select.
- The bottom two sections are similar to the formula boxes we are used to seeing except that we have the added ability to refer to field values in different rows.

3.5 Text Input

F	The <i>Text Input</i> tool allows creation of datasets to be used by typing in the data.							
	Group	Input	Output					
Figure 3-7 – Text input	In/Out	None	Data stream					
An <i>Action</i> tool can conne modify how this tool wo	ect to the <i>Lightni</i> rks in apps and	An <i>Action</i> tool can connect to the <i>Lightning Bolt Anchor</i> to modify how this tool works in apps and macros.						

Properties Window:

The *Text Input Configuration* window allows us to create a data source that is internal to the module. The configuration has five options.



Figure 3-8 – Text input configuration

• *Import* allows us to import a dataset.

- *Copy* allows us to copy the cells we highlighted in the table.
- *Paste* allows us to paste values into the table.
- *Delete* allows us to delete everything, or the row or column we have selected.
- *Insert* allows us to create a row or column in the data set.

An example of what this would look like is provided in the previous image.

3.6 Gender Swapped

From +]					
То	Alteryx Consultants					
Cc						
Subject	Gender Swapped					
Attached	Unisex Names.txt					
Hey,						
A major newspaper is considering writing an article on the change in assignment of names to different sexes, and they would like our help finding out if there is anything definitive that they can say.						
They provi the count	ded the attached file that has unisex baby names from 1880 to 2013 and of children given each name by sex.					
This is a pr	etty interesting dataset, so I am going to work on this with you.					
The first things they want to know are: Are there any names that between 1880 and 1889 were completely assigned to one sex and then completely assigned to another between 2004 and 2013? If so, which were they?						
Thanks,						

Let us refer to the *Unisex Names.txt* file in the folder *Chapter* 3 – *Unisex Names.* It is a tab-delineated file. We need to make sure the file pop-up window looks like the image represented here. Let us save the workflow as *Gender Swapped*.

© Re	ead it as a built in	type			
[Alteryx database	(*.yxdb)			-
	Delimiter Comma Other:) Tab	○ Space	None	

Figure 3-9 - Tab delineated pop-up window

Add a *Browse* tool and a *Select* tool to the input. After running the workflow, we should be able to see the records in the file.

Res	Results - Browse (18)				→ 中 X		Gender Swapped.yxmd* ×
	4 of 4 Field	ds 🕶 🖋 📔	Cell Viewer	•	D• 🖬 🖬	• »>	
Ä	Record #	Name	Year	Female	Male	*	
	1	Addison	1880	[Null]	19		
	2	Adrian	1880	[Null]	18		
	3	Allie	1880	105	31		
	4	Allison	1880	[Null]	6		
	5	Alva	1880	17	70		Unisey Names tyt
	6	Antonia	1880	18	[Null]		Unisex Hames.cxc
	7	Asa	1880	[Null]	60		
	8	Ashley	1880	[Null]	8		
	9	Aubrey	1880	[Null]	21		
	10	August	1880	[Null]	210		

Figure 3-10 - Gender Swapped Browse

If we take a look at the data, we will see that some numeric fields have *nulls* in them. Whenever values are missing, Alteryx fills those cells with a "[Null]" placeholder. We could leave these values as nulls; however, it is easier to work with numeric fields when they do not have nulls in them. Instead, what we will do is impute the nulls to 0s.

Unfortunately, Alteryx did not recognize these fields as numeric, so we need to make modifications in the *Select* tool so that *Female* and *Male* are converted into *Double*.

Since we know we need to work with two 10-year periods, we should also convert *Year* to *Double* because it will make our formulas simpler and faster when we filter and flag the data.

_	_						
		Field	Туре		Size	Rename	Description
		Name	V_String	•	254		
		Year	Double	•	8		
	V	Female	Double	•	8		
		Male	Double	•	8		
		*Unknown	Unknown	-	0		Dynamic or Unknown Fields

Figure 3-11 – Gender Swapped Select

If we add the *Imputation* tool with the settings as seen in the image, the tool will convert all of the null fields in *Female* and *Male* fields to zeros.

Imputation Options Fields to impute All No Year Year Female Male Male Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Imputed value indicator field	P	utation (19) - Config	uration			× 4
Fields to impute All No Vear Female Female Male Incoming value to replace Null() User Specified Value Verage Average Median Mode User Specified Value 0.00000 Image Include imputed value indicator field	Ir	mputation O	otions				
Vear ✓ Female ✓ Male ✓ Male Incoming value to replace ● Null() ● User Specified Value Replace with value ● Average ● Median ● User Specified Value ● User Specified Value ● 0.00000 ↓	1	Fields to imp	ute			<u>All</u>	None
 Female Male Incoming value to replace Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Include imputed value indicator field 		Year					
Incoming value to replace Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Include imputed value indicator field		Male					
Incoming value to replace Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000							
Incoming value to replace Null User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Image Include imputed value indicator field							
Incoming value to replace Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Include imputed value indicator field							
 Null() User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 Include imputed value indicator field 		- Incoming v	alue to repla	ice			
 User Specified Value Replace with value Average Median Mode User Specified Value 0.00000 + Include imputed value indicator field 		Null					
 Oser specified value Replace with value Average Median Mode User Specified Value 0.00000			sified Value				
Replace with value Average Median Mode User Specified Value 0.00000 Include imputed value indicator field		O User spe	chied value	***			
 Average Median Mode User Specified Value 0.00000 Include imputed value indicator field 		Replace wit	h value				
 Median Mode User Specified Value 0.00000 Include imputed value indicator field 		O Average					
 Mode User Specified Value 0.00000 Include imputed value indicator field 		Median					
User Specified Value 0.00000		Mode					
0.00000		User Spe	cified Value				
Include imputed value indicator field		0.00000					
Include imputed value indicator field							
		Include im	puted value	indicator fi	eld		
—	1			and the second second			

Figure 3-12 – Gender Swapped Imputation configuration



Figure 3-13 - Gender Swapped data stream

Now let's see how the data has changed.

Res	ults - Brow	vse (18)				→ ₽	×	Res	ults - Brow	wse (20)				- ₽ X
=	4 of 4 Field	ds • 🖌	Cell Viewer	- 1			>>	=	4 of 4 Field	ds 🕶 🕪 🛛	Cell Viewer	• † 1	. •	••• »
Ä	Record #	Name	Year	Female	Male			-	Record #	Name	Year	Female	Male	*
	1	Addison	1880	[Null]	19				1	Addison	1880	0	19	
	2	Adrian	1880	[Null]	18				2	Adrian	1880	0	18	
	3	Allie	1880	105	31				3	Allie	1880	105	31	
	4	Allison	1880	[Null]	6				4	Allison	1880	0	6	
	5	Alva	1880	17	70				5	Alva	1880	17	70	
	6	Antonia	1880	18	[Null]				6	Antonia	1880	18	0	
	7	Asa	1880	[Null]	60				7	Asa	1880	0	60	
	8	Ashley	1880	[Null]	8				8	Ashley	1880	0	8	
	9	Aubrey	1880	[Null]	21				9	Aubrey	1880	0	21	
	10	August	1880	[Null]	210				10	August	1880	0	210	

Figure 3-14 - Gender Swapped Browse - Before and After

We can create a filter that limits the data set to *Years before 1890 or years after 2003* to isolate the two date ranges 1880 to 1889 and 2004 to 2013. We will add a filter to the end of the data stream. Use the expression box below to verify the Boolean formula.

```
Expression:
```

```
// [Year] < 1890 || [Year] > 2003
// or
[Year] < 1890 OR [Year] > 2003
```

Figure 3-15 - Gender Swapped expression box

Notice that in this expression, we have two lines starting with //. These two forward slashes mean that the line is a comment. It is a way to tell Alteryx to ignore that specific line from the calculation. We used it to show the two intended methods of filtering this data, but it can be used to maintain old versions of the formulas or leave descriptions of the formula for future benefit.



Figure 3-16 – Gender Swapped Data stream

The next thing we can do is create a *flag* for the first and last 10 years for the dataset. We want to do this to create a comparison between the two time frames. In order to do this, we will create a string field that will have *First 10 Years* or *Last 10 Years* in the field.



Figure 3-17 - Gender Swapped Formula configuration

We need to change the field type to string with a conditional formula to test if the record is in the first 10 years. Since the data is already filtered, we do not need to test the other values.

Now that we identified the data with the *Year Bucket* field, we can summarize the data to find out the total number of babies given each name during the 10 years we want to analyze.

	Field	Туре			
Þ	Name	V_String			
	Year	Double			
	Female	Double			
	Male	Double			
	Right_Female	Double			
	Dista Marta	Daubla			
	Right_Male	Double			
Actio	Year Bucket	String]	Output Field	
Actio	Year Bucket	Add Action]	Output Field Name	
Actio	Field Name	Add Action GroupBy]	Output Field Name Name	
Actio	Right_Male Year Bucket ns: Field Name Year Bucket	Add Action GroupBy GroupBy]	Output Field Name Year Bucket	
Actio	Right_Male Year Bucket ns: Field Name Year Bucket Female	Add Action GroupBy GroupBy Sum]	Output Field Name Name Year Bucket Sum_Female	

Figure 3-18 – Gender Swapped Summarize configuration

The data stream after *Summarize* looks as below:



Figure 3-19 - Gender Swapped Data stream after summarize

We will next run the workflow and look at the *Browse* tool to see the data we have.

its - Brov	vse (6)			*	φ×
4 of 4 Field	is 🔻 🖌	Cell Viewer 👻	1 1	🕒 • 🔒 • 🕀	• >>
Record #	Name	Year Bucket	Sum_Female	Sum_Male	
1	Addison	First 10 Years	0	170	
2	Addison	Last 10 Years	81476	2803	
3	Adrian	First 10 Years	0	148	
4	Adrian	Last 10 Years	1603	71931	
5	Adriel	Last 10 Years	258	4376	
6	Aidan	Last 10 Years	1025	65368	
7	Aiden	Last 10 Years	1179	131797	
8	Alexis	First 10 Years	0	11	
9	Alexis	Last 10 Years	90995	23277	
10	Ali	Last 10 Years	2930	8056	
	4 of 4 Field Record # 1 2 3 4 5 5 6 7 8 9 9	4 of 4 Fields Addison	4 of 4 Fields Cell Viewer Record # Name Year Bucket 1 Addison First 10 Years 2 Addison Last 10 Years 3 Adrian First 10 Years 4 Adrian Last 10 Years 5 Adriel Last 10 Years 6 Aidan Last 10 Years 7 Aiden Last 10 Years 8 Alexis First 10 Years 9 Alexis Last 10 Years 10 Ali Last 10 Years	4 of 4 Fields Cell Viewer Cell View	A of 4 Fields Cell Viewer I Image: Constraint of the state

Figure 3-20 - Gender Swapped Browse configuration

From the *Browse* tool, we can see that we have the data structured in a way we need. However, since we were asked to find the baby names that were previously only assigned to one sex and now are only assigned to the other, we need a way to check each of the names to see if they meet these criteria.

We will do this by using *Multiple Row Formula* to identify which records should be kept. Since we are using a Multi Row

Formula, we need to make sure the data is in the right order. So, we will first use a *Sort* tool.

	Name		Order	
	Name	-	Ascending	-
	Year Bucket	-	Ascending	-
*		-		-
•	Year Bucket	• •	Ascending	▼▼

Figure 3-21 - Gender Swapped Sort configuration



Figure 3-22 – Gender Swapped data stream after Sort

Now that we have made sure the data is sorted, we need to tackle the complex question. In order to do so, we will break it down into smaller, simpler ones:

- Which names occur in both the first and last 10-year lists?
- Which names have at least one count of zero?
- Which names have a female count alternate between something and zero?
- Which names have a male count alternate between something and zero?

By asking these four relatively simple questions, and filtering out the data that does not meet the criteria, we can answer the complex question we have been asked. We will use the same field as we move forward to filter the data, which we will call *Potential Swap*. We will create it in the first *Multi-Row Formula* tool and update it in all subsequent ones. Let us see how we can create this formula.

At this point, much of this formula should be familiar. We are looking at a conditional statement that tests two things, and if both are true, we write *Remove*; else, we write *Check*. What is different about this is that we have special operators in the field names *Row-1*: and *Row+1*:. These allow us to look at the row (record) above and below the current one so we can use the value there. In this case, we are checking to see if the name of the current row (*[Name]*) equals the previous or following rows.

Oreate New Field	Туре	Size	
Potential Swap	String	▼ 254	
Num Rows Values for	r Rows that don't Exist		
1 🚔 0 or Empt	у		
Group By (Optional)			
Name			A
Year Bucket			
Sum_Female			Clea
Variables Functions	Saved Expressions		
• Row +0 - Active Rov	N		
⊕ Row +1			
Expression:			
IF [Name] != [Row-	1:Name] AND [Nam	e] != [Row	+1:Na
THEN "Remove"			
ELSEIF [SUM_FEMAIE THEN "Check"] == 0 or [Sum_M	aie] == 0	
ELSE "Demove"			

Figure 3-23 - Gender Swapped Multi-row formula configuration

Those familiar with Boolean logic may have understood what the formula here actually does. The formula checks to see if the [*Name*] does not match both the previous and following rows. This is a logically equivalent statement called the *contrapositive*, and we are bringing it up here to demonstrate two things: First, *!*= is the *logical operator* that means *does not equal*. The second is to show if we are having trouble with a logical statement. There are multiple ways one can approach it.

Between each *Multi-Row Formula*, we can add a *Filter* to improve speed by removing records we know are not needed for

analysis. Let us add a series of alternating *Filters* and *Multi-Row Formulas* to answer the remaining three questions.

All *Filter* configuration windows for the next three steps should look like the image represented here.

Potential S	wap 🔻	Equals 🔹 C	heck
Variables	Functions	Saved Expressions	1
⊡- Const	ants		
Const.	ants		

Figure 3-24 - Gender Swapped Filter configuration

The remaining *Multi-Row Formula* properties windows should look like the provided image, with the following table of equations in the expression box.



Figure 3-25 - Gender Swapped remaining Multi-row formula

2 IF $[Sum_Female] == 0$ or $[Sum_Male] == 0$

THEN "Check"

ELSEIF [Name] == [Row-1:Name] AND ([Row-1:Sum_Female] == 0 or [Row-1:Sum_Male] == 0) THEN "Check"

ELSEIF [Name] == [Row+1:Name] AND ([Row+1:Sum_Female] == 0 or [Row+1:Sum_Male] == 0) THEN "Check"

ELSE "Remove" ENDIF

3 IF [Name] = [Row-1:Name] AND [Sum_Female] != 0 AND [Row-1:Sum_Female] != 0

THEN "Remove"

ELSEIF [Name] = [Row+1:Name] AND [Sum_Female] != 0 AND [Row+1:Sum_Female] != 0

THEN "Remove"

ELSE "Check" ENDIF

4 IF [Name] = [Row-1:Name] AND [Sum_Male] != 0 AND [Row-1:Sum_Male] != 0

THEN "Remove"

ELSEIF [Name] = [Row+1:Name] AND [Sum_Male] != 0 AND [Row+1:Sum_Male] != 0

THEN "Remove"

ELSEIF [Name] = [Row+1:Name] AND [Sum_Male] !=0 AND [Row+1:Sum_Male] !=0

THEN "Remove"

ELSE "Check" ENDIF

Figure 3-26 - Gender Swapped Formula

At this point, the data stream is quite long, so we have moved the tools below one another to make it easier to see.



Figure 3-27 - Gender Swapped data stream

We will notice that a *Browse* tool has been added here. Typically, if the formulas are being tested, then a *Browse* tool should be used on both the *True* and *False* outputs of every *Filter* tool. However, since these formulas were already tested, we are just going to take a look at the end result.

Res	ults - Brov	vse (15)				▼ ₽	×
	5 of 5 Field	ds 🔻 🖌	Cell Viewer	+ +		🕛 - 🕞 - 💷 -	>>
Ä	Record #	Name	Year Bucket	Sum_Female	Sum_Male	Potential Swap	
	1	Addison	First 10 Years	0	170	Check	
	2	Adrian	First 10 Years	0	148	Check	
	3	Alexis	First 10 Years	0	11	Check	
	4	Allison	First 10 Years	0	70	Check	
	5	Andrea	First 10 Years	51	0	Check	
	6	Angel	First 10 Years	0	8	Check	
	7	Antonia	First 10 Years	276	0	Check	
	8	Asa	First 10 Years	0	532	Check	
	9	Ashley	First 10 Years	0	80	Check	
	10	Ashton	First 10 Years	0	12	Check	

Figure 3-28 - Gender Swapped Browse configuration

We can now see that Beverly and Hilary are the only names that meet our criteria.

However, just like in the *Freestyle Skiing* example, we should make Alteryx give us exactly the results we will give to the newspaper so we do not make a mistake. This time, all we need to do is *Summarize* by the *Name* field to have the results that we are looking for.

	Field	Туре	
Þ	Name	V_String	
	Year Bucket	String	
	Sum_Female	Double	
	Sum_Male	Double	
	Patantial Course		
Actio	ns:	String	Add 🔻
Actio	Potential Swap	String	Add Output Field Name
Actio	Potential Swap	String Action GroupBy	Add Cutput Field Name Name
Actio	Potential Swap	Action GroupBy	Add Cutput Field Name Name
Actio	Potential Swap	Action GroupBy	Add Output Field Name Name

Figure 3-29 - Gender Swapped Summarize configuration

Running the *Browse* tool shows us the following:



Figure 3-30 - Gender Swapped Browse after summarize

The *Gender Swapped* data stream should look like the image on the next page when it is complete.



Figure 3-31 - Gender Swapped data stream when complete

3.7 What About Me?

From +	
То	Alteryx Consultants
Cc	
Subject	FW: What About Me?
That is pre	tty interesting.
Now that would be i	we have taken a look at which names were completely reassigned, it nteresting to find out if your name has a history of being unisex.
A quick loc going to us	ok will show you that Michael doesn't even appear in the data set, so I'm se the name Andrea, but feel free to use your name.
Let's see w the name y	hat the yearly percent breakdown for male and female children was for you choose.
Thanks,	

We will add a few things to this list to adhere to best practices, but the necessary steps we need are:

- 3. Import the data.
- 4. Clean up the fields.
- 5. Input a name.
- 6. Limit the records by that name.
- 7. Create calculated fields that show us the percent breakdown.
- 8. Export the data

We already know what issues we have in the dataset, so we can copy the last data stream until the *Imputation* tool.



Figure 3-32 - Reuse Gender Swapped for What about me

Now that we have this, the next step is to add the name to the data stream.

If we add a *Text Input* with a column called *Name*, and name as the only record, below is what we will get.



Figure 3-33 –What about me Text input

Since the *Text Input* will show us all of the data, we do not need to add a *Browse*. But it is a good practice to add the *Select* so that we can ensure that the data types from both data streams match.

Unisex Names Select Text Input Select



Figure 3-34 - What about me Select configuration

Though we have different types in our data field, we do not need to convert them because they are both strings. If they were not, it would cause an issue in joining the data. Due to the way we have approached this problem, we can use the *Join* tool to filter the data to the appropriate records. So we will take only the joined section and not keep the name field coming out of the right (*Text Input*).



Figure 3-35 - What about me Join configuration

Running the *Browse* tool for the *J* output will show us the following:

ults - Brov	vse (4)			• 1	+ >
4 of 4 Field	ds 🔻 🖌 🚽	Cell Viewer 👻	† ↓		• >
Record #	Name	Year	Female	Male	-
1	Andrea	1881	5	0	
2	Andrea	1884	7	0	
3	Andrea	1885	6	0	
4	Andrea	1886	9	0	
5	Andrea	1887	7	0	
6	Andrea	1888	7	0	
7	Andrea	1889	10	0	
8	Andrea	1890	6	0	
9	Andrea	1891	8	0	
10	Andrea	1892	16	0	

Figure 3-36 - What about me Browse configuration



Figure 3-37 - What about me Browse configuration

If we see the image here, we see that no records have been returned. It is because the name entered in the *Text Input* does not match anything in the *Unisex Baby Names* dataset. Change the *Text Input* name to *Andrea* to follow along more easily.

At this point, we have covered steps 1 through 4. The next step is for us to create the percent breakdown for both male and female babies in each year. Since we intend to perform the same calculation on two different fields, we can use the *Multi-Field Formula* tool to accomplish this. If we add the *Multi-Field Formula* tool after the *J* output from the *Join* with the following configuration, we will be able to create the percent of total for Male and Female babies each year.



Figure 3-38 –What about me Multi-field formula configuration

Now, all that's left is to write the file out to *What About Me.csv*.

The *What About Me* data stream should look like the below image when it is complete.


Figure 3-39 - What about me data stream when complete

3.8 What's In A Name?

From +	
То	Alteryx Consultants
Cc	
Subject	FW: What's In A Name
Hev	
So L board boa	k from the neuropener, and they decided to rup the article. But they are curious about
something that	at they think will be an interesting spin.
See the email	below.
I don't have ti	me to help you out this time, but I'm interested in the result. Let me know when you've
got it.	
Thanks,	
From: Erin	
To: Michael D	avis
Subject: Unis	ex Baby Names
Hello,	
What you spu numbers work	n up was good, but I think that we are going in a different direction, as long as the < out.
For some cont	text:
When someor	e tells us their name, we make assumptions about them, one of the first being their sex.
I want to know them concern	v the 25 names most likely to be heard and have an incorrect assumption made about ing sex, based on this data.
When I say "n of referring to "Unknown" ar	nost likely" let's call it the most amount of years when names had at least a 25% chance males and females. (If there are ties, they should be alphabetical. You can remove nd "Baby" from the list of names).
Deadline's ap Erin	proaching,

CHAPTER 4 The Direct Approach

То	Alteryx Consultants
Subject	What's The Policy On That?
Hey,	
As I'm	sure you're aware, we are working for a small company.
Frank policie open t	has asked me to build a process that helps us stay on top of our new s. I will be using a software program that, given a link, can automatically he file or web address that it points to. But first, I need to prep the data.
I'm goi functio	ng to have you sit with me on this so you can see some more of Alteryx's onalities.
They o	nly want policies that were published in the last 30 days to show up.
Thanks	Ъ.

4.1 Tools

Tools	Concepts
Append Fields	Appending Data
Date Time Now	Cartesian Product
DateTime	Working With Dates
Directory	Working With Directories

4.2 Append Fields

	The <i>Append</i> tool adds each record from <i>S</i> (<i>source</i>) to the end of each record in <i>T</i> (<i>target</i>).				
Eigure 4-1 - Append fields	Group	Input	Output		
figure 1 1 - Append fields	Join	See below	See below		

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input S: Any data stream.

Input T: Any data stream.

Output: A data stream of records that has the Cartesian product from inputs *T* and *S*, with the records of *S* appended to the end of *T* records.

This tool is very useful for appending a single data value onto a full dataset, such as appending the mean of a field to every row for calculation purposes. Since we are taking the Cartesian product of two data streams, we run the risk of creating a very large data stream, which could cause memory errors.

The *Append Fields Configuration* window is identical to the *Select Configuration* window, with two exceptions.

		Input	Field	Туре		Size	Rename	Descrip
•	~	Target	FullPath	V_WString	-	254		
	-	Target	Directory	V_WString	-	254		
	-	Target	FileName	V_WString	-	260		
	-	Target	Short Fil	V_WString	-	14		
	~	Target	Creatio	DateTime	-	19		
	-	Target	LastAc	DateTime	-	19		
	-	Target	LastWri	DateTime	-	19		
		Target	Size	Int64	-	8		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
		Target	Attribut	Bool	-	1		
<								

Figure 4-2 - Append fields configuration

- There is an additional element in the metadata section called *Input*. It identifies if the data is coming from the *Target* (*T*) or *Source* (*S*) inputs for the tool.
- There is a *Warn/Error on Too Many Records Being Generated,* which allows us to decide if and how we should be alerted to a high rate of replication of the *Target* field.

4.3 Date Time Now

	The <i>Date Time</i> system time w starts executin	<i>Now</i> tool ge hen the moo g.	ts the dule
Figure 4-3 – Date time now	Group	Input	Output
	In/Out	None	String

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: A string representing the date or the date and time, depending on the settings.

The tool is useful if we want to limit our data stream based on the current time, or if we want to publish the time of a report or dataset when created.

The *Date Time Now Configuration* window allows us to select the format of the string required in the output.



Figure 4-4 – Date time now configuration

The following image represents the format of the strings available. See *Appendix G* for a list of character descriptions.



Figure 4-5 - Date time now formats

4.4 DateTime

	The <i>DateTime</i> tool converts between String and Date format fields			
	Group	Input	Output	
Figure 4.6 - Date time	Parse	See below	See below	

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: A data stream with a date field in a *string, date,* or *date/time* field.

Output: The original data stream with a new field. The new field cane be a *string* representing the date or the date/time depending on a *date* or *date/time* field that was represented as a string field.

The tool can only convert the formats in the menu. We need to write a *Formula* that does string conversions in addition to using this tool if we need a different type of date format.

The DateTime Configuration window has four components.



Figure 4-7 - Date time configuration

- *Conversion Mode* allows for converting from or to a string field.
- Formatted Input String Field (Input Date/Time Field to be *formatted*) is the field we want to convert.
- *Format of Input String (Desired Format of Output String)* is the format the input string is in.
- *Output Date/Time Field (Output Formatted String Field)* lets us name the field we are creating.

4.5 Directory

	The Directory tool creates a data stream that has the contents of a directory.					
	Group	Input	Output			
Figure 4-8 - Directory	In/Out	None	See below			

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: A data stream containing metadata about all of the files in a given directory.

We can use special characters to look for a file meeting a pattern. For example, *.* will return all files, and *a* *.*csv* will return all *.csv* files starting with the letter *a*. These filematching commands are different from the regular expression syntax, which we will discuss in *Expensive Beauty Products*.

The *Directory Configuration* window has three components.



Figure 4-9 - Directory configuration

- *Directory* lets us navigate to a folder.
- *File Specification* allows us to type the generic format for a file name to be found in the *Directory*, where * denotes zero or more characters or spaces, while ? denotes exactly one character or space.
- *Include SubDirectories,* when checked, includes each of the subfolders when it's looking for files. Otherwise, it will only look in the directory that we have navigated to.

4.6 What's The Policy On That?

То	Alteryx Consultants
Subject	What's The Policy On That?
Hey,	
As I'm	sure you're aware, we are working for a small company.
Frank policie open t	has asked me to build a process that helps us stay on top of our new is. I will be using a software program that, given a link, can automatically he file or web address that it points to. But first, I need to prep the data.
l'm go functio	ing to have you sit with me on this so you can see some more of Alteryx's onalities.
They o	only want policies that were published in the last 30 days to show up.
Thank	S.

Since we need to search for files, we are going to be using the directory tool. Our finalized policies are published as *pdf* files, so we can use that to limit our search. We have three divisions responsible for publishing policies right now, and they all publish them to subfolders in *Chapter 4 – The Direct Approach > What's The Policy On That.*

The plan is to:

- Bring in a list of all of the *PDF* files in the directory.
- Get the current date.
- Remove files created more than 30 days ago.
- Export the data to What's The Policy On That.csv.

When we bring the *Directory* tool onto the canvas and navigate to the *Policy* folder in the *Directory* file browse, we can use the expression *.*pdf* to find all *PDF* files in the directory. But since we need to search all of the sub-folders (subdirectories), we need to check the box at the bottom of the configuration window.





Figure 4-11 Policy Data stream

As usual, we will add a *Browse* and *Select* tool following the data connection. Let's look at the *Select* tool to see what's in the data stream.

	Field	Туре		Size	Rename	Description
] FileName	V_WString	•	260		
V	FullPath	V_WString	•	32767		
C] Directory	V_WString	•	32767		
] ShortFileName	V_WString	•	14		
] Creation Time	DateTime	•	19		
] LastAccessTime	DateTime	•	19		
/ /	LastWriteTime	DateTime	٠	19		
	Cute 78 Dat	Date		71	**	****
	1 OGUV	Date	•			1

Figure 4.12 – Policy Select Configuration

We can see the data stream has 19 fields. Whenever we use the *Directory* tool, it will return these 19 metadata fields about each of the files that were found. See *Appendix K* for details on each. In our case here, we will only be working with the *FullPath, FileName,* and *LastWriteTime* fields. Now that we have the appropriate information from the list of files, the next thing we need is to add the *current date* into the workflow. This is where the *Date Time Now* tool comes in. Please note here that the annotations are hidden to save space.



Figure 4-13 - Policy Date time now configuration

We will set the output format of the *Date Time Now* input to *yyyy-MM-dd* so it's easy for us to know which format it's in when we are converting it to the cutoff date 30 days ago. It's important to note that the *Date Time Now* tool creates a single field with a single record, which is a string representing the date when the module runs.

Running the *Browse* and *Select* tools show us the following information.



Figure 4-14 - Policy browse configuration for Date time now

Options - 1					TIP: To reorder multiple rows:			
_		Field	Туре		Size	Rename	Desc	
	•	DateTimeNow	String	•	10			
1	-	*Unknown	Unknown		0		Dynar	

Figure 4.15 - Policy Select Configuration for Date time now

In order to convert this string into a *DateTime* field, we need to use the *DateTime* tool. So if we add it right after the *Select* tool like the following image, we can create a new field called *Today* that has today's date as a *Date* field.



Figure 4.16 - Policy Using Date time to convert String to Date Time format

Now we need to find out what the date 30 days ago was. In order to do this, there's a simple *Formula* function we can use. If we look under the *Functions* tab of the *Formula Configuration window*, we see that one of the branches is called *DateTime* and it has the formulas we can use to manipulate *Date*, *Time*, or *DateTime* fields. We want to find out what the date 30 days before today was. We need to remember that we must add a negative interval of days when using the *DateTimeAdd* function to obtain a previous date.



Figure 4-17 - Policy using Formula to find Date 30 days ago

A description and example of how the tool works is displayed when we click on the function. A double-click on it shows us to add the DateTimeAdd(dt,i,u) expression to the *Expression* box.

- *dt* should be replaced with a *Date*, *Time*, or *DateTime* field.
- *i* should be replaced with the interval we want to add. (-30 for 30 days ago).
- *u* should be replaced with a unit keyword from the list in Appendix H. In this case, it is "days."

Now that we have these two data streams ready to be combined, the question becomes: What method do we use? A *Union* would allow us to combine the data. Unfortunately, since we will be performing a calculation to see if the dates were in the last 30 days, we need the *Cutoff Date* in every record not added to the end of the dataset. A union will not work.

A *Join* would allow us to do this if we had a field we could match, which means we could use a function on both data streams that just writes 1 to the field, and then join on that. But that's difficult to explain and maintain.

Fortunately, Alteryx has a tool called *Append Fields* that does exactly what we want to do with the *Formulas* and a *Join*, in an optimized way. We shall connect the list of *PDFs* to the *Target* (*T*) input and the date to the *Source* (*S*) input. Since we are changing the structure of the data by combining two data streams, we should put a *Browse* tool following the *Append Fields* tool to follow the output at this stage.



Figure 4.18- Policy Append Configuration to Use Formulas with Join

As we can see, *DateTimeNow* and *Today* fields can be excluded when we were joining the data. This is because they are no longer needed, as we only need to compare the *LastWriteTime* to the *Cutoff Date* to determine if the file should be included in the list.

Now we can compare the *LastWriteTime* to the *Cutoff Date*. If it is before the cutoff date, we don't need those records. The way we're going to approach this is a little different from what we've done so far. Let's write a formula in the *Filter* tool to satisfy the condition.



Figure 4-19 - Policy Comparing Last Write Time to Cutoff Date

This looks normal. However, notice that the formula is true any time *LastWriteTime* is less than (before) the *Cutoff Date*. This is because we can easily use the *False* (*F*) output to finish the data stream similar to the way we can use the *True* (*T*).

Now, let's output the data to a file called *What's The Policy On That.csv*.

The *What's The Policy On That* data stream should look like the below when it's complete.



Figure 4.20- Policy data stream when complete

4.7 Where, Oh Where Have The Three Files Gone?

То	Alteryx Consultants
Subject	Where, Oh Where Have The Three Files Gone?
Hey,	
I know	this is a little unorthodox, but it will be a good way to score points with me and my boss.
Usuall	y, I would take care of things like this, but Frank needs me in a meeting.
Frank	has lost some important files, and they don't remember what the files are called.
Frank and th drive i	knows they saved the files somewhere on their network drive under a particular folder each, at you will recognize the names of the files as soon as you find them, but searching the network s something Frank doesn't have time to do.
l need a PDF "Chapt indepe	you to build a directory and sub directory search for an Excel file (.xlsx) in the "Excel Files" folder, file (.pdf) in the "PDF files" folder, and an image file (.png) in the "Images" folder in ter 4 – The Direct Approach" > "Looking for Files." Look at the results of each of the three endent file lists, and identify the ones we need.
After t so tha	hat, bring the data together into an output file with the two columns "File Name" and "Full Path", t by the time the meeting is over, we can direct Frank right to the files.
I really	appreciate this.

CHAPTER 5 Cultural Musings

То	Alteryx Consultants					
Subject	Culturally Rich					
Attached	1 2010 GDP Per Capita Per State.xlsx;					
Hey,						
We have so the more lik	me downtime, and I would like to explore the idea that the richer the state, ely it is to have museums.					
We are goin the 2010 Ce	We are going to use the attached list of 2014 Q3 Museums data, 2010 GDP data, and the 2010 Census data.					
If you don't have the 2010 Census data package yet, you can download it at <u>http://downloads.alteryx.com/data.html</u> . Just walk through the installer, and when you see the uninstall screen, make sure nothing is selected.						
Let's walk th	nrough the data prep.					
Thanks.						

5.1 Tools & Concepts

Tools	<u>Concepts</u>		
Allocate Input	Using Census Data		
Find Replace	Data Mapping		
Join Multiple	Complex Joins		
Text To Columns	Splitting field members at delimiters		



The *Allocate Input* tool allows us to get input from demographic data using specific packages.

	Group	Input	Output
Figure 5-1 – Allocate Input	Demographic Analysis	None	See below

Note: Specific packages have to be downloaded to use this tool. The *Census Data* is a free package to use here.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: A data stream that contains the selected fields from an Allocate Dataset.

The *Allocate Input Configuration* window has three core components and two sub-components for selection and customization.



Figure 5-2 – Allocate Input Configuration

Choose a Dataset drop-down allows us to select the dataset associated with one of the packages we have downloaded.

Use Allocate Workspace check box helps us define the fields and geographies we want to use.

The selection window is broken down into tabs. We see different tabs depending on whether or not the *Use Allocate*

Workspace is checked. When unchecked, we see the *Pick Geography*, *Pick Variables* and *Special Fields* tabs.

- *Pick Geography* tab allows us to select the geographic level of the data required.
- *Pick Variables* tab allows us to select the demographic variables needed in the output.
- *Special Fields* tab allows us to select the required special fields relevant to our selections in the output.

When *Use Allocate Workspace* is checked, we see the *Workspace* and *Special Fields* tabs.



Figure 5-3 – Allocate Input Configuration (contd.)

- *Workspace* can import and modify the *Allocate Workspace* file (.aws), which allows us to save tree-based selections.
- *Special Fields* allows us to select the special fields needed in the output.



The *Find Replace* tool allows us to replace information in a data stream by entering the *F* (find) (target) input with information that matches in the *R* (replace) (source) input.

Figure 5-4 – Find Replace	Group	Input	Output
	Join	See below	See below

Note: The *Find Replace* tool is useful in renaming field entries or for replacing specific information in a data stream. In simple words, it works like a lookup function.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input F: A data stream with at least one string field

Input R: A data stream with at least two fields: one mapping to the string field being replaced in the *F* input, and the other field that's required in the output by replacing the original.

Output: The *F* input data stream with an updated field or the *F* input data stream with additional fields from the *R* input.

The *Find Replace Configuration* window has two main components: *Find* and *Replace*.



Figure 5-5 – Find Replace Configuration

- *Find* allows us to define the matching criteria.
 - *Find Within Field (Target)* is the field we want to match in the data stream connected to the *F* input.

• *Search Term Field (Source)* is the field that we want to match in the data stream connected to the *R* input.

We can select one of the three ways to match the two fields:

- Beginning of Field will match only if the Search Term Field (Source) matches the beginning of the Find Within Field (Target).
- *Any Part of Field* will match if the Search Term Field record is found anywhere in the *Find Within Field* record.
- *Entire Field* will match if the entire *Search Term Field* record matches with the *Find Within Field* record.
- *Case Insensitive Find* allows us to match a field irrespective of uppercase or lowercase in both the sources.
- *Match Whole Word Only* allows us to match only if the *Search Term Field* matches to a word or set of words in the *Find Within Field*. For example, if this option were checked, *apple* would not match *pineapple*, but if it were unchecked, it would match.
- *Replace* allows us to define what the replacement field is
 - *Replace Found Text With Field* allows us to select a field from the R input with which to replace the values.
 - Replace Multiple Found Items (Find Any Part of Field only) allows us to choose if we want multiple occurrences of a string to be replaced in a single record.

- *Append Field to Record* allows us to add specific fields to the data instead of just the replaced value.
- 5.4 Join Multiple



The *Join Multiple* tool will perform a full outer join or an inner join between data streams sharing a set of key fields.

in Multinle	Group	Input	Output
in Munipic	Join	See below	See below

Note: If there are multiple records that match each other (based on the joining field(s)), records from the original data streams be replicated.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

The *Input* must have at least two data streams with the same mapping fields.

The *Output* will be a data stream with all records joined by the key fields. It can either include all records from each data stream or only those that match across all data sources.

The *Join Multiple Configuration* window looks and functions similarly to the *Join Properties* window with only the following three differences.

	Muse	aums		F	opul	ation			GDP per Capita	
1	State	,	- 4	⇒ N	lame	•	•	4	State	•
•			- 🗢	•			-	\$		
Ca	artesian	Joins:								
Er	tor on m	ultidime	ensional joi	ins o	f mor	e than	16 Recor	rds		
O	nly Outp	out Rec	ords that	Join	from	All Inp	uts	>		
	Inp	out	Field	1	уре	Size	Renam	e	Description	CIICK a
•	Inp	out seums	Field NAME	T V	ype •	Size	Renam	e	Description	
•	Inp	out seums seums	Field NAME ALTNA	ד V . V	ype •	Size 254 254	Renam	e	Description	
Þ	Inp	out seums seums seums	Field NAME ALTNA ADDR	T V . V	ype •	Size 254 254 254	Renam	e	Description	
	Inp Mut Mut Mut	out seums seums seums seums	Field NAME ALTNA ADDR City2	T V . V V	ype	Size 254 254 254 254 254	Renam	e	Description	
	Inp Mu Mu Mu Mu Mu	out seums seums seums seums seums	Field NAME ALTNA ADDR City2 ZIP5	T V . V V V	ype	Size 254 254 254 254 254 254	Renam	e	Description	
	Inp Imp Imp	out seums seums seums seums seums seums	Field NAME ALTNA ADDR City2 ZIP5 PHONE	T V V V V V	ype • • • • • •	Size 254 254 254 254 254 254 254	Renam	e	Description	
	Inp Imp Imp	seums seums seums seums seums seums seums	Field NAME ALTNA ADDR City2 ZIP5 PHONE WEBU	T V V V V V V V	ype 	Size 254 254 254 254 254 254 254 254	Renam	e	Description	
	Inp Imp Imp	seums seums seums seums seums seums seums seums	Field NAME ALTNA ADDR City2 ZIP5 PHONE WEBU LATIT	T V . V V V V V V V	ype • • • • • • • • • • • • • • • • • • •	Size 254 254 254 254 254 254 254 254 254	Renam	e	Description	
	Ing V Mus	out seums seums seums seums seums seums seums seums	Field NAME ALTNA ADDR City2 ZIP5 PHONE WEBU LATIT LONGI	T V V V V V V V V V V	ype • • • • • • • • • • • • •	Size 254 254 254 254 254 254 254 254 254 254	Renam	e	Description	

Figure 5-7 – Join Multiple Configuration

- It will have as many columns to join by, as there are input data streams.
- It has the ability to alert in case of too many replications.

- Instead of displaying Venn diagrams for the potential outputs, there is a checkbox (*Only Output Records that Join from All Inputs*) for deciding if we want the intersection of all of the datasets (the *J* output in the *Join* tool Inner Join) or the full outer join (when unchecked).
- 5.5 Text To Columns



The *Text To Columns* tool will break string fields into multiple string fields based on a delimiter.

Figure 5-8 – Text to columns

Group	Input	Output
Parse	See below	See below

Note: This tool is useful in breaking down strings like email addresses or transaction elements so as to be able to look at the component pieces. It uses a single character to split the fields.

For example, if we type @ | into the *Delimiters* text box, *Text To Columns* will split fields at @ and again at |, and not when it encounters @ | together.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

The *Input* data stream should have at least one string field.

The *Output* will have the original data with either additional columns or rows generated by splitting the original field.
Properties Window:

The *Text To Columns Configuration* window has the following five core components.

Û T	ext To Columns (15)	- Configuration 💿 👻 🕂 🗙
*	Field to Split City	Delimiters
₽ -	Split to Columns # of Columns:	2
\$	Extra Columns: Output Root Name:	Leave Extra in Last Field City
0	Split to Rows Advanced Options	
	Ignore Delimiters in Ignore Delimiters in Ignore Delimiters in Ignore Delimiters in Skip Empty Fields	Quotes Single Quotes Parenthesis Brackets
Ø		

Figure 5-9 – Text to columns configuration

- *Field to Split* allows us to select the field to be split from the incoming data stream.
- *Delimiters* are the list of special characters or separators used to split the field by.

- *Split to Columns* allows creation of new columns for each split and has the following three secondary questions:
 - *# of Columns* allows us to select the number of new columns to be created.
 - *Extra Columns* is a drop-down menu that allows us to decide what happens to that data that would be split into more than the number of columns selected.
 - *Output Root Name* is the beginning of the name of the columns created by the split.
- *Split to Rows* changes the behavior such that new rows are created instead of additional columns for every split that is found.
- *Advanced Options* allow us to decide if there are more granular ways to control the splits.

5.6 Culturally Rich



The files required for this chapter have been placed in clearly marked sub-folders in *Chapter 5 – Cultural Musings > Culturally Rich*. We will start by bringing in each of the four data sources to see what we have. Let us open all three files and then connect to the *Census* data.

	Results - Brow	wse (18)	▼ 🖓	×
	2 of 2 Field	ds 👻 🖌 Cell Viewe	er 🕶 📄 • 🔂 • 🕀 •	, >>
	Record #	State	2010 GDP per Capita	*
	1	District of Columbia	148729	
	2	Delaware	61677	
	3	Alaska	60873	
	4	North Dakota	46867	
	5	Connecticut	55269	
	6	Wyoming	56689	
10 GDP Por	7	Massachusetts	51889	
TO ODF FEI	8	New York	52242	
oita Per	9	New Jersey	49026	
to allow	10	Oregon	45375	
te.xisx	11	Washington	45631	
le=`Sheet1\$`	12	Virginia	47042	-
e oncesty	4		•	

Figure 5-10 – Culturally Rich – GDP data



Figure 5-11 - Culturally Rich - Museum Data

As can be seen in these two files, the names of the states appear in different formats, with the *GDP per Capita* file having the entire state name spelled out, and the *Museum* data having only the two-letter state abbreviation within the city field.

	Results - Brow	rse (20)	er 🕶 🕆 🙏 51 mecords disc	- 0	×
	Record #	State	Census Replacements	Abbreviation Replacement	
	1	Alabama	Alabama	AL	
	2	Alaska	Alaska	AK	-
	3	Arizona	Arizona	AZ	1
	4	Arkansas	Arkansas	AR	
	5	California	California	CA	
	6	Colorado	Colorado	co	
	7	Connecticut	Connecticut	СТ	
	8	Delaware	Delaware	DE	
	9	District of Columbia	District of Columbia	DC	
	10	Florida	Florida	FL	
State Man yley	11	Georgia	Georgia	GA	
State Map.XISX	12	Hawaii	Hawaii	HI	+
CA11 101 101	4			,	

Figure 5-12 – Culturally Rich – State Map

However, the third file has three separate columns with either the complete state name or its abbreviation. This is because *State Map.xlsx* is a file created especially for the purpose of field mapping by linking the data sources from the *GDP*, *Museum* and *Census* data sets, which all have state identifiers in different formats.

The following are the observations from the data in the three files we connected to.

- The *GDP per Capita* data does not need any preparation before the join.
- The *State* associated with each museum needs to be parsed out of *City*, and then mapped to the *GDP per Capita* name.
- The *State Map.xlsx* file can be used for mapping all data sources together.

Let us now bring in the *Census* data in order to plan what needs to be done with that data stream.

We shall use an *Allocate Input* tool on the canvas, then choose the *US Census* 2010 – *Most Recent Vintage* dataset. Let us

select the *States* option under *Pick Geography,* and click the checkbox to select all states.



Figure 5-13 – Culturally Rich Allocate Input configuration – Pick Geography

The purpose of bringing in the *Census* data is to compare the *GDP* to the number of museums in each state instead of the *GDP per Capita*. We can use the two data sources to generate the GDP value.

Now that the goal is identified, figuring out the variable needed becomes easy.



Figure 5-14 – Culturally Rich Allocate Input configuration – Pick Variables

Under *Population* and *US Census (2010)*, the last option is *Total Population*. This signifies the total number of people that have lived in each chosen State (Geography).

Since we need only this information, click on the *Special Fields* (as shown in image below) and uncheck all options.



Figure 5-15 – Culturally Rich Allocate Input configuration – Special Fields

Now that we have started with the final data stream, let us take a look at the data that comes out of the *Census* data connection.

3 of 3 Field	ds 🔻 🐦 🛛 Cell View	er 🔻	
Record #	Name	Key	P0010001
1	Alabama	01	4779736
2	Alaska	02	710231
3	Arizona	04	6392017
4	Arkansas	05	2915918
5	California	06	37253956
6	Colorado	08	5029196
7	Connecticut	09	3574097
8	Delaware	10	897934
9	District of Columbia	11	601723
10	Florida	12	18801310
11	Georgia	13	9687653
12	Hawaii	15	1360301

Figure 5-16 – Culturally Rich Allocate Input Browse configuration

The following two points about this data stand out:

- Even though we have only selected one geography and one variable, we still have three fields. This is because the *Key* field uniquely identifies all geographies so that even if we have regions with the same name, they have unique identifiers. In this case, because they are States, we do not need the *Key*.
- There is a field called *P0010001*. This is because the data is stored with keyed column headers. We will simply rename the field to *2010 Population* in the *Select* tool, as shown in the following image.

		Field	Туре		Size	Rename	Description
Þ		Name	V_WString	-	100		
		Key	V_WString	-	32		
	V	P0010001	Int64	-	8	2010 Population	Population.
		*Unknown	Unknown	•	0		Dynamic o.

Figure 5-17 – Culturally Rich Census Population data Select configuration

Now, the remaining preparation is to map all of the state names from the different data sources back to the *GDP per Capita State* names before joining these data streams.

Let us start with mapping the *Census* data since it requires only one step before the join. In order to do this, we will use the *Find Replace* tool. Connect *Census* data stream to the *Find* (*F*) input and the mapping data stream to the *Replace* (*R*) input as shown in the following image.



Figure 5-18 – Culturally Rich Connecting census data to Find and State Map data to Replace

Now, let us take a look at the settings of *Find Replace* tool.



Figure 5-19 – Culturally Rich first Find Replace configuration

Since the map has been structured in such a way that the field *Census Replacements* has all of the *Census State* names in it, we will be looking for the *Entire Field* matches of the *Census Replacements* in the *Census Name* field. We also want to make sure that we replace the fields found with the *State* field (which has the *State* names from the *GDP* data). Since the *Census* data stream and the *GDP* data stream share a mapping field, let us get the *Museums* data ready.

We will first remove all of the fields except *Name* and *City* in the *Select* tool following the *Input Data* tool. Next, we will split the data apart. To do this, use the *Text to Columns* tool and add it to the end of the *Museums* data stream with the following settings.

Field to Sp	olit	Delimiters
City		
Split to	Columns	Comm
# of C	olumns:	2
Extra	Columns:	Leave Extra in Last Field 🔹
Outpu	t Root Name:	Location
Split to	Rows	
Advanced	Options	
Ignore Ignore	Delimiters in Delimiters in Delimiters in	Quotes Single Quotes Parenthesis
Ignore	Delimiters in I moty Fields	Brackets

Figure 5-20 – Culturally Rich Text to columns configuration to split at a separator

Then, let us look at the updated data after it has been modified with the *Text to Columns* field.

4 of 4 Field	ds 🔻 🖌 🛛 Cell Viewer 👻 🏌 🌵 35,144 rec	ords displayed, 1.2 MB		
Record #	NAME	City	Location1	Location2
1	WESTERVELT WARNER MUSEUM	TUSCALOOSA, AL	TUSCALOOSA	AL
2	U. S. ARMY AVIATION MUSEUM	FORT RUCKER, AL	FORT RUCKER	AL
3	WESTERVELT-WARNER MUSEUM OF AMERICA	TUSCALOOSA, AL	TUSCALOOSA	AL
4	U.S. SPACE & F C - 11		HUNTSVILLE	AL
5	WELLBORN MI Cell With leading	ng spaces	ALEXANDER CITY	AL
6	VON BRAUN ASTRONOMICAL SUCIETY	HUNTSVILLE, AL	HUNTSVILLE	AL
7	USS ALABAMA BATTLESHIP COMMISSION	SPANISH FORT, AL	SPANISH FORT	AL
8	SLOSS BLAST FURNACE SITE	BIRMINGHAM, AL	BIRMINGHAM	AL
9	STURDIVANT HALL	SELMA, AL	SELMA	AL
10	THREE NOTCH MUSEUM	ANDALUSIA, AL	ANDALUSIA	AL
11	SLOSS FURNACES NATIONAL HISTORIC LAND	BIRMINGHAM, AL	BIRMINGHAM	AL
12	SPRINGVILLE MUSEUM	SPRINGVILLE, AL	SPRINGVILLE	AL

Figure 5-21 – Culturally Rich Modified Museum data

When we look at the *Browse* tool, we can see that *Location2* has a *red triangle* at the top-right corner of the cell. It is an alert indicating that Alteryx recognizes a potential issue with the data. If we hover the mouse over the cell, we will see a little text-box pop-up that reads "This cell has leading spaces". This means the first character of the cell is a space (attributed to the split and a space after "," delimiter) and that we should address it. We will do this by using the *Formula* tool to create a new field, we will call *State*.



Figure 5-22 – Culturally Rich Trimming Museum data using Formula

This time, we will be creating a string. If we look under the *String* functions, we can find three functions designed to remove one string from another. We will use the *Trim* function because we do not know if there are cells having trailing spaces which would cause issues with mapping.

Since we are removing spaces from the field, we could simply write the formula *Trim([Location2])* because the function removes white spaces by default. However, for the purpose of

clarity and readability, the best practice is to explicitly define all the variables. In this case, it would mean writing the formula as *Trim*([*Location2*]," ").

Now that we are left with clean two-character *State* abbreviations, we can map them to the *GDP State* names. Once again, we will bring a *Find Replace* tool onto the canvas and connect the mapping data to the *Replace* input.



Figure 5-23 – Culturally Rich second Find Replace configuration Now that we have prepared all the contributing data streams, this is what the data stream should look like.



Figure 5-24 – Culturally Rich Consolidated Modified data streams

Notice that we have crossed the data streams. From a technical perspective, it does not matter if they are crossed. However, practically, the workflow becomes harder to read when the data streams are crossed frequently.

Best practice is to create workflows with as few crossovers as possible. In this case, we leave it as-is because there will invariably be at least one cross over in this module. We could use a couple of *Join* tools now, but since all data streams share the same key (*State* name), we can use a *Join Multiple* tool to join all three data streams at once. We will connect the *GDP Per Capita*, *Museum*, and *Population* data streams to a *Join Multiple* input in the same order. We can then rename the connectors similarly and configure it as shown in the image.

One thing to note is that Alteryx renames conflicting fields with the data stream name and an "_". For example, since there are two fields that are named *State*, the second occurrence is renamed with the data stream as an identifier: *Museum_State*.

[(GDP Per Ca	apita	М	useu	m		Population		1
	1	State		⊧⇒ S	tate	-		Name	•	
	•		-	\$ \$		Ŧ	4		-	
	Carte	sian Joins:								
ſ	Error	on multidime	ensional join	s of mo	ore that	an 16 R	ecords			
	Delv	Output Pag	ordo that l	oin from	- AII	Innuto				
_	-			-		-	-			
		Input	Field	Туре	,	Size	Ren	ame	Descrip	tion
Þ	V	Input GDP Pe	Field State	Type V_St		Size 255	Ren	ame	Descrip	tion
Þ	√	Input GDP Pe GDP Pe	Field State 2010 G	Type V_St. Doub		Size 255 8	Ren	ame	Descrip	tion
Þ	 ✓ ✓ ✓ 	Input GDP Pe GDP Pe Museum	Field State 2010 G NAME	Type V_St. Doub V_St.	le •	Size 255 8 254	Ren	ame eum Name	Descrip	tion
•	 ✓ ✓ ✓ ✓ 	Input GDP Pe GDP Pe Museum Museum	Field State 2010 G NAME City	Type V_St. Doub V_St. V_St.	le •	Size 255 8 254 254	Ren	ame eum Name	Descrip	tion
•		Input GDP Pe GDP Pe Museum Museum Museum	Field State 2010 G NAME City Locatio	Type V_St Doub V_St V_St	e •	Size 255 8 254 254 254	Ren	ame eum Name	Descrip	tion
Þ		Input GDP Pe GDP Pe Museum Museum Museum	Field State 2010 G NAME City Locatio	Type V_St. Doub V_St. V_St. V_St. V_St.	le •	Size 255 8 254 254 254 254	Ren	ame eum Name	Descrip TextToC TextToC	tion
		Input GDP Pe GDP Pe Museum Museum Museum Museum Museum	Field State 2010 G NAME City Locatio State	Type V_St. Doub V_St. V_St. V_St. V_St. V_St.	le •	Size 255 8 254 254 254 254 254 254	Muse	ame eum Name eum_State	Descrip TextToC TextToC	Colu
•		Input GDP Pe GDP Pe Museum Museum Museum Museum Populati	Field State 2010 G NAME City Locatio Locatio State Name	Type V_St. Doub V_St. V_St. V_St. V_St. V_St. V_St. V_W	le •	Size 255 8 254 254 254 254 254 254 100	Ren Muse Muse	ame eum Name eum_State	Descrip Text ToO Text ToO	tion Colu
		Input GDP Pe GDP Pe Museum Museum Museum Museum Populati Populati	Field State 2010 G NAME City Locatio Locatio State Name 2010 P	Type V_St. Doub V_St. V_M Int64	· · · · · · · · · · · · · · · · · · ·	Size 255 8 254 254 254 254 254 254 8 100 8	Ren Muse Muse	ame eum Name eum_State Ilation_Name	Descrip TextToC TextToC	colu colu

Figure 5-25 – Culturally Rich Join Multiple configuration



This is what the data stream would look like.

Figure 5-26 – Culturally Rich Unified data stream

Once we perform the following steps, our data preparation will be complete.

• Add a field called *Total GDP* (product of GDP Per Capita and Population).

	Outpu	ut Field	Туре	Size	Expression	
1	2010	GDP	Double •	8	[2010 GDP per	ÌÌ
•			Double -	8		
						1
I						
Vari	ables	Functio	ns Saved Ev	oraccia		
+: +:	Existin Consta	g Fields ants	Javed Ex	5165510	ns	
	Existin Consta	g Fields ants		JIESSIO	ns	
+ + xpre	Existin Consta ession: L0 GDP	g Fields ants	apita] * [2010	Population]	
* *	Existin Consta ession: LØ GDI	g Fields ants P per C	apita] * [2010	Population]	
***** ****	Existin Consta ession: LØ GDP	g Fields ants P per C	apita] * [2010	Population]	
**** ****	Existin Consta ession: LØ GDP	g Fields ants	apita] * [2010	Population]	

Figure 5-27 – Culturally Rich Total GDP calculation

• Export the data to a file called *Culturally Rich*.

The *Culturally Rich* data stream should look like this when it is complete.



Figure 5-28 – Culturally Rich Complete data stream

5.7 Culturally Divided

To <u>.</u>	Alteryx Consultants
S <u>u</u> bject	Culturally Divided
Hey,	
I want to cre each state by appropriate of the museu	ate a map in another tool that will approximate the cultural appreciation of y identifying what the museum's per capita rates are. This may not be an measure, but I'm curious. Note that this time, we will need to take a count ums per state, not just apply the state data to each museum.
It would also capita. (Aboy	be nice if I had a flag for the comparison to the average museum's per ve average, average, and below average).
I need to run	to a meeting; do you think you can get this done by the time I get back?
Thanks.	

Chapter 6 The Sport Report

То	Alteryx Consultants
Subject	The Sport Report
Attached	Major Sport Teams.xlsx
Hey,	
l got a call fr exactly wher United States	om a sports bar chain. They want to see a rough report that could show them e the major professional baseball, basketball, football, and soccer teams in the s and Canada are so they can target new locations.
I want you to	help me build a report with the following properties.
Every	sport should start on a new page.
There It sho	should be a generic company header and footer (we can use the default for now). uld say the name of the sport.
There	should be a map of all of the stadium locations for that sport.
There assoc	e should be a bar graph that has the total count of teams by the governing iation.
There the di each	e should be a table that has the name of the association, the league, or conference, vision, the team, the city, the home park, the street address, and the zip code, for team. (The table should be allowed to take in as many pages as necessary.)
Thanks.	

6.1 Tools & Concepts

In this chapter, we will be covering the following Tools and Concepts:

Tools	<u>Concepts</u>
Charting	Report Building
Create Points	Spatial Objects
Layout	Maps
Render	
Report Footer	
Report Map	
Report Header	
Table	

To the reader,

We will be covering the basics of report building in Alteryx. However, since building reports in Alteryx is an advanced technique, we will not be exploring the Report tools as completely as the tools in other sections.

USEReady

6.2 Charting

0	The <i>Charting</i> tool creates a chart object for reporting.			
	Group	Input	Output	
Figure 6-1 - Charting	Reporting	See below	See below	
An <i>Action</i> tool can connect to the <i>Lightning Bolt Anchor</i> to modify how this tool works in apps and macros. Application questions can be connected to the <i>Top Black Question Anchor</i> to use those answers in this tool.				
Input: A data stream with at least one numeric field.				
<i>Output:</i> A data stream with a record for every element in fields grouped by, a column for each grouping field, and a field called <i>Chart</i> with a <i>Chart object</i> in each record.				
This tool creates a reporting object.				

Properties Window:

The *Charting Configuration* window has four components: *Chart Type, Select Field(s) To Chart, Preview,* and *Setup Chart Appearance.*

	Iration	Province (Click to only and):
Chart Type: Bar		Freview (Click to enlarge).
Select Field(s) To Chart:	Θ-	
Ser	ries 1	Series 1
Bar [Nor	ne] 🗸 🗸	
Label Field: [None]	~	
Setup Chart Appearance General	:	Reset to Defau
Setup Chart Appearance General Font Details Borders, Backgrounds	3D Interior	Reset to Defau
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout	3D Interior Gradient Coloring	Reset to Defau
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis	3D Interior Gradient Coloring Width (inches)	Reset to Defau No 6
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles	3D Interior Gradient Coloring Width (inches) Height (inches)	Reset to Defau No 6 4
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watemark	3D Interior Gradient Coloring Width (inches) Height (inches) Resolution	Reset to Defau No 6 4 2x (192 dpi)
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) Height (inches) Resolution Title	Reset to Defau No 6 4 2x (192 dpi)
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) Height (inches) Resolution Title Title Size	Reset to Defau No 6 4 2x (192 dpi) 14
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) Height (inches) Resolution Title Title Size Subtitle	Reset to Defau No 6 4 2x (192 dpi) 14
Setup Chart Appearance General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) Height (inches) Resolution Title Title Size Subtitle Subtitle Size	Reset to Defau No 6 4 2x (192 dpi) 14 12

Figure 6-2 – Charting Configuration

Chart Type:	Bar
Select Field	Area Stacked Area Column
	Stacked Column
	Bar
Bar	Stacked Bar
	Line
	Combination: Area And Line
	Combination: Column And Line
	Tomado
	Pareto
	Box And Whisker
	Scatter
	Bubble
	Polar
	Radar
	Pie
4	

Figure 6-3 - Chart Types

Chart Type allows us to select from a list of charts.

For each of these charts, the *Setup Chart Appearance* will reflect the appropriate settings. We will be walking through the basic settings when we use a bar chart.

Select Fields to Chart is where we can select each of the individual items that represent a series in the data. When working with a *Bar* chart, this means we can take additional columns to create additional bars for each field. At the bottom of this window, there is an ability to select the *Label Field*, which allows us to change the label on the bars.

Preview shows us a sampled version of what our resulting chart will look like. It only shows the formatting and does not refer to the real data stream. If we click on the image, it opens in a pop-up window so we can get a clearer look at what we are designing.

The *Setup Chart Appearance* has two main components: the list on the left side, which allows us to select which aspect of the chart we want to modify, and the settings on the right side that we can modify. Most of the settings we will see in the following options are fairly self-explanatory, and some may have secondary questions that appear when we select them. We will be covering all of these options at a conceptual level.

Setup Chart Appearance	:	Reset to Defaults
Font Details	2D Interior	No
Grouping	3D Interior	
Legend, Data Layout	Gradient Coloring	No
Vertical Axis Horizontal Axis	Width (inches)	6
Top Horizontal Axis Series 1 Styles	Height (inches)	4
Watermark	Resolution	2x (192 dpi) 🔹
	Title	
	Title Size	14
	Subtitle	
	Subtitle Size	12
	Primary Font	Microsoft Sans Serif B I

Figure 6-4 - Chart Appearance - General

Under *General* setting, the first two settings change the formatting of the bars. *3D interior* makes the bar appear 3D while *Gradient Coloring* makes the bars look rounded. With *Width* and *Height*, we can set the dimensions of the graph image. *Resolution* allows us to set the quality of the image. *Title* and *Subtitle* allow

us to type a title onto the graph, while *Title Size* sets the size of the font. *Primary Font* allows us to set the default font used for the graph.

Setup Chart Appearance		Reset to Defaults
Font Details Borders, Backgrounds	Change Title Font	No
Grouping Legend Data Layout	Color	Black
Vertical Axis Horizontal Axis	Alignment	Center 🗸
Top Horizontal Axis Series 1 Styles	Change Subtitle Font	No
Watermark	Color	Black
	Alignment	Center
	Change Legend Font	No

Figure 6-5 - Chart Appearance - Font Details

Font Details allow us to change the font settings for the Title, Subtitle, and Legend.

Setup Chart Appearance:		
General		Reset to Defaults
Font Details Borders, Backgrounds	Border Chart	No
Grouping Legend, Data Layout Border	Border Area	No
Vertical Axis Horizontal Axis	Background Color	White
Top Horizontal Axis Series 1 Styles	Area Color	White
Watermark	Interior Color	White

Figure 6-6 - Chart Appearance - Borders, Backgrounds

Borders, Backgrounds allow us to change the color and size of borders throughout the graph (excluding the bars).

Setup Chart Appearance: General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Top Horizontal Axis Series 1 Styles Watermark	All Clear
---	--------------

Figure 6-7 - Chart Appearance - Grouping

With *Grouping*, for each unique element in the field(s) selected, a different chart will be created (and a different record will be created in the output).

-Setup Chart Appearance		
General		Reset to Defaults
Borders, Backgrounds	Show Chart Legend	Yes 🔹
Grouping Legend, Data Layout	As Separate Image	No
Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Alia	Position	Top 🔹
	Alignment	Center
Watermark	Border Legend	No
	Plot Data By:	Record

Figure 6-8 - Chart Appearance - Legend, Data Layout

In *Legend, Data Layout,* the first five tools allow us to customize the legend. *Plot Data By* changes the way the data is colored within the graph. The default is to color by series; h. However, if we change this setting from *Record* to *Field*, we will be able to change the color of the data by a specific field.

- Setup Chart Appearance		
General		Reset to Defaults
Borders, Backgrounds	Font	Microsoft Sans Serif B I
Grouping Legend, Data Layout	Size	8
Vertical Axis Horizontal Axis	Color	Black
Top Horizontal Axis Series 1 Styles	Axis Title	
Watermark	Font	Microsoft Sans Serif B I
	Size	10
	Color	Black
	Show Grid Lines	Yes 🔹

Figure 6-9 - Chart Appearance - Vertical Axis

Vertical Axis allows us to customize the font of the *Axis Title* and labels, as well as turn the gridlines on or off.

General		Reset to Defaults
Borders, Backgrounds	Font	Microsoft Sans Serif B I
Grouping Legend, Data Layout	Size	8
Vertical Axis Horizontal Axis	Color	Black
Top Horizontal Axis Series 1 Styles	Axis Title	
Watermark	Font	Microsoft Sans Serif B I
	Size	10
	Color	Black
	Show Grid Lines	Yes 🗸
	Prefix	
	Suffix	
	Custom Origin	No
	Manual Increment	No
	Manual Minimum	No
	Manual Maximum	No
	Use Highlight Line	No
	Custom Decimal Places	No

Figure 6-10 - Chart Appearance - Horizontal Axis

Horizontal Axis allows us to customize the horizontal axis. The top half is the same as the *Vertical Axis* settings. However, this being a numeric axis, it gives us the ability to customize the range, set a reference (highlight) line, and change the number format.

Setup Chart Appearance	:		
General]		Reset to Defaults
Borders, Backgrounds	Use Top Axis	No	•
Grouping			
Vertical Axis			
Horizontal Axis			
Top Horizontal Axis			
Watemark			

Figure 6-11 - Chart Appearance - Top Horizontal Axis

The *Top Horizontal Axis* allows us to set, define, and customize a secondary axis for the chart.

Setup Chart Appearance General		Reset to Defaults
Font Details Borders, Backgrounds	Style Mode	Standard 🗸
Grouping Legend, Data Layout Label Data Poin	Label Data Points	No
Vertical Axis Horizontal Axis	Series Color	R=179, G=0, B=35
Top Horizontal Axis Series 1 Styles Watermark		

Figure 6-12 - Chart Appearance - Series X Styles

The *Series* X *Styles* allows us to set the style for the series X so we can differentiate them.



Figure 6-13 - Chart Appearance - Watermark

Watermark allows us to set a watermark on the graph.

6.3 Create Points

	The <i>Create Points</i> tool converts			
	decimal latitudes and longitude			
	columns into a new field called			
	<i>Centroid</i> with a record.	<i>Centroid</i> with a point object for each record.		
Figure 6.14 Create Points				
rigule 0-14 - Cleate Follits	Group	Input	Output	
rigure 0-14 - Create Points	Group Spatial	See below	See below	
An <i>Action</i> tool can be co	Group Spatial	See below	See below It Anchor to	
An <i>Action</i> tool can be co modify how this tool wo	Group Spatial onnected to the rks in apps and	See below Lightning Bo macros.	See below	

Input: A data stream with latitude and longitude columns.

Output: The incoming data stream with an additional field containing point objects.

This tool creates a spatial object field.
Properties Window:

The *Create Points Configuration* window has two core components.



Figure 6-15 – Create Points Configuration

The *X Field* (*Longitude*) and *Y Field* (*Latitude*) are dropdowns that allow us to select the longitude and latitude fields from the input data stream.

The single value list following the field selection allows us to define how the latitude and longitude are defined.

- *Fields are Lat/Long Floating Point* means the fields are in a standard decimal format.
- *Fields are Lat/Long Integers (x 1,000,000)* means the fields have been standardized to an integer format.
- *Fields are Projected Floating Points* allows us to identify a specific formatting method to pinpoint spatial fields.

6.4 Layout



An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Any data stream with at least one report object field.

Output: The original data stream with a single *Layout* object that combines other report objects. Depending on the settings, the original report object fields may be removed.

This tool creates a reporting object.

Properties Window:

The *Layout Configuration* window has the following five core components.

Layout Mode:	Each Individual Record	
		All
Group By:		Clear
Include Sour	ce Fields in Output	
Layout Configu	ration	
Orientation:	Horizontal	
Layout Width:	Percentage v 100 🜩 Perc	ent
Layout Height:	Automatic 🗸	
Border		
Separator		
	0	

Figure 6-17 – Layout configuration

- *Layout Mode* allows us to define the method used to combine the different report objects.
- *Group By* (when using the *Each Group of Records* option from the *Layout Mode* drop-down) allows us to define what is to be combined by the group of records that share a field(s).
- *Include Source Fields in Output* allows us to decide whether or not the original report fields that have been combined should be in the outgoing data stream (assuming we have chosen the *Each Individual Record* from the *Layout Mode* drop-down).
- *Layout Configuration* allows us to set some parameters for the look and feel of the resulting *Layout* object(s).

Width:	Automatic 🗸 🗸	
Alignment (V):	Middle	~
Alignment (H):	Center	~
Fill Color		

Figure 6-18 - Per Column Configuration

Per Column Configuration has two setups. *Each Individual Record* mode gives us the ability to select the reporting objects that we want to combine across a record, while the other two modes allow us to set the field(s) that we want to combine when creating a single record per group.

6.5 Render

	The <i>Render</i> too allows us to se reporting that	l creates a fi e the structu we have des	ile that 1red signed.			
	Group	Input	Output			
Figure 6-19 - Render	Reporting	See below	File			
An <i>Action</i> tool can be connected to the <i>Lightning Bolt Anchor</i> to modify how this tool works in apps and macros.						

Application questions can be connected to the *Top Black Question Anchor* to use those answers in this tool.

Input: Any data stream with at least one report object.

Output is a file that has the structured report we've designed.

If we require more than one report object, we need to combine them with the *Layout* tool and use the *Layout* object.

Properties Window:

The *Render Configuration* window has three core components.

nder (10)	Configura	tion	▼ ₽
Render Conf	iguration		
Output Mode	e: Temporary	y PDF Document	~
Output File:			
Group Da	ata Into Separa	ate Reports	
Field To	Group On:	Sport	~
Modify F	Filename By:	Appending Group To Filename	~
Report Data			
Data Field	Layout		~
Separator	Insert Section	n Breaks Between Records	~
✓ Header	Header		~
✓ Footer	Footer		~
	0 🔹 He	eader/Footer edge distance (points) ader and footer on first page	
Report Style			
Paper Size	Letter (8.5"	x 11")	~
Orientetien	Portrait () Landscape	
Orientation			

Figure 6-20 - Render Configuration

Render Configuration allows us to set the type and location of the report output.

Output Mode allows us to define a temporary file in one of a few different formats, as well as define a permanent file to write to.

Output file lets us define a path and filename for the file.

Group Data Into Separate Report, when selected, allows us to create multiple output reports based on a grouping field selected in *Field To Group On*.

Modify Filename By allows us to define how each of the files should be named in the output.

Under *Report Data, Data Field* allows us to select the field in the body of the report.

Separator allows us to set what is between each record.

Header and *Footer* allow us to define the header and footer fields respectively.

Header/Footer edge distance (points) is the number of points everything is from the edge of the paper.

Show header and footer on first page allows us to turn those aspects on or off for the first (often title) page of the report.

Report Style allows us to set the paper and layout properties of the final report.

Paper Size allows us to define page dimensions.

Orientation defines the way the report is laid out.

Margins (in.) define the margin width in inches.

Edit Default Styles allows us to change the default fonts and coloring.

6.6 Report Footer

F	The <i>Report Footer</i> tool allows us to create a footer to be used in reports.					
	Group	Input	Output			
Figure 6-21 – Report Footer	Reporting	Data stream	See below			
An <i>Action</i> tool can connect to the <i>Lightning Bolt Anchor</i> to modify how this tool works in apps and macros.						
<i>Output:</i> The original data stream, with the additional field <i>Footer</i> , which has a <i>Layout</i> object in each record.						
This tool creates a reporting object and should be used after report elements have been joined together.						

Properties Window:

The *Report Footer Configuration* window allows us to define the footer.



Figure 6-22 – Report footer configuration

Copyright text is intended for a copyright line on each page.

Information Text is a left-aligned text field that can include hyperlinks in the format <Link text to display |URL>.

Include page numbers allows to set the visibility of the page number at the bottom-right corner of the report.

6.7 Report Map



An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros

Application questions can be connected to the *Top Black Question Anchor* to use those answers in this tool.

Input: Any data stream with at least one spatial object field.

Output: A data stream with a record for every element in fields grouped by, a column for each grouping field, and a field called *Map* with a *Map object* in each record.

This tool creates a report object.

Properties Window:

The *Report Map Configuration* window has two main components. The *Preview* window gives us a look at the map to be generated.

	port Hap (35) - Co Preview Dick to enlarge		Refresh	Settings Data Layers Lagand 0 • 0 1 1 0 • • Sold from 645 steam A Sole A Data Theme Base Layers - Lines • Base Layers - Lines	Layer Name: Data Connectio Show in Legens Zoom to Layer Disable:	points from data stream m. (#1) 2. 2.	
5	Settings Data Laye Map Size (W x H): Resolution: Scale: Reference Base Map: Background Color: Map Drop Shadow: Expand Extent:	m Legend 8 by 6 Inches 3x (96 dpt) Meles CoudMade - Most Recert Virtage R=253. 0=254. 0=255 Ne 10 5: w1 a Minimum Vidth of 0.5	v v miles	Settions Data Lavers	Legend		
Setti	ngs Data Layers	Bats Connection Name: Spatial Field: Grouping Field: Thematic Field: Label Field:	#1 [All SpatialD • Sport • Association • Association •	Position Background Color Text Color Font Font Size Icon Width Icon Height Numeric Comma Set	Us Us Us Us Us parators	ione] e Default e Default e Default e Default e Default e Default o	Reset to Defaulta

Figure 6-24 - Report Map Configuration

The bottom of the window is broken into four tabs:

- The *Settings* tab allows us to set high-level options including dimensions, resolution, scale units, the background map to be used, background color (if not using a map), whether or not the objects should have a drop shadow, and some zoom settings.
- The *Data* tab allows us to select data streams and add them as data layers. When an incoming connection or file connection is selected, we can define how the layers that use that connection will behave. *Data connection name* is how it will be identified. *Spatial Field* allows us to select one or all of the spatial object fields in the dataset to define.

Grouping Field will create a new map for each unique element in the selected field. *Thematic Field* allows us to color the elements based on a particular field. If a string field is to be used, we need to set the theme to *unique* and type in the values. *Label Field* allows us to select the field we will display the text from on the map.

- The *Layers* tab allows us to modify specific aspects of the map. We can add and remove layers to the map using the plus and minus symbols, as well as change the floating order of the objects. By selecting different objects in the tree, we can modify various settings. Some of these settings will be covered in the following exercise.
- The *Legend* tab allows us to set a number of options related to including a legend with the map. By default, the legend is off and accepts the default values.

6.8 Report Header



Header, which has a *Layout* object in each record.

This tool creates a reporting object and should be used after report elements have been joined together.

Properties Window:

The *Report Header Configuration* window allows us to design a header for the report.

ort	Header (11) - Configuration 👻
Cont	figuration
The this	Report Header macro will append the created header to your supplied data new field can then be used in a Render tool as the Header for your report.
Ent	er the report title:
The	Sport Report
	Include the date in your report header.
	Specify the date format:
	Month dd, yyyy 🗸 🗸
(Use the Alteryx logo Specify custom image:
	Browse to the image to use in the report header:
	The size of the Alteryx logo is 150 by 55 (W x H) pixels, if you use an image of a different size or ratio, you may encounter layout issues. You may need to resize your image outside of Alteryx.

Figure 6-26 – Report header configuration

We can type in a title, decide if we want a date and set its format, and also decide if we want an image logo in the header.

6.9 Table

0 3	The <i>Table</i> tool creates a data stream with one or more tables.				
	Group	Input	Output		
Figure 6-27 - Table	Reporting	Any data stream	See below		

An *Action* tool can connect to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Application questions can be connected to the *Top Black Question Anchor* to use those answers in this tool.

Output: A data stream with a record for every element in the fields grouped by, a column for each grouping field, and a field called *Table* with a *Table object* in each record.

This tool creates a report object, which can be used with other report objects in the table tool.

Properties Window:

The *Table Configuration* window allows us to create a report table.

Table Mode Basic			*
	Pivot (CrossTab)		
Sport			A
âroup By			Clea
Table Configuration			
Table Width Automa	atic 🗸		
Pivot Style Method	dology over Variable		
Bar Graph [No Va	lid Fields]	V	Settings.
Show Column Headi	ngs		
✓ Sport Dynamic or Unknow	n Fields Nignment (H) Borders Prefix Suffix Dec. Places Column Rules	Automatic V Center V None V 0 ¢ Create	

Figure 6-28 – Table configuration

Table Mode has the following two options:

- *Basic* will create a table from the rows and columns we have.
- *Pivot (CrossTab)* allows us to use fields we have created using a *CrossTab* tool in order to create a series of more flexible tables.

Group By will create a different table for each of the unique elements in the grouping fields.

Table Configuration allows us to modify the appearance of the table. *Table Width* allows us to set a standard width or let it vary with the available space. *Pivot Style* allows us to display elements of the pivoted data from a *Crosstab* tool in different ways. *Bar Graph* allows us to create a spark bar for each element in the table. *Show Column Headings* adds a label to the top of each of the columns.

Per Column Configuration allows us to select columns to be display in the table and to set the format for each column. *Column Rules* allows us to create a conditional formatting for the column.

Create Row Rule allows us to create a record level formula.

Default Table Settings allows us to change the default formatting of the table.

6.10 Major Sports Teams

То	0	Alteryx Consultants
Subjec	ct	The Sport Report
Attach	ed	Major Sport Teams.xlsx
Hey,		
l got a exactly United	call fro where States	om a sports bar chain. They want to see a rough report that could show them e the major professional baseball, basketball, football, and soccer teams in the s and Canada are so they can target new locations.
I want	you to	help me build a report with the following properties.
•	Every	sport should start on a new page.
:	There It sho	should be a generic company header and footer (we can use the default for now). uld say the name of the sport.
•	There	should be a map of all of the stadium locations for that sport.
•	There assoc	should be a bar graph that has the total count of teams by the governing iation.
•	There the di each	should be a table that has the name of the association, the league, or conference, vision, the team, the city, the home park, the street address, and the zip code, for team. (The table should be allowed to take in as many pages as necessary.)
Thanks	5.	

When building reports in Alteryx, we need to remember the following key points:

- Most report objects should be created in individual data streams and then brought together. The exceptions are *Layout, Header,* and *Footer* objects.
- Report objects should be in the same record for each grouping.

Let us start by connecting to the data (*Chapter 6 – Sport Report*) > Major Sports Teams.xlsx.



Figure 6-29 - Opening Sport Report

As in any other workflow, we will bring in the data and ensure that it is in the correct format.

		Field	Туре		Size	Rename	Descriptio
Þ	•	Sport	V_String	-	255		
	✓	Association	V_String	•	255		
	•	League / Conference	V_String	•	255		
	•	Division	V_String	•	255		
	•	Team	V_String	•	255		
	•	City	V_String	-	255		
	•	Home Park	V_String	•	255		
	•	street address	V_String	•	255		
	•	zip	V_String	-	255		
	•	Latitude	V_String	•	255		
	•	Longitude	V_String	•	255		
	•	*Unknown	Unknown	-	0		Dynamic

Figure 6-30 - Sports Report- Data formatting

Since we have *Latitude* and *Longitude* fields, we can easily start tackling the mapping aspect of our report. We start by adding a *Create Points* tool to the end of the workflow and using the accompanying configuration.



Figure 6-31 Sport Report - Mapping and creating points

Then, when we add a *Browse* tool and run the workflow, we will see that we have a field called *Centroid* in the last column with a green point object in each record. This is a special spatial field, and we will use this field in order to plot the locations of each team's home field.



Figure 6-32 - Sport Report - run module using Browse

The resulting data stream will look like the below.



Figure 6-33 - Sports Report Data Stream

Since we are only going to use the one input for this *Report Map*, we can leave the connection labeled #1. We need to think about the way the report is going to look, so we should consider the overall layout. It will likely flow better if the map and chart are placed next to each other. This way, we know that image will not take more than half of the page.

In the following pages, the map configuration will allow us to create the desired maps.



Figure 6-34 – Report map setting configuration

The following are the Report Map Data and Layers settings.

Note that we want to group by sport to make one map for each.

Settings Data Layers	Legend			
File Connections	s Da Select data connection Th Lat	ta Connection N tial Field: uping Field: ematic Field: bel Field:	ame: #1 Centroid Sport [None] [None]	Specify Centroid Map per sport

Figure 6-35 - Sport Report - Report Map Data configuration



Figure 6-36 - Sport Report - Report Map Layers configuration

Normally, we would not put a *Browse* tool after *Report* tools, but it will help us understand the output of the *Report* tools.



Figure 6-37 - Sport Report - Browse

If we run the module, we should see the output in the *Browse* tool at the end of the data stream.

We see that we have four records, one for each of the four sports we grouped by, identified by the *Group* column at the right. A *Map* field contains the *Report* object map for each sport. The *BoundingRect* spatial field can be used to create new maps that are zoomed to the same area as these. Since we will not be creating any more maps, we can ignore this field.

Also, we should note that we can preview the map and report it if we want by selecting the tabs along the top.



Figure 6-38 - Sport Report - Chart creation using Browse

Now that we have our maps, let us create our charts. We will be making a bar graph for each sport, showing the number of teams in each association. This means that the first thing we should do is summarize the data by sport and association, then take the count of the team names.



Figure 6-39 - Sports Report - Chart creating using Select

Since we want to have clean field names while reporting, we will change the name of the *Count* column to *Number of Teams*.



Figure 6-40 – Sport Report - Adding Charting Tool to end of data stream

Now that we have the data prepared we can add a *Charting* tool to the end of this data stream and apply the settings as shown in the image.

Chart Type: Bar		Preview (Click	k to enlarge):
Select Field(s) To Chart:	Θ-	Teams per	Governing Body
Series 1	Series 2	Record 3	
Bar Number of	Teams 🔻 [None]	Record 2	
		Record 1-	
•	•		
Label Field: Association	on 🗸		
Setup Chart Appearance		- 112	
General		Res	et to Defaults
Font Details			
Borders, Backgrounds	3D Interior	No	Similar to Repo
Borders, Backgrounds Grouping Legend, Data Lavout	3D Interior Gradient Coloring	No	Similar to Repo map configurati
Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis	3D Interior Gradient Coloring Width (inches)	No No 3.5	Similar to Repo map configurati
Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis	3D Interior Gradient Coloring Width (inches) (Height (inches) (No 3.5 3.5	Similar to Repo map configurati
Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) (Height (inches) (Resolution (No No 3.5 3.5 3.5 3x (288 dpl)	Similar to Repo map configurati
Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) (Height (inches) (Resolution (Title	No No 3.5 3.5 3.5 3.2 3.2 3.2 3.2 3.2 3.2 3.2 3.2	Similar to Repo map configurati
Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark	3D Interior Gradient Coloring Width (inches) (Height (inches) (Resolution (Title Title Size	No No 3.5 3.5 3x (288 dpi) Teams per Gover 14	Similar to Report map configurati

Figure 6-41 - Sport Report – Charting Tool

The images below show the grouping, legend, and data layout settings made to the *Charting* tool.

Setup Chart Appearance: General Font Details Borders, Backgrounds Grouping Legend, Data Layout Vertical Axis Top Horizontal Axis Series 1 Styles Watermark Fields to group (For one chart per run, select Sport Association Number of Teams Group by Sp	none): All ort ear
---	--------------------------

Figure 6-42 - Sport Report - Setup Chart Appearance, Grouping set to Sport

General Fort Details		Reset to Defaults
Borders, Backgrounds	Show Chart Legend	No
Grouping Legend, Data Lavout	Plot Data By:	Record
Vertical Axis Horizontal Axis Top Horizontal Axis Series 1 Styles Watermark		

Figure 6-43 - Sport Report – Setup Chart Appearance – Setting Legend, Data Layout

Settin	igs Data Layers Legend	
		Reset to Defaults
	Position	[None]
	Background Color	Use Default
	Text Color	Use Default
	Font	Use Default
	Font Size	Use Default
	Icon Width	Use Default
	Icon Height	Use Default
	Numeric Comma Separators	No

Figure 6-44 – Sport Report – Setting Legend settings (refer Report Map tool)

Let us understand the settings we have changed. The highlevel settings are easy to understand. Since we wanted to create a bar chart, we selected *Bar*. Since we wanted to compare the number of teams, we selected that field for our series, and as we wanted to know the governing body related to each bar, we chose *Association*.

Under the *General* settings, we set the width and height to the same dimensions as the maps because we want them to fit side by side. We increased the resolution to have a cleaner image. We provided a descriptive title for the chart with a reduced text size so it fits on one line and we have more room for the chart. We grouped by *Sport* to make one chart for each. We removed the chart legend because we do not need it in this case. Now that we have covered our maps and charts, we should set a title for each section. We have made a map and chart for each of the sports, but we haven't mentioned anywhere what the sport actually is. We need an object that we can put at the top of the page, with the sport in it.



Figure 6-45 – Sports Report – Data Stream Branching

In order to create this object, we will branch off into another data stream, as shown in the previous image. We will use the *Summarize* tool to group by *Sport* so that we have a single record for each. We are then going to add a *Table* tool to create our *Report* objects.

	Sport			A
Group By				Cle
Table Coofigurat	ion			
Table Width	Automatic	~		
Pivot Style	Methodology ove	er Variable		
Bar Graph	[No Valid Fields]	~	Settings
Sport	Labranov Dialda	Width	Automatic	~
Sport		Width	Automatic	~
Dynamic or	INCOMP FIGURE			
Dynamic or	Unknown Fleids	Alignment (H)	Center	×.
Dynamic or U	JINKNOWN Fleids	Alignment (H) Borders	Center None	~
Dynamic or I	Jinknown Fleids	Alignment (H) Borders Prefix	Center None	~
Dynamic or	Jinknown Fleids	Alignment (H) Borders Prefix Suffix	Center None	>
Dynamic or I	Unknown Fleids	Alignment (H) Borders Prefix Suffix Dec. Places	Center None	>

Figure 6-46 - Sport Report - grouping by sport using Table

Since we want a *Table* object for each sport, we will group by Sport. Since we are using *Sport* as a section title, we will be able to identify the sections we are interested in. Additionally, since we want to ensure that only *Sport* makes it into this table, we can deselect the *Dynamic or Unknown Fields* option.

	Reset to Defaults
Data Font	Arial
Font Size	20
Text Color	Black
Background Color	White
Alternate Color	R=240, G=240, B=240
Rows Per Group	1

Figure 6-47 - Sport Report - setting title style using Style Editor

Since this will be the title in each section, we should increase its font size.

At this point, the following image is what our data stream will look like.



Figure 6-48 - Sport Report - data stream group by each sport

We now need to create a table for the demographic information about the teams. Since this table will be in the report, it should be sorted so the consumer can easily see what's going on.

	Name		Order		
•	Sport	V	Ascending	~	
	Association	~	Ascending	~	
	League / Conference	~	Ascending	~	
	Division	V	Ascending	~	

Figure 6-49 - Sport Report - Applying sort to data

Let us sort the data like we see in the *Sort Configuration* window. This will allow the report consumer to easily scan the list for the team they are interested in.

Group By	Sport Association League / Confere Division Team City	Group by	Sport	
Table Configura	tion			
Table Width	Automatic	~		
Pivot Style	Methodology over	er Variable		
Bar Graph	[No Valid Fields]	~	Settings
Show Colum	n Headings			
 ✓ Division ✓ Team ✓ Home Park ✓ Street addree ✓ City ✓ Zip ✓ Sport Clatitude Longitude Dynamic or 	ss Unknown Fields	Borders Prefix Suffix Dec. Places Column Rules	Vone O Create	*

Figure 6-50 – Sports Report – team information using Table

Using these settings, we can create a table with the desired useful information about each team.



Figure 6-51 - Sport Report - with separate data streams

Now that we have created each of the objects for the body of our report, we need to bring them into a single data stream so we can format the layout, header, and footer.

In the *Browse* tool following the map we created, we saw a single record for each sport because that's the field we grouped by. Adding a *Browse* tool after each of the report objects would also result in the same, except that instead of the field being called *Group*, it would retain the title *Sport*. Since we know that the sport name makes a unique key in all of our data streams, we can use the *Multiple Join* tool to bring them together.
01	om	, opeoine							
		Мар	Chart	Т	able			Subtit	e
	1	Group 🗸	🗇 🏟 Sport 🗸	♦ ♦ S	port	Y	\$	Sport	~
	•	~	\$\$ V	\$		Y	\$		\sim
(Carte	sian Joins	:						
	Error	on multidim	ensional joins of r	nore than	16 R	eco	ords		
_	-		*	TIP: 10	reor	der	multip	le row	s: select, rig
-	-		*	TIP: 10	reor	der	multip	ole rows	s: select, rig
		Input	◆ Field	Туре	o reor	ter	Size	le rows	s: select, rig Rename
•	~	Input Map	◆ Field Map	Type V_String		ter	Size	83647	s: select, rig Rename
Þ	✓ ✓	Input Map Map	 Field Map BoundingRect 	Type V_String Spatial) bj	∙ •	Size 21474 53687	83647 0911	s: select, rig Rename
•	> > >	Input Map Map Map	 Field Map BoundingRect Group 	Type V_String Spatial V_String) bj	• •	Size 21474 53687 255	83647 0911	Rename
•	> > >	Input Map Map Map Chart	 Field Map BoundingRect Group Sport 	Type V_String SpatialO V_String V_String	b reord Dbj g	• 2 • 2 • 2	Size 21474 53687 255 255	83647 0911	Rename
•	> > > > >	Input Map Map Map Chart Chart	 Field Map BoundingRect Group Sport Chart 	Type V_String SpatialC V_String V_String V_String	b reord Dbj g		Size 21474 53687 255 255 21474	83647 0911 83647	Rename
•	> > > > > > > > > >	Input Map Map Chart Chart Table	 Field Map BoundingRect Group Sport Chart Sport 	Type V_String SpatialC V_String V_String V_String V_String	g Dbj g g g g		Size 21474 53687 255 255 21474 255	83647 0911 83647	Rename
•	> > > > > > > > >	Input Map Map Chart Chart Chart Table Table	 Field Map BoundingRect Group Sport Chart Sport Table 	Type V_String SpatialC V_String V_String V_String V_String V_String	Dbj g g g g g g g g g g g g g g g g g g g		Size 21474 53687 255 255 21474 255 21474	83647 0911 83647 83647	Rename
•	> > > > > > > > >	Input Map Map Chart Chart Chart Table Subtitle	 Field Map BoundingRect Group Sport Chart Sport Table Sport 	Type V_String SpatialC V_String V_String V_String V_String V_String V_String V_String	Dbj 9 9 9 9		Size 21474 53687 255 255 21474 255 21474 255	83647 0911 83647 83647	Rename Table_Spo
•	> > > > > > > > > >	Input Map Map Chart Chart Table Table Subtitle Subtitle	 Field Map BoundingRect Group Sport Chart Sport Table Sport Table 	Type V_String SpatialC V_String V_Strin	Dbj g g g g g g g g g g g g g g g g g g g		Size 21474 53687 255 255 21474 255 21474 255 21474	83647 0911 83647 83647 83647	Rename Table_Spo Subtitle_Sp

Figure 6-52 - Sport Report - Multiple Joins using Sport name

If we had named the incoming connections the way we see them on the previous image, we would see something similar to the above *Join Multiple Properties* window. Let us use the same setup as above to keep the information that will help us downstream.

=	4 of 4 Field	ds ▼ 🖋 Cell Viewer ▼ 🕴 🗍 4 re	cords displayed, 100 KB		
A	Record #	Group	Chart	Table	Subtitle_Table
-	1	Baseball	Chart - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab
	2	Basketball	Chart - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab
	3	Football	Chart - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab
	4	Soccer	Chart - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab

Figure 6-53 - Sport Report - using Browse to get data after Join

At this point, we should have five fields: The *Sport* field for our understanding, and the four report object fields use in the workflow.

Now that we have the objects for the body of our report, we need to incorporate the following into the layout:

- The sport should be at the top of each section.
- We designed the map and chart object to fit side by side, but there is no rule for where they should go.
- We have a table that can vary in length because it has one row for each team. In such a situation, the table is typically put at the bottom of the report body.

• Based on what we see here, we need two *Layout* tools. The first is a horizontal *layout* tool that will align the map and chart objects. The second *Layout* tool is to align the section title, map-chart layout, and team information table.



Figure 6-54 – Sport Report – Updated Data Stream

Here are the settings for the first *Layout* tool.

your (13) - con	ingulation		•
Layout Mode:	Each Individual Record		
Group By:	Group Sport Table_Sport Subtitle_Sport		All Clear
Include Sourc	e Fields in Output		
Layout Configura	ation		
Orientation:	Horizontal		
Layout Width:	Fixed	✓ 7 inches	
Lavout Height:	Automatic	~	
Separator	Dia-la	2 A sivele	
V Separator	Black	2 💌 pixeis	
Cell Padding:	0 🖶 pixels		
Per Column Con	figuration		
Î	Create a l	ine between ma	p and chart
Map Chart	Width:	Automatic 🗸 🗸	
Table	Alignment (V):	Middle	~
Table	Alignment (V): Alignment (H):	Middle	* *
Table	Alignment (V): Alignment (H):	Middle Center	> >

Figure 6-55 – Sport Report – using Layout to align maps, team info table etc.

A *Browse* will show us the four columns for the first *Layout* tool.

sults - Brov	wse (20)				- -
4 of 4 Field	ds 🕶 🕪 Ce	Il Viewer 👻 🕴 🕴 4 records displaye	d, 100 KB		● · ● · ● ·
Record #	Sport	Table	Subtitle_Table	Layout	~
1	Baseball	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
2	Basketball	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
3	Football	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
4	Soccer	Table - View Browse Tool Report Tab	Table - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	

Figure 6-56 - Sport Report - Data after applying Layouts

The image below shows the settings for the second *Layout* tool.

Layout Mode:	Each Individual Recon	d	
	Sport		All
Group By:			Clea
Include Sour	ce Fields in Output		
Layout Configu	ration		
Orientation:	Vertical Merge Conte	ents to Line up Table Columns	
Layout Width:	Percentage	✓ 45 Percent	
Layout Height:	Automatic	~	
Border	Black	2 🌲 pixels	
Separator	White	1 🌲 pixels	
Cell Padding:	0 🌩 pixels 🕨	Add a border	
Per Row Config	guration		
Layout	Height:	Automatic 🗸 🗸	
Subtitle_Ta	Alignment (V):	Middle	~
	Alignment (H):	Center	\sim
	Fill Color		
		Ensure the ord	er

Figure 6-57 – Sport Report – applying Layout configuration

Rerunning the flow now would show the two columns in the second *Layout* tool.

Res	ults - Brov	vse (21)	
=	2 of 2 Field	ls 🔻 🖌 🛛 C	Cell Viewer 👻 👔 🌡 🛛 4 records displayed, 122 KB
Ä	Record #	Sport	Layout
	1	Baseball	Layout - View Browse Tool Report Tab
	2	Basketball	Layout - View Browse Tool Report Tab
	3	Football	Layout - View Browse Tool Report Tab
	4	Soccer	Layout - View Browse Tool Report Tab

Figure 6-58 - Sport Report - selected data after applying Layout



The data stream after having the Layouts in place would look like the below.

Figure 6-59 - Sport Report - data stream with Layouts in place

Now that we have combined all of the elements of the body of the report into a single report object, we can add a *Report Header* and *Report Footer* to the report. We will use the default settings for both because the client wants to see what would typically be there. We are ready to render the report, but before we do that, let us see what the data we are feeding into the *Render* tool looks like.

Res	ults - Brov	vse (22)				▼ 무 🗙
=	4 of 4 Field	is • 🗸 🛛 (Cell Viewer 🔹 🕴 👃 4 records displa	ayed, 520 KB		● • □ • □ •
Ä	Record #	Sport	Layout	Header	Footer	~
	1	Baseball	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
	2	Basketball	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
	3	Football	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	
	4	Soccer	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	Layout - View Browse Tool Report Tab	

Figure 6-60 – Sport Report – Data selected after adding Report Header and Report Footer

As we can see, there is no special format for the *Header* and *Footer* fields. This means we could put any object into the *Header* and *Footer* options when we render the report. However, the *Report Header* and *Report Footer* tools create appropriately sized objects that are designed for standard report information.

Render Configuration	
Output Mode: Choose a	Specific Output File
Output File: Sport rep	port.pdf
Group Data Into Separa	ite Reports
Field To Group On:	Sport Specify output filer
Modify Filename By:	Appending Group To Filename
Report Data	
Data Field Layout	•
Separator Insert Section	n Breaks Between Records
V Header Header	
Eooter	
He Show hea	Header and footer input
Report Style	
Paper Size Letter (8.5"	x 11") 🔻
Orientation () Portrait ()	Landscape
Margins (in.) 0.5 inch Ma	rgins
	Edit Default Styles

Figure 6-61 - Sport Report - applying Render to specify path

In the *Render Properties* window, we will specify that we are creating an output file, and specify a path for it. Since we want to start each sport on a new page, we will use the section break option. Finally, we will turn the Header and Footer fields on and define them. The Major Sports Teams data stream should look like the following when its complete.



Figure 6-62 - Sport Report data stream on completion

6.11 Football!

То	Alteryx Consultants
Subject	FW: FOOTBALL!
Hey,	
The sports draft of the long as we following : breaking it Additionall summarize	bar just called me back. They said that the first e report shows promise, and are likely to use us, as can quickly modify the report to show the Limit to a report about Football. Instead of down by sport, break it down by the association. y, instead of summarizing the chart by association, it by division.
I am in mee show them make these	etings for the rest of the afternoon and I cannot the turn around, but, I think you know enough to e changes.
If we can g would guar	et it back to them in the next 20 minutes, I think it rantee that they go with us.
Thanks.	

CHAPTER 7

Expensive Beauty Products

То	Alteryx Consultants
Subject	Expensive Beauty Products
Attached	Body Care Co.csv
Hey,	
See the ema	ail below from Jan at Body Care Co.
Body Care (tend to ask	Co. is a company that I work with pretty regularly. They have Alteryx and us to work on projects when they don't have the bandwidth.
I think worl topics.	king together on this would be a good way to introduce you to a few new
Thanks.	
I've got son	nething fairly straightforward for you this time.
I am using o swamped, s would be g	lata that looks like the attached data set in order to run a lot of analyses. I'm so if you could build me an Alteryx flow that I can use going forward, that reat!
I am using a running too	a dashboarding tool that needs to be highly responsive. Unfortunately, it is slow because I don't have an easy way to identify the selected data.
Below is the Would you	e logic to create a flag for a number of different things I need to look at. be able to create a process so the data can be modified to have these flags?
1) Lotions (2) Moisturi 3) Baby Pro	"Lotion" is in the description) zing Products ("Moisturizing" is in the description) oducts ("Baby" is in the description)
4) Acne Pro 5) Create gi looking into	ducts ("Anti-Acne" or "Anti-Oil" is in the description) roups for all products that are named 90% identically by word (we are o similarly named products to see how their sales correlate).
Thanks in a Jan	dvance,

7.1 Tools & Concepts

Tools	<u>Concepts</u>
Fuzzy Match	Fuzzy Logic
Record ID	Regular Expressions
RegEx	Pre-Calculation
Unique	Data Flags
Tool Container	Removing Duplicates
	Indexing
	Group Tools
	Organize Module

7.2 Fuzzy Match

	The <i>Fuzzy Match</i> tool allows us to apply fuzzy logic to a data stream to match similar items.			
	Group	Input	Output	
Figure 7-1 – Fuzzy Match	Join	See	See	

Note: The *Fuzzy Match* tool has multiple different methods of determining what matches. If the reader is not familiar with fuzzy logic, then it is suggested to research these techniques in order to understand how each works.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: Any data stream with a text field

Output: A data stream has the key field and selected *Fuzzy Match* information fields.

Properties Window:

The *Fuzzy Match Configuration* window has the following five core components.



Figure 7-2 – Fuzzy Match Configuration

- *Merge/Purge Mode* allows us to decide if we need to compare all of the records in the list, or to compare records that have different sources. This means if we have data that is broken into logical buckets, we can compare across those buckets without comparing within them.
- *Record ID Field* is the field we will match on.

- *Match Threshold* is the minimum match percentage to have a positive result.
- *Match Fields* allows us to select fields and methodologies to run the analysis for matches.
- *Advanced Options* allows us to modify the output of the *Fuzzy Match*.

Match Style	Edit Match Options		
Custom	Match Style		
Address	Custom Celete		
Address No Suite	Preprocess		
Company Name	[None]		
Custom	Generate Keys		
Name	Whole Field (Case Insensitive)		
Name w/ Nicknames Phone ZIP Code	Generate Keys for Each Word Don't Generate Keys for the following words		
	Don't Generate Keys for Single Letter Words Ignore if Empty Maximum Key Length: 100 Match Function		
	Words & Digits: Levenshtein Distance 🔹		
	When Using Word Based Match, also use:		
	Character (No Spaces)		
	Word Frequency Statistics (Word Match Only)		
	[None]		
	Nickname/Abbreviation Table (Word Match Only) Penalty		
	Match Threshold: 90 * % Match Weight: 100 *		
	OK Cancel Help		

Figure 7-3 – Fuzzy Match Configuration Edit Match options

Under *Match style,* there are some predefined matching methods. These match styles are Alteryx's suggested fuzzy logic methodologies for the type of things they name.

Custom allows us to define a methodology, which can be saved as a permanent setting if we are an administrator or super user.

The *Edit* option under the *Match Fields* section opens the *Match Options* window and offers more customization. There are many options here that allow us to define a matching algorithm. Many of the options are self-explanatory. However, if we are not familiar with fuzzy logic, we should research Levenshtein Distance and Jaro Distance before designing our own match methodology.

7.3 Record ID

5	The <i>Record ID</i> tool creates an index column in our data stream.			
3	Group	Input	Output	
Figure 7-4 - Record ID	Preparation	Any data stream	See below	
Note: This is useful for creating an index				

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: The original data stream with a new field, which numerically identifies each record.

Properties Window:

The *Record ID Configuration* window has four basic components:

6 F	tecord ID (17) - Configuration 👻 무 🗙	\$
%	Starting Value	
•	Field Name Group ID	
8 1	Field Type Size Int32 ●	
0	Field Position First Field	

Figure 7-5 – Record ID Configuration

- *Starting Value* is the number we want to assign to the first record (which will be incremented following that record).
- *Field Name* is the resulting field name for our index.
- *Field Type* allows us to select the type and size of the resulting field.
- *Field Position* allows us to set the new field to the beginning or end of the dataset.

7.4 RegEx



The <i>RegEx</i> tool allows us to run
regular expressions on a field in our
data stream.

	Group	Input	Output
RegEx	Parse	See below	See below

Note: Alteryx uses the Perl regular expression syntax. See Appendix I.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Input: A data stream with a string field.

Output: The original data stream and additional columns or rows depending on the selections made in the tool.

Properties Window:

The *RegEx Configuration* window has five basic components:

F	Field to Parse		
	Test		•
F	Regular Expression		
1	s		▼
	Case Insensitive		
(Dutput Method		
	Replace		•
ſ	Properties		
	Replacement Text		
		[▼
	Conv Upmatched Text to Output		
	Copy oninatched Text to Output		

Figure 7-7 – RegEx Configuration Replace method

- *Field to Parse* is the text field we want to manipulate.
- *Regular Expression* is the expression we use in order to identify sections of the string.
- *Case Insensitive* allows us to specify if case is important to our search.
- *Output Method* allows us to select what the regular expression will result in.
- *Properties* define specific options depending on which *Output Method* is selected.

Replace allows us to replace the text every time the *Regular Expression* is found in the *Replacement Text* string.

Copy Unmatched Text to Output will maintain any text before, between, or after the occurrences of the *Regular Expression*. If this option is not checked, the output will only have the *Replacement Text* as many times as the mentioned *Regular Expression* is found. In the above case, every time a space (\s) is in the field, it is replaced by the string characters "- -" (two hyphens).

R	egEx (2) - Configurati	ion 👻 무 :
	Field to Parse	
2	Test	-
	Regular Expression	
1	*[2/*]	
	Case Insensitive	
	Output Method	
	Tokenize	-
1	Properties	
	Split to Columns	
	# of Columns:	3
	Extra Columns:	Drop Extra with Warning 👻
	Outsuit Baset Manage	Test
	Output Root Name:	TOSE

Figure 7-8 – RegEx Configuration Tokenize method

Tokenize allows us to split the field we have, into an additional column or row every time we find the *Regular Expression* in the field.

In this case we are creating a new record for every word in a field. Since the regular expression will look for the longest occurrence of the format we need to specify that we do not want any spaces in the unknown characters between the beginning and ending of a word.

We can see that $\langle [^{s}]^* \rangle$ looks for *words* and creates a new record with all other data replicated for each of the resulting words. This is because:

\< identifies the beginning of a word</pre>

[^\s] means any character that is not a space

* means repeat the previous thing 0 or more times

\> identifies the end of a word

Field to Parse				
Pho	ne Number (xx	x) xxx - xxxx		
Regu	lar Expression	1		
(\()()(\/s)()(z\-\s)()				
7 C	ase Insensitive	•		
Dutp	ut Method			
Pars	e			
Prov	perties			
FIO	penies			
Out	tput Fields			
	Output Field	Туре	Siz	e Expression
1	Output Field RegExOut1	Type String	Siz	e Expression (\()
1 2	Output Field RegExOut1 RegExOut2	Type String String	▼ 16 ▼ 16	e Expression (\() ()
1 2 3	Output Field RegExOut1 RegExOut2 RegExOut3	Type String String String	Siz ▼ 16 ▼ 16 ▼ 16 ▼ 16	e Expression (\() () (\)\s)
1 2 3 4	Output Field RegExOut1 RegExOut2 RegExOut3 RegExOut4	Type String String String String	Siz • 16 • 16 • 16 • 16 • 16	e Expression (\() () (\)\s) ()
1 2 3 4 5	Output Field RegExOut1 RegExOut2 RegExOut3 RegExOut4 RegExOut5	Type String String String String String	Siz • 16 • 16 • 16 • 16 • 16 • 16	e Expression (\() () (\)\s) () (\s)-\s)

Figure 7-9 – RegEx Configuration Parse method

Parse allows us to define *Marked Groups* in order to break down fields with specific syntax into new columns, which can be named in the *Output* fields section.

In this example image, the tool isolates the numbers from a list of phone numbers formatted as (*XXX*) *XXX* – *XXXX*.

Marked groups are identified by parentheses. The marked group (X) would select X and create a column.

The expression is made of six marked groups:

(\() Since "\(" means the "(" character, the first group is an opening parenthesis.

(...) Since "." means any single character the second group is the three characters following the previous group (which would be the first three digits of the phone number).

 $(\)\s)$ Since "\)" means the ")" character, the third group is the closing parenthesis and space before the second set of numbers.

(...) As seen earlier, it represents the next three characters. So we have the second three digits of the phone number in the fourth group.

 $(\s-\s)$ Since "-" is literally the hyphen character, the fifth group is the "space hyphen space" between the second and third set of digits.

(....) gets us the last four digits of the phone number.

egex (2) - Configuration	▼ ₽
Field to Parse	
Test	
Regular Expression	
.bhow.	
Case Insensitive	
Output Method	
Match	
Properties	
Match Status Field Name	
Test_Matched	
Error if not Matched	

Figure 7-10 – RegEx Configuration Match method

Match creates a Boolean field such that anytime the *Regular Expression* is in the field, the result is true else it is false.

In this case we have the regular expression, .*world.*. This means that the word *world* with any characters before or after it.

The new field created is called *Test_Matched*, which can be used downstream.

Figure 7-11 - Unique	The <i>Unique</i> too separate uniqu records based select.	ol allows us le or duplica on the field	to ate (s) we
	Group	Input	Output
	Preparation	Any data stream	See below

Note: Unique records are kept if the record is the first to pass through the data stream. Unique differs from grouping by the same fields in a *Summarize* tool. It keeps all of the data associated with the record instead of just the used fields.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output U: a data stream where each record is unique based on the fields that were selected.

Output D: a data stream where each record is a duplicate of the one in the U (unique) output based on the fields that were selected.

Properties Window:

The *Unique Configuration* window is a list of each of the fields in the dataset.

🛞 U	Inique (7) - Configuration	- ₽ X
	Unique Fields	
\odot	✓ Test1 Test2	All
•		Clear
-		
->		
0		
0		
0		

Figure 7-12 – Unique Configuration

We need to select all those we want to use to identify unique records.

7.6 Tool Container

Test Container ()	The <i>Tool Container</i> allo group tools or segmen stream for easy identif functionality within th It helps organize the w modules.	r allows gments o lentificat nin the w the work	s us to of a data ation of workflow. rkflow into	
Figure 7-13 – Tool Container	Group	Input	Output	
	Documentation	None	None	

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

It can either be enabled or disabled when the flow is executed

Properties Window:

The *Tool Container Configuration* window contains all the settings and format options for the container

car Car	Container (7)	- Configuration	· 무 › ID
Tes	Test Container		7
			Reset to Defaults
	Text Color	R=49, G=76, B=74	
	Fill Color	R=236, G=242, B=242	
	Border Color	R=49, G=76, B=74	
	Transparency	25	 ▼
	Margin	Medium	-

Figure 7-14 - Tool Container Configuration

Caption provides a title for the container

Color formats can be specified for the container, text within the container, and border.

The Margin option allows the user to set the size of the box.

The *Disabled* checkbox allows the user to control the execution of the flow or functions within the container. If the checkbox is selected, the data stream does not execute any tools in the container. The same can be controlled using an *Action* tool.

7.7 Expensive Beauty Products

То	Alteryx Consultants
Subject	Expensive Beauty Products
Attached	Body Care Co.csv
Hey,	
See the em	ail below from Jan at Body Care Co.
Body Care (tend to ask	Co. is a company that I work with pretty regularly. They have Alteryx and us to work on projects when they don't have the bandwidth.
I think worl topics.	king together on this would be a good way to introduce you to a few new
Thanks.	
I've got son	nething fairly straightforward for you this time.
I am using o swamped, s would be g	data that looks like the attached data set in order to run a lot of analyses. I'm so if you could build me an Alteryx flow that I can use going forward, that reat!
I am using a running too	a dashboarding tool that needs to be highly responsive. Unfortunately, it is o slow because I don't have an easy way to identify the selected data.
Below is the Would you	e logic to create a flag for a number of different things I need to look at. be able to create a process so the data can be modified to have these flags?
 Lotions (Moisturi Baby Pro Acne Pro Create ge looking into 	"Lotion" is in the description) zing Products ("Moisturizing" is in the description) oducts ("Baby" is in the description) oducts ("Anti-Acne" or "Anti-Oil" is in the description) roups for all products that are named 90% identically by word (we are o similarly named products to see how their sales correlate).
Thanks in a Jan	dvance,

The problem that Jan is describing is important to be aware of when we create data for dashboard and reporting tools. These tools are expected to create visuals in real time based on dynamic user inputs. Ideally, the performance is fast enough that the dashboard consumers do not observe a lag in the data load.

In order to achieve this goal, dashboard creators need well-designed and usually tidy data. This means that as much should be built into the data preparation process as possible. In Jan's case, they are doing some complex string calculations in order to filter the data appropriately.

Best practice is that any calculation the dashboard creator needs to run at a record level should be done in the data preparation stage.

To improve performance, pre-calculation steps are performed and it is a fundamental part of good data preparation.

Even when we are not preparing the data for analytics tools, it is important to pre-calculate fields the end user may need. This is because building and vetting a calculation once in a workflow means there will be a significant reduction in risk caused by human error downstream.

Let us imagine that we work for a large retailer that wants to know the profit ratios for each transaction. We may have 20 different analysts working on this project and reproducing the results every month. If these analysts are each rewriting the same formula (*Profit/Sales*) every month, it means we need to make sure they do not accidently type (*Sales/Profit*) in 240 different formulas per year. Instead, if we know they need transaction level profit ratio, we can create a standard formula in the data preparation process that writes this value into the source data that they all use.

Besides the pre-calculation, we can see from this email that Jan is only interested in a module. This is the case because they work with Alteryx and will do some modifications, like replacing the data connection and outputs. We will often come across things like this when we are supporting other Alteryx users because they either do not know how or do not have time to build a portion of a module.

Considering the information in the email, it seems Jan was talking about flagging item names that fall into each of those buckets. We will use two techniques to achieve this. For requirements one through four, we will write regular expressions that create flags for matches and mismatches. And for the fifth, we will use fuzzy logic to define grouping for each of the tools.

Regular expressions are pattern-matching formulae that allow us to define a pattern to look for in a string. Alteryx uses the Perl programing language syntax for its regular expressions, so any documentation related to Perl regular expressions will help us get up to speed. A cheat sheet is provided at the end. (See Appendix I.)

Fuzzy logic is essentially a logical process that allows computers to say, "Eh, close enough!" As far as Alteryx is concerned, this is limited to string parsing; however, more generally this extends to programming a definition of *tall*, *heavy*, and *blue*. There are various methodologies of matching strings. Alteryx uses the *Levenshtein Distance* and *Jaro Distance* to achieve this. The nuances of the differences are well documented and it is suggested that the reader learn more about them if they plan to use Alteryx's *Fuzzy Match* tool.



Figure 7-15 – Expensive Beauty Products Body Care Company data

If we connect to the file in *Chapter 7 -- Expensive Beauty Products > Body Care Co.csv*, we see that Alteryx is reading all of the fields in this file as strings.

_	_	1					
		Field	Туре		Size	Rename	Description
▶		Date	V_String	-	254		
	V	Store Number	V_String	-	254		
		Item Name	V_String	-	254		
		Unit Cost	V_String	•	254		
		Average Units Sold	V_String	-	254		
		Units Sold Per Day	V_String	-	254		
	V	*Unknown	Unknown	-	0		Dynamic

Figure 7-16 – Expensive Beauty Products Select Configuration

Ideally, we should change the last three fields to numeric. But since we are only going to use *Item Name* and we don't know what the rest of Jan's process looks like, we should leave the fields the way they are.

Since we need to make sure Jan knows what each piece of this module is doing, we have to make sure that it is well documented. The first thing we will do is put our initial connection into a tool container that indicates that the entire set of tools should be replaced when the connection is changed.



Figure 7-17 – Expensive Beauty Products Input data – to be replaced in a container

Let us take a look at the contents of this file.
6 of 6 Fiel	ds 🔻 🖋	Cell Viewer 🔻	188,868 records displayed, 2.8 MB			D - D -	• 🛨 •
Record #	Date	Store Number	Item Name	Unit Cost	Average Units Sold	Units Sold Per Day	
1	1/1/2014	1	Vitamin Restoring Anti-Aging Cream 1 Ounce	1.99	0	3	
2	1/1/2014	1	Anti-Itch Cream 1 Ounce	1.99	0	3	
3	1/1/2014	1	Anti-Itdh Lotion 5 Ounce	1.99	1	2	
4	1/1/2014	1	Natural Protection Sunscreen SPF 50 3 Ounce	10.99	0	3	
5	1/1/2014	1	Baby Calming Bubble Bath 8 Ounce	3.99	0	3	
6	1/1/2014	1	Baby Lotion 8 Ounce	2.99	0	2	
7	1/1/2014	1	Baby Lotion 20 Ounce	5.99	0	1	
8	1/1/2014	1	Baby Soothing Anti-Itch Cream 3 Ounce	3.99	0	3	
9	1/1/2014	1	Baby Wash And Shampoo 18 Ounce	2.99	0	2	
10	1/1/2014	1	Lavender Body Wash 14 Ounce	2.99	0	2	
11	1/1/2014	1	Stress Relief Body Wash 10 Ounce	2.99	0	1	
12	1/1/2014	1	Anti-Acne Medicated Body Cream 2 Ounce	11.99	0	2	

Figure 7-18 – Expensive Beauty Products Input data – Browse configuration

Next, let us create flags using regular expressions. We will create a new *Tool Container* called Regular Expressions with a *RegEx* tool in it.



Figure 7-19 – Expensive Beauty Products RegEx – "lotion" flag – Match method

The first requirement we want is a flag to identify if *lotion* is in the *Item Name* field. We can use the expression *.*lotion.** to identify any field that has the string "lotion" in it.

This regular expression checks to see if the literal pattern "0 or more characters followed by 'lotion' followed by 0 or more characters" is in each record and returning true or false.



Figure 7-20 – Expensive Beauty Products RegEx – "moisturizing" flag – Match method

A familiarity with Body Care Co's data lets us know their item names always have spaces between words. This means we can actually make our expression more specific as shown in the image, with .*\<*moisturizing*\>.*

This regular expression reads "0 or more characters followed by the word 'moisturizing' followed by 0 or more characters."

The special characters \< and \> in regular expressions take into account specific characters preceding or ending a word to identify if the entire word is matched.

Following the same pattern, we can create a field called *Baby Product*.



Figure 7-21 – Expensive Beauty Products RegEx – "baby" flag – Match method

We then make another for acne products.



Figure 7-22 – Expensive Beauty Products RegEx – a complex flag – Match method

Our last flag is more complicated, because we have two different things that could define acne products.

We will combine what we know about Boolean statements and regular expressions to create the expression

```
(.*anti\-acne.*) | (.*anti\-oil.*)
```

This regular expression is checking to see if the pattern "0 or more characters followed by 'anti-acne' followed by 0 or more characters, or 0 or more characters followed by 'anti-oil' followed by 0 or more characters' exists.

As we can see, regular expressions can get complex, but there are always multiple ways to create them. In this case, we could have also used .**anti\-(acne|oil)*.*, where it would have looked for the pattern "0 or more characters followed by 'anti-' followed by 'acne' or 'oil' followed by 0 or more characters".

1	0 of 10 Fi	elds 💌 🐦	Cell Viewer *	11	188,858 records displayed, 2.9 MB								+
R	tecord #	Date	Store Number		Item Name	Unit Cost	Average Units Sold	Units Sold Per Day	Lotion	Moisturizing	Baby Product	Acne Product	
1		1/1/2014	1	Vitamin	Restoring Anti-Aging Cream 1 Ounce	1.99	0	3	False	False	False	False	
2		1/1/2014	1	Anti-Itch	Cream 1 Ounce	1.99	0	3	False	False	False	False	
3		1/1/2014	1	Anti-Itdl	h Lotion 5 Ounce	1.99	1	2	True	False	False	False	
4		1/1/2014	1	Natural	Protection Sunscreen SPF 50 3 Ounce	10.99	0	3	False	False	False	False	
5		1/1/2014	1	Baby Ca	Iming Bubble Bath 8 Ounce	3.99	0	3	False	False	True	False	
6		1/1/2014	1	Baby Lo	tion 8 Ounce	2.99	0	2	True	False	True	False	
7		1/1/2014	1	Baby Lo	tion 20 Ounce	5.99	0	1	True	False	True	False	
8		1/1/2014	1	Baby So	othing Anti-Itch Cream 3 Ounce	3.99	0	3	False	False	True	False	
9		1/1/2014	1	Baby Wa	ash And Shampoo 18 Ounce	2.99	0	2	False	False	True	False	
1	0	1/1/2014	1	Lavende	r Body Wash 14 Ounce	2.99	0	2	False	False	False	False	
1	1	1/1/2014	1	Stress R	elief Body Wash 10 Ounce	2.99	0	1	False	False	False	False	
1	2	1/1/2014	1	Anti-Acn	e Medicated Body Cream 2 Ounce	11.99	0	2	False	False	False	True	

Figure 7-23 – Expensive Beauty Products Boolean filed (flags) for RegEx Match

We can see that we have the four Boolean fields to identify each of the four sets of products we have defined.



Figure 7-24 – Expensive Beauty Products Adding comments to indicate contents in Tool Container

We can now add a *Comment* tool that describes what is happening in the *Tool Container* and move onto the fuzzy logic.



Figure 7-25 – Expensive Beauty Products Updated data stream

We need to define the groups of products with similar names. The first thing we should do is to create a unique list of product names. Let us introduce a *Summarize* tool within a container and connect it to data input stream.



Figure 7-26 – Expensive Beauty Products second data stream for fuzzy match

Date V_String Store Number V_String
Store Number V_String
Item Name V_String
Unit Cost V_String
Average Units S V_String
Field Action Field Name
▶ Item Name GroupBy ▼ Item Name

Figure 7-27 – Expensive Beauty Products Summarize configuration

Let us run the workflow and take a look at the unique list.

Res	ults - Brov	wse (13) 👻 🖓	×							
	1 of 1 Field	ds 🔻 🐦 🛛 Cell Viewer 👻 👔 🛊 拱 🕇	*							
Ä	Record #	Item Name	*							
	1	1% Hydrocortisone Cream 1 Ounce								
	2	After Shave Lotion 4 Ounce								
	3	Anti-Acne Body Cream 12 Ounce								
	4	Anti-Acne Facial Bar 4 Ounce								
	5	Anti-Acne Facial Cream 4 Ounce								
	6	Anti-Acne Medicated Body Cream 2 Ounce								
	7	Anti-Acne Medicated Foam 7 Ounce								
	8	Anti-Acne Medicated Pads 30 Sleeves								
	9	Anti-Acne Moisturizing Cream 4 Ounce								
	10	Anti-Itch Cream 1 Ounce								
	11	Anti-Itch Creem 1 Ounce								
	12	Anti-Itch Hand Cream 3 Ounce								

Figure 7-28 – Expensive Beauty Products Browse configuration after Summarize

Now that we have a unique list of products, we can determine which of those products have similar names.



Figure 7-29 – Expensive Beauty Products Fuzzy Logic configuration

If we add a Fuzzy Match tool with the above settings, we know we are matching *Item Name* at a minimum 90% threshold. The settings here define a matching algorithm using a word and digit-based *Levenshtein distance method*. This algorithm is looking at the whole field by word (except for single-character word; up to 100 words per field) and keeping anything that it finds with at least a 90% match.

Because of the nature of this methodology, we will introduce an issue in the output data stream. That will result in duplicate records for the fields that match based on multiple keys. In order to correct this, we will isolate the unique records to be kept.

If we add a *Unique* tool after the *Fuzzy Match*, we can see a new field has been created called *Item Name2*. The new field allows us to see which entries match.





A quick run using the *Browse* tool shows us the matched items.

:=	2 of 2 Field	ds ▼ 🖌 🛛 Cell Viewer ▼ 🕴 🖡	🕛 • 🕞 • 🖃 •	>>
ä	Record #	Item Name	Item Name2	
	1	Anti-Itch Cream 1 Ounce	Anti-Itch Creem 1 Ounce	
	2	Anti-Itch Lotion 5 Ounce	Anti-Itdh Lotion 5 Ounce	
	3	Baby Siothing Bubble Bath 4 Ounce	Baby Soothing Bubble Bath 4 Ounce	
	4	Baby Sunscreen SPF 55 4 Ounce	Sunscreen SPF 55 4 Ounce	-

Figure 7-31 – Expensive Beauty Products Viewing matching items



Figure 7-32 – Expensive Beauty Products Record ID for matched items

We can see by looking at the resulting data stream that we have four groups of items, and three of those are a result of misspellings in the data.

What we can do is provide a numeric grouping for each of these matches, which will solve Jan's problem of data comparison. Since each of the groups is unique, we can use the *Record ID* field to identify them.

Now we can add a comment that will allow Jan to understand what we did, and then we will be ready to combine the data streams.



Figure 7-33 – Expensive Beauty Products Adding comment to indicate the process

In order to combine the data streams, we need to join the data twice so that each part of the group can be flagged with the *Group ID*.



Figure 7-34 – Expensive Beauty Products First Join for left group members The *Join* tools settings are shown below.



Figure 7-35 – Expensive Beauty Products Join configuration for "Item Name"



Figure 7-36 – Expensive Beauty Products Second Join for right group members

	Le	ft		Right			
1	It	em Name	- 4	Item Name2			
*			* **			-	
				💛 Joi	in	on Ite	em Nan
	_		- (_	
	L		D(R	
Op	tions	+ 1	1	TIP: To reord	ler i	multiple	e rows: sele
			•				
		Input	Field	Туре		Size	Rename
	V	Left	Unit Cost	V_String	•	254	
	V	Left	Average Unit	V_String	•	254	
		1.0	Units Sold Pe	V_String	•	254	
		Left					
	V	Left	Lotion	Bool	•	1	
	 <	Left Left Left	Lotion Moisturizing	Bool Bool	•	1	
	 <	Left Left Left Left	Lotion Moisturizing Baby Product	Bool Bool Bool	• •	1 1 1	
		Left Left Left Left Left	Lotion Moisturizing Baby Product Acne Product	Bool Bool Bool Bool	• • •	1 1 1 1	
		Left Left Left Left Left Left	Lotion Moisturizing Baby Product Acne Product Group ID	Bool Bool Bool Bool Int32	• • •	1 1 1 1 4	
		Left Left Left Left Left Left Right	Lotion Moisturizing Baby Product Acne Product Group ID Group ID	Bool Bool Bool Int32 Int32	• • • •	1 1 1 4 4	Right_Gro
		Left Left Left Left Left Left Right	Lotion Moisturizing Baby Product Acne Product Group ID Group ID Item Name	Bool Bool Bool Int32 Int32 V_String	• • • •	1 1 1 4 4 254	Right_Gro
		Left Left Left Left Left Left Right Right	Lotion Moisturizing Baby Product Acne Product Group ID Group ID Item Name Item Name2	Bool Bool Bool Int32 Int32 V_String V String	• • • •	1 1 1 4 4 254 254	Right_Gro Right_Iter

Figure 7-37 – Expensive Beauty Products Join configuration for "Item Name2"

12 of 12 Fie	lds 🔻 🎺	Cell Viewer	▼ 1 ↓ 18	8,868 records disp	played, 1.5 MI	B 🚺 - 🔂 - 🖸
Record #	Lotion	Moisturizing	Baby Product	Acne Product	Group ID	Right_Group ID
1	False	False	False	False	[Null]	[Null]
2	False	False	False	False	[Null]	[Null]
3	False	False	False	False	[Null]	[Null]
4	False	False	False	False	[Null]	[Null]
5	False	False	False	False	[Null]	[Null]
6	False	False	False	False	[Null]	[Null]
7	False	False	False	False	[Null]	[Null]
8	False	False	False	False	[Null]	[Null]
9	False	False	False	False	[Null]	[Null]
10	False	False	False	False	[Null]	[Null]
11	False	False	False	False	[Null]	[Null]
12	Fal	53/50	5-1	\sim		
10000		-			Cardina - martine -	antino, antino, an
43030	Irue	False	False	False	2	[Null]
43031	True	False	False	False	2	[Null]
43032	True	False	False	False	2	[Null]
43033	True	False	False	False	2	[Null]
43034	True	False	False	False	2	[Null]
43035	True	False	False	False	2	[Null]
43036	True	False	False	False	2	[Null]
43037	True	False	False	False	2	[Null]
12028	Te	Falso	Cal.	~		
000944	Taise	Edise	- Inte	Faise	- INTERIO	
88845	False	False	True	False	[Null]	3
88846	False	False	True	False	[Null]	2
88847	False	False	True	False	[Null]	3
88848	False	False	True	False	[Null]	3
88849	False	False	True	False	[Null]	3
88850	False	False	True	False	[Null]	3
88851	False	False	True	False	[Null]	3
88852	False	Falce	True	Falca	[Null]	3
88853	False	Falce	True	Falca	[Null]	3
88854	Falca	Falca	True	Falca	[Noll]	2
	I DINE	I DINE (Contraction Contra	T T T N LI T T	Const.

Figure 7-38 – Expensive Beauty Products Browse configuration after Joins

If we applied the previous settings along with the default *Union* settings, we should see that we have two new fields at the end of the data stream. We need to combine these fields in order to give Jan a simple dataset to work from.





We can create a conditional formula called *Similar Products* that will bring the fields together:

IF IsNull([Group ID]) AND IsNull([Right_Group ID]) THEN Null() ELSEIF IsNull([Group ID]) THEN [Right_Group ID] ELSE [Group ID] ENDIF





In order to finish this data stream, we can add a select statement that has the two *Group ID* fields removed. If we then use a *Tool* container, we can very clearly show where any downstream tools should be connected.

The *Expensive Beauty Products* data stream should look like the following image when it is complete.



Figure 7-41 – Expensive Beauty Products Complete data stream

7.8 More Flags

То	Alteryx Consultants					
Subject	More Flags					
Hey,						
Check out th	Check out the response we got!					
Think you ca	n handle it?					
Hey,						
That last dat the analysis.	aset you created was great. I have a few more items that I need to add to					
Can you add	a flag for the following?					
1) All sun pro 2) All itch rel "hydrocortis 3) All shamp 4) All conditi	otective products (includes "SPF" in the description) ief (includes "anti-itch", "anti-itdh" (I noticed an issue with the data), or sone" in the description) oos (includes "shampoo" in the description) oners (includes "conditioner" in the description)					
Thanks agair Jan	۱,					

Chapter 8 Applications Wanted

То	Alteryx Consultants
Subject	Let's Build An App For That!
Attached	All Recorded Traffic Tickets.csv
неу,	
I need to pu	Ill you in on a project with the Baltimore Ticket Team.
They have a	asked us if we can build an application that will allow them to do the following:
1. select	t a file with violations
2. select	t a date range to limit the data
3. have	it create output file(s) of their choice
a.	in-state plates
b.	out-of-state plates
C.	in-state and out-of-state plates (as a single or multiple files).
I think we w	vill hear back from them soon if we build it to these exact standards. We should design
this analytic	application such that the default is to select Maryland, yet it also gives the flexibility
to select an	y state. Also include a text box so they can limit to a specific location of interest.
Should be f	un.

=

8.1 Tools & Concepts

In this chapter, we will be covering the following Tools and Concepts:

<u>Tools</u>	<u>Concepts</u>
Action	Building Applications
Condition	
Date	
Date Filter	
Drop Down	
File Browse	
Radio Button	
Text Box	

8.2 Action

	The <i>Action</i> tool modifies the values in other tools.				
	Group	Input	Output		
Figure 8-1 – Action	Interface	None	None		

Application questions can be connected to the top *Black Question Anchor*.

A condition tool can be connected to the *Diamond Anchor*.

The *Lightning Bolt Anchor* can be connected to a tool to update that tool based on the *Action* tool's settings.

Action tools are special tools that update settings for the connected tools in the workflow.

Properties Window:

The Action Configuration window has only one component.



Figure 8-2 – Action configuration with list of Action types

Selecting an action type allows us to configure the changes in the tool connected to the action.

The list of options appear depending on the tool that is connected, and we can see a list of tools that behave in the same way as seen in Appendix J.

The image is a comprehensive list of action types. The methods are self-explanatory, and we will almost always use the default action associated with each tool.

8.3 Condition

	The <i>Condition</i> trigger differ on the condi	n tool allow cent results tion in the t	s us to depending ool.				
\	Group	Input	Output				
Figure 8-3 - Condition	Interface	None	None				
At least one Action tool s	At least one <i>Action</i> tool should always follow this tool.						
Multiple Application questions can be connected to the top <i>Black Question Anchor</i> .							
Diamond Anchor T connec	cts to an Actio	n tool and p	provides				

True result if the condition is true.

Diamond Anchor F connects to an Action tool and provides False result if the condition is false.

Properties Window:

The *Condition Configuration* window allows us to create a Boolean expression using the connected questions.



Figure 8-4 - Condition Configuration

In this case, since connections 1 and 2 are both *Radio Button* tools, their values are either true or false.

The expression is true if either of these options were selected.

• If the expression is true, then the true (T) path has a true value passed down it, and the false (F) path has a false value passed down it.

• If the Expression is false, then the true (T) path has a false value passed down it, and the false (F) path has a true value passed down it.

8.4 Date

	The <i>Date</i> too selection from	l allows for d n a calendar	late menu.
	Group	Input	Output
Figure 8-5 - Date	Interface	None	None

The *Black Question Anchor* connects the results to other tools, which have incoming anchors.

The *Date* tool creates a string of the selected date in the format YYYY-MM-DD.

Properties Window:

The Date Configuration window has only one component.



Figure 8-6 – Date Configuration

Enter the text or question to be displayed allows us to write the prompt that the end user sees.

8.5 Date Filter

۶	The <i>Date Filter</i> tool provides the ability to filter based on dates.			
	Group	Input	Output	
Figure 8-7 – Date Filter	Preparation	See below	See below	
An Action tool can be co	onnected to the	Lightning Bo	lt Anchor to	
nodify how this tool works in apps and macros.				
<i>Input:</i> A data stream with at least one date field.				
<i>Output:</i> The input data st range.	tream limited t	o appropriat	e date	

Properties Window:

The Date Filter Configuration window has two components.



Figure 8-8 - Date Filter Configuration

- The *Select the date field to be used in filtering* allows in selecting the date field in the incoming data source.
- *Date selection method* allows us to define which method the date filter will use to limit the data.
- *Select the start and end dates* allows us to select two dates on a calendar and only keep the dates restricted between them.

ubseque eceding ta to inc	nt periods a periods befo	fter that da
ubseque eceding ta to inc	nt periods a periods befo	fter that da
ta to inc	clude	
u Fri	Sat	
5 27	28	
5 6	7	
2 13	14	
9 20	21	
6 27	28	
2 3	4	
23/2015		
e the en	d date to inc	clude in the
	u Fri 27 62 13 20 527 23 23 23 23/2015 e the er	u Fri Sat ; 27 28 ; 6 7 2 13 14) 20 21 ; 27 28 2 3 4 2/2015 e the end date to in

Figure 8-9 - Date Filter Configuration - Using Start and End Dates

- Use start date and include subsequent periods after that date allows us to select a date to start from and then define how long after that date we want to include.
- Use end date and include preceding periods before that date allows us to select a date to end the data and then define how long before that date we want to include.
8.6 Drop Down



The <i>Drop Down</i> tool allows for the
selection of a single value from a
provided list.

Group	Input	Output
Interface	None	Data stream

The *White Question Anchor* receives information from other tools to populate the list.

The *Black Question Anchor* connects the results to other tools that have incoming anchors.

The result of this question is the string of the selected option.

Properties Window:

The *Drop Down Configuration* window has three basic components.



Figure 8-11 - Drop Down Configuration

- *Enter the text or question to be displayed* allows us to write the prompt that the end user sees.
- *List Values* allows us to select a method of populating a list and any related properties.
- *Default Selection Name Partial OK* allows us to set a default value by typing in a unique part of the option.
- *Field Types* are the list of field types.
 - *External Source Must contain NAME & VALUE fields (can be relative path)* uses a reference file that has both a Name and a Value field.

- *Field from Connected Tool* allows us to connect to a tool and select one of the fields from that data connection, as long as they match the file type in the multi-select list.
- Manually set Values (Name: Value one per line) allows us to create a list of options in the properties box. The format for this is: *Text to display: Resulting value in data*.

Allocate Datasets	
Properties	File Browse in this Analytic App/Macro Properties
Geocoder Datasets Properties	Show only Fields of the following types: Ø Bool Ø Byte None
Drivetime Datasets Properties	
Behavior Analysis Datasets Properties	V Float V Double V String V WString
Reference Base Maps Properties	✓ V_String ✓

Figure 8-12 - Dataset types

- *Allocate Datasets* allows us to select from the Allocate dataset list we have available.
- *Geocoder Datasets* allows us to select from the Geocoder dataset list we have available.
- *Drivetime Datasets* allows us to select from the Drivetime dataset list we have available.
- *Behavior Analysis Datasets* allows us to select from the behavior dataset list we have available.
- *Reference Base Maps* allows us to select from the base maps we have available.
- *File Browse in this Analytic App/Macro* connects to a file browse tool and allows for selection of one of the fields from that data connection, as long as they match the file

type in multi-select list. Include [None] means we are providing an option that will result in no selection.

8.7 File Browse

	The <i>File Browse</i> connect to a file a predesignated	tool allo of choic l file.	ws to e instead of
	Group	Input	Output
G Figure 8-13 – File Browse	Interface	None	None
The Black Question Ancho	<i>r</i> connects the real	sults of q	uestions to

anchors on other tools.

The tool can be used to update any tool that used a file connection. When using this tool, the end user should know what information should be included in the data so the module runs correctly.

Properties Window:

The File Browse Configuration window has three components.



Figure 8-14 - File Browse Configuration

- *Enter the text or question to be displayed* allows us to write the prompt the end user sees.
- *Save As Dialog* allows us to output the data as a dialog box instead of a file.
- *Standard Database File Formats* are a typical file browse.
 - *Spatial Formats Only* will allow only file types that can have spatial objects.
 - *Generate Field Map for Input Data* forces the end user to align the fields in the selected data source with the fields in the development data source.
- *Report Formats* option allows us to select only report formats that are compatible with the *Render* tool.
- *Arbitrary File Specification* allows us to define specific file formats that we want to be accessed or preferred.

8.8 Radio Button

	The <i>Radio But</i> single select c <i>Interface desig</i>	<i>tton</i> tool creates a option for use in the <i>mer</i> .			
	Group	Input	Output		
Q Figure 8-15 - Radio Button	Interface	None	None		
The Black Question Ancho anchors on other tools. The question creates a Bo along with all other singl no groups are defined).	<i>r</i> connects the polean. It will be select option	results of qu e a single se s in the grou	estions to elect option 1p (all, if		

Properties Window:

The *Radio Button Configuration* window has three components.



Figure 8-16 - Radio Button Configuration

- *Enter the text or question to be displayed* allows us to write the prompt the end user sees.
- *Default Value*, when selected, will use this question as the default selected.
 Best practice is to make sure this question is first in the list of values.
- *Collapse Group When Deselected* allows us to create a group of questions that only display when the question is selected.

8.9 Text Box



Figure 8-17 - Text Box

The *Text Box* tool allows the user to enter a string of choice.

Group	Input	Output
Interface	None	None

The *Black Question Anchor* connects the results of questions to anchors on other tools.

The questions produce a string output.

Properties Window:

The Text Box Configuration window has five components.



Figure 8-18 - Text Box Configuration

- *Enter the text or question to be displayed* allows us to write the prompt the end user sees.
- *Default Text* is a string that will populate the text box.
- *Mark Text (for Passwords),* when checked, replaces all visible text characters in the response with "*".
- *Multiline*, when checked, allows the end user to create a multiple line input and makes the text entry larger.
- *Hide control (for API development)* gives the ability to hide the entry field from the interface designer but allows the API to continue to edit the entry.

8.10 Let's Build an App For That!

То	Alteryx Consultants
Subject	Let's Build An App For That!
Attached	All Recorded Traffic Tickets.csv
Hey,	
I need to p	ull you in on a project with the Baltimore Ticket Team.
They have	asked us if we can build an application that will allow them to do the following:
1. selec	t a file with violations
2. selec	t a date range to limit the data
3. have	e it create output file(s) of their choice
а	. in-state plates
b	. out-of-state plates
C	in-state and out-of-state plates (as a single or multiple files).
I think we w this analyti to select ar	will hear back from them soon if we build it to these exact standards. We should design c application such that the default is to select Maryland, yet it also gives the flexibility ny state. Also include a text box so they can limit to a specific location of interest.
Should be t	fun.

Alteryx allows us to develop Analytic Applications that make it so the workflow can be modified by user inputs. Apps have a couple of special properties that allow us to:

- Use an interface to run the workflow instead of opening the file in Alteryx Designer.
- Publish the workflow to an Alteryx Gallery, giving users without Alteryx Designer the ability to run workflows.

			=		Fie	ble	Туре		Size	Rename	Desc	cription
_			- 1	•	🗸 tag		V_String					
					✓ stat	e	V_String	•	254			
4		- (0		V loca	ation	V_String	•	254			
					Viol	Code	V_String	٠	254			
Kec	orded		ā		✔ De:	scription	V_String		254			
піс	ickets.	SV			Viol	Fine	V_String	-	254			
					Viol	Date	V_String	•	254		1	
					✔ *Ur	nknown	Unknown	•	0		Dyna	mic or Unkn
ults - B	rowse (2)											÷ 1
7 of 7 1	ields 🔹 🖌	Cell Viewer	- 1 1 1,018,97	8 records	displayed,	31 MB						Ø- 8- 1
Record	# tag	state	locat	tion		violCode	2		Description		violFine	violDat
1	KZW022	MD	PARK HGTS @ VIOLE	TAVE.		30	Red Light Vi	olatio	n	\$	75.00	11/08/2002 12:2
2	LCK831	MD	EDMONDSON @ ATH	IOL / WO	ODRIDGE	30	Red Light Vie	olatio	n	s	75,00	01/15/2004 07:5
3	LSN017	MD	RUSSELL (NB) @ HA	MBURG	ST.	30	Red Light Vi	olatio	n	5	75.00	02/29/2004 02:1
4	MGC390	MD	NORTHERN PARKWA	Y @ FAL	LS RD.	30	Red Light Vie	olatic	n	\$	75.00	03/04/2004 07:0
5	MLG502	MD	FRANKLIN ST. @ CAT	THEDRAI	STREE	30	Red Light Vi	olatio	n	5	75.00	03/31/2004 12:1
6	155M003	MD	REISTERSTOWN ROA	D @ PA1	TERSON	30	Red Light Vir	olatio	n	5	75.00	04/10/2004 05:2
		1.00			D. 81.45	20	Ded Gehtin.	alatie			75.00	05/20/2004 02/2
7	155M003	MD	PARK HGTS. AVE. B H	1AYWAR	DAVE.	20	Red Light Vi	oracit	Alt	3	12:00	03/30/2004 02:2.

Figure 8-19 – Traffic Tickets – Opening, selecting and viewing data

We are going to approach this by laying out the workflow and then adding the interface tools that allow for the modification. So, let us get started by connecting to the file *All Recorded Traffic Tickets.csv.* Since they haven't asked us to do anything with the Fine amount field, we can leave it in a string format.

However, since we know we need to limit the data by dates, we are going to want the *violDate* in a date format. If we take a look at the format in *violDate*, we see it is a string-formatted date with an *AM/PM* flag. If we look at the *DateTime*, tool we can see that we do not have a matching date format, so we need to use a formula tool to manipulate the string first.



Figure 8-20 - Traffic Tickets - Conversion from string to date



Figure 8-21 - Traffic Tickets - Input String Format

Since we are only going to allow the end user to select dates, we can drop the time portion of the field. This means that we only need the first 10 characters. The formula is: *LEFT([violDate],10)*



Figure 8-22 - Traffic Tickets - Using Date filter tool

Now that we have a field in date format, we can create the *Date Filter* tool. Because we only have one date type field, *Filter Date* populates in the field selection. We are going to allow the end user the ability to select the beginning and end date, so we can select that option and leave all other settings alone. It does not matter that both dates are referring to today because we are going to replace these values.

Note: While we are testing, it may make it hard to see what is going on if we do not select a date range that is reflected in the data.



Figure 8-23 - Traffic Tickets - Using RegEx to flag string

Now we can add the feature they didn't ask for. We can write a regular expression that allows us to capture partial or full location names to limit the data. We want to look at the location field and then create a regular expression that will tell us if the field matches the user input. We can use the following expression along with a *text box interface* tool to do this.

.*something.*

Using the above expression, we will update *something* with the user-specified value so we can flag the appropriate records. (If the user wants everything, they can leave the field blank resulting in the *expression* .*.*, which will mark all fields as true.)



Figure 8-24 - Traffic Tickets - match with flagged string

Now that we have flagged which records match the desired location, we can use the *location_Matched* field that was generated in the *RegEx* tool to filter the data. If we use the basic filter for the *location_Matched*, we can see the field is true and the function becomes the expression *location_Matched* = 1.

Now that we have limited the data to the appropriate locations, we can flag which state they may want to analyze in detail. For now, we can use a custom filter expression in order to finish building the core data stream.

[state] = "Replace with State"



Figure 8-25 - Traffic Tickets - getting three possible outputs

The last step in the workflow is to create the three possible outputs. We can add them to the end of the data stream where appropriate.

- Connect the *Selected State Violations.csv* to the True output on the state *filter*.
- Connect the *Non-Selected State Violations.csv* to the False output on the state *filter*.
- Connect the *All Matching Violations.csv* to the True output on the Location Match *filter*.

We can put the three outputs in *Tool Containers*. This is not for organizational purposes like we saw earlier. When we are building applications, we can enable and disable *Tool Containers* based on the user selection. In this case, we are going to disable all three tool containers so the only time the output is created is when we have selected that option.

Caption		ID
Selected State		14
		Reset to Defa
Text Color	DarkSlateGray	
Fill Color	R=236, G=242, B=242	
Border Color	DarkSlateGray	
Transparency	25	
Margin	Medium	

Figure 8-26 - Traffic Tickets – using Tool Container to disable it except when appropriate option is selected

We can disable the containers using the *Disabled* option in the bottom-left corner of the properties windows. This option only shows the container and not the contents and will turn off all the three outputs.



Figure 8-27 - Traffic Tickets - selected options



Figure 8-28 - Traffic Tickets - Basic data stream

We now have the basic data stream designed for the analytic app. We will have to modify some of the tools, but almost everything is ready for the transition. Let us look at the list of interactions we are providing the Baltimore Ticket Team.

- 1. Select files.
- 2. Choose the date range.
- 3. Name a location (optional).
- 4. Select a State (Maryland, by default).
- 5. Select which combination of the three output files is to be returned.



Figure 8-29 - Traffic Tickets - File Browse

The first question we can address is which file should be selected. If we drop a *File Browse* tool onto the canvas, we can see that the tools show their *lightning bolt* and *question mark* anchors indicating we are now working with an *Analytic Application or Macro*. We can then change the text so it reads *Select a file* and use the remaining default settings.

We know we want to connect the file selection question to the *Input Data* tool; however, the *lightning bolt anchor* only connects to *Action* tools. We could bring in the tool, but Alteryx has a clever way of handling it: When we connect a question directly to a *lightning bolt anchor*, it brings an *Action* tool onto the canvas with all of the default settings in place. In this case, it will do exactly what we want, and we can move onto the next question. (Click the Action tool On and Off to confirm the settings.)



Figure 8-30 - Traffic Tickets - using Action Tool for Date

The next question is slightly more complicated. We need to set the beginning and end dates for the selected timeframe. We can drag in two *Date* tools and label one *From* and the other *To* so we know which dates refer to the beginning and end of the range. This time, when we connect the questions to the *Date Filter* tool, there is no indication what value should be updated. What we want to do is update two values in the *Date Filter*. However, since there are multiple inputs, we need to select which value each action should update.

- From's *Action* tool should update the variable *start*-*end.start*
- To's Action tool should update the variable start-end.end

Now that we have this filter getting its values updated, we can move onto the next question.



Figure 8-31 - Traffic Tickets - using Text Box to specify location

This time, we are going to add a *Text Box* tool with the prompt *Type a location here (optional)*. This will allow the end user to type anything they want to limit the locations required. We want this tool to modify the regular expression in the *RegEx* tool.



Figure 8-32 - Traffic Tickets - using Action with RegEx to specify string

We can see the *Action* tool is giving us an error, and if we take a look, we need to specify what should be updated. In this case, we know that in the *RegEx* tool, there is the expression string *.*something.** We want to replace *something* with the value the end user types in. Thus, we can use the *Replace a specific string* option to replace the string *something* within the regular expression. Now let's see how we can modify the state filter.



Figure 8-33 – Traffic Tickets – using Drop Down for selecting an individual State of Interest

This time, we are going to feed the answer of a question directly into a tool for use in the expression. We are going to create a *Drop Down* question with the prompt *Select an Individual State of Interest*. We are going to use the second file we received from the Baltimore Ticket Team to import the list of possible state codes from a file called *All Violation State Options.csv*. Finally, we are going to set the default value to *MD* so the team can quickly run the in-state analysis.

Custom I	Filter			
Variables	Functions	Saved Expressions		
		4	Click to see field	
Expression:			Ivanies	

Figure 8-34 - Traffic Tickets - using Formula to filter the data

We are almost done with this question, except that we need to update the formula used to filter the data. We can replace the original formula with the one in the expression mentioned. (If we do not rename the connection, it should say [#1] instead of [Selected State].)

There is only one more question, but this one is the most complicated. This last question involves selecting which output(s) get created. Since the end user may want to output any combination of three different files (we will force them to select at least one), there are five different options we need to provide.



Figure 8-35 - Traffic Tickets - choosing from 5 different option

Create five Radio Button tools with the default settings other than the labels:

- 1. Selected State File
- 2. Non-Selected State File
- 3. All Violations File
- 4. Selected State and Non-Selected Files
- 5. All three Files

The goal here is to develop a structure where anytime we want a particular file created, we enable the *Tool Container(s)*. In order to activate the tool containers, we need *Action* tools connected to them that enable the container if the incoming connection is true.

Since multiple values can result in each file getting produced, we have complex conditions that need to be met. In

order to evaluate these conditions, we need to add a *Condition* tool before each *Action* tool.

Each of these Condition tools should be connected in the following way:

- 1. Condition connected to the Selected State Output
 - Selected State File (connection #1)
 - Selected and Non-Selected State Files (connection #2)
 - All three files (connection #3)
- 2. Condition connected to the Non-Selected State Output
 - Selected State File (connection #1)
 - Selected and Non-Selected State Files (connection #2)
 - All three files (connection #3)
- 3. Condition connected to the All Violations Output
 - All Violations File (connection #1)
 - All three Files (connection #2)

The expressions in each of the three conditions should be:

- 1. [#1] OR [#2] OR [#3]
- 2. [#1] OR [#2] OR [#3]
- 3. [#1] OR [#2]



Figure 8-36 - Traffic Tickets - Interface designer

Now that we have the workflow developed, we can see the *Interface Design*. The interface is how the end users are going to interact with the workflow. As we can see from the previous page, the *Interface Designer* shows us the list of questions we created in the workflow. We could leave the questions like this, yet it provides a nicer user experience if we organize the tools.



Figure 8-37 - Traffic Tickets –Choosing Tab from the drop down

We can create a new tab by clicking *add drop down* and selecting *Tab*. We can then click on questions we want to move in order to have related questions together.



Figure 8-38 - Traffic Tickets – Entering text to be displayed

We can rename this tab by clicking on the tab at the top of the list and changing the text in the *Tab configuration window*.



Figure 8-39 - Traffic Tickets – Location information

We can then add a new tab and rename it *Location Information*. Once it is created, we can move the questions in the previous image to that tab that is about Location (of the driver's home state or violation).

Inte	erface Designer - Lay	out View	
Ħ	File and Date Range	Location Information	Output
+2:	Select a file		ar an

Figure 8-40 - Traffic Tickets - Adding a tab called Output

Now we can create one more tab called *Output*. But before we move questions, we need to add a *Group Box* from the *Add* menu.



Figure 8-41 - Traffic Tickets – Moving all questions into a group using LabelGroup

Wha	at files do you want output?
0 5	elected State File
1 0	Ion-Selected State File
0 5	ingle All Violations File
0 5	elected and Non-Selected State Files
© A	ul three Files

Figure 8-42 - Traffic Tickets - Grouped questions moved to Output tab

Now we will move all the required questions down into the *Group Box* by selecting the question and clicking down until our list looks like the image.

Once we have all of our output questions grouped together, it is easy to move them across into the Output tab. Then we can click on the *Group Box* title and move that.

After completion, we should have the three tabs for our questions.

elect	a file						Select an Individual State of Interest
from 4 Sun 22	March, 2015 Mon Tue Wed Thu Fri Sat 23 24 25 26 27 28				Fri 27	, Sat 28	Type a location here (optional)
1 8 15 22 29	2 9 16 23 30	3 10 17 24 31	4 11 18 25 1 Today:	5 12 19 26 2 3/23/	6 13 20 27 3 2015	7 14 21 28 4	File and Date Range Location Information Output What files do you want output?
0							Selected State File
4		Ma	rch, 21	015		•	Non-Selected State File
22 1	23 2	24 3	25 4	Thu 26 5	Fri 27 6	Sat 28 7	Single All Violations File
8	9	10 17	11 18	12 19	13 20	14	Selected and Non-Selected State Files
22 29	23 30	24 31	25 1	26 2	27 3	28 4	All three Files

Figure 8-43 - Traffic Tickets – all three tabs

For better representation purposes, some of the connections are made wireless. The final workflow after adding the Interface tools looks as shown in the image.



Figure 8-44 - Traffic Tickets - data stream when complete

8.11 To Summarize or Not to Summarize: That is the Question



Chapter 9 Where's the Joe?

То	Alteryx Consultants								
Subject	Where's the Joe?								
Attached	Company Locations.csv								
Hey,									
campaign.									
They have provided us a map of all of their global locations to work with.									
They want to depending o	They want to see the location as well as a set radius around all of their stores, which should be configurable from 1 mile to 10 miles, depending on the user's preferences.								
They want a report to be created with a map of all U.S. locations drawn within a selected area, as well as the individual maps for each state for stores in that selected area.									
This is a good time to introduce you to a concept that will save you a lot of time down the road. Many requests have similar core elements, and it saves a lot of time when you can reuse them. Thus, Alteryx allows you to make macros.									
We should d	nould develop a macro that will allow you to pass map inputs as well as a numeric field for the radius before we tackle this project.								
Thanks.									

9.1 Tools & Concepts

In this chapter, we will be covering the following Tools and Concepts:

Tools	Concepts						
Macro Input	Building Macros						
Macro Output	Using Macros						
Мар	Mapping						
Map Input							
Numeric Up Down							
Spatial Match							
Trade Area							


Figure 9-1 - Macro Input

The Mac	ro Input	tool allows for
creation	of an	incoming data
connectio	n for a mac	cro.
Group	Input	Output
Interface	See below	Any data

Note: The *Macro Input* tool receives data from the calling workflow.

The bottom *Black Question Anchor* can be connected to a question anchor on top of a tool to pass the result to the subsequent tool.

Input: The sample data connection or the calling module's data stream.

The *Macro Input Configuration* window has four basic components.

Macro Input (2) - Configuration	▼ 무 >
Template Input (For Test as a Standard Workflow)	
Text Input	
Edit Data [1 Row & 1 Column]	
File Input	
Configure	
Input Name:	
Input2	
Anchor Abbreviation:	
Show Field Map	

Figure 9-2 - Macro Input Configuration

- *Template Input (For Test as a Standard Workflow)* is the data we will use to build the macro's data stream. The fields here need to match the incoming connections. We can use either a text input or a file input for the source.
- *Input name* is the name of the data connection.
- *Anchor Abbreviation* allows us to type a single character to be displayed on the macro's input connection. Typically, this is used to identify which data stream should be connected to which connector, if we have multiple inputs.

• *Show Field Map* allows us to manually match the fields to be used in the data stream if the names do not match the original data source.

9.3 Macro Output



The *Macro Output* allows us to create an outgoing data connector for a macro.

" ~ `	Group	Input	Output
Figure 9-3 – Macro Output	Interface	Any data stream	See below

Note: The *Macro Output* tool returns data back to the calling workflow.

Output: The incoming data stream into the workflow that called this macro.

The *Macro Output Configuration* window has two components as shown in the following figure.

\$3 N	Aacro Output (3) - Configuration	- ₽ X
2	Output Name:	
3	Output3	
P	Anchor Abbreviation:	
0		
EĮ.		
0		
0		

Figure 9-4 - Macro Output Configuration

- *Output Name* allows us to rename the output data connector.
- *Anchor Abbreviation* allows us to type a single character to be displayed on the macro's output connection. Typically this is used to identify which data will be coming out of the connector if we have multiple outputs.

9.4 Map



The *Map* tool allows us to create spatial object(s) on a map for use in the data stream.

	Group	Input	Output
0	In/Out	None	None
Figure 9-5 - Map			

Note: The *Map* tool creates spatial objects to be used with a Map Input tool.

Application questions can be connected to the top White Question Anchor to update other questions.

The bottom Black Question Anchor can be connected to a question anchor on top of a tool to pass the result to the subsequent tool.

An Action tool can be connected to the Lightning Bolt Anchor to change the spatial objects created.

The Map Configuration window has six basic components.

Base N	/ap
Clou	dMade - Most Recent Vintage
Sele	ct/Draw
0	1 realure
0	1 or More Features
Zoo	m To
۲	Default
0	Custom: Edit Bounds
0	Reference File
Refe	erence Layer
۲	None
0	File Browse Question (must be connected)
0	Reference File
	File Path
	Label Field
Draw	
Drav	w Mode Properties
Use	r Can Draw: 📝 Points 🔲 Lines 📄 Polygons
	User Can Label Features

Figure 9-6 - Map Configuration

Enter the text or question to be displayed allows us to write the prompt the end user sees.

- *Base Map* allows us to select the map the end user will be able to use in order to understand where they are clicking.
- *Select/Draw* allows us to decide if the end user is allowed to create single or multiple spatial objects.
- *Zoom To* allows us to decide what area of the map should be shown by default.
- *Reference Layer* allows us to use a spatial file in order to feed spatial information into the map. We would want to do this so end users can select created objects or use those objects for reference points when they create their own. If we choose to add a reference layer, we can feed it in from a *File Browse* question or set a specific file to always use. When we are using either of these modes a Label Field must be selected (if we use the *File Browse* option setting the label field is done in the interface).
- *Mode* allows one of two input methods. *Draw* allows the end user to create objects on their own, while *Select* allows them to choose objects in the reference layer. When we are in *draw* mode, we have the ability to allow the end user to create points, lines or polygons and decide if the user can create custom labels for the objects.

9.5 Map Input

F	The <i>Map Input</i> allows us to use map as a source to create spatial objects.			
	Group	Input	Output	
	In/Out	None	See below	
Figure 9-7 – Map Input				
<i>Note:</i> This tool creates s	patial objec	ts.		
Those. This tool creates s	patial objec	.15.		

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: A data stream with spatial objects and a label field.

The Map Input Configuration window has four basic components.



Figure 9-8 – Map Input Configuration

• *Mode* allows us to select between *Draw*, where we create the objects, and *Select*, where we can use objects from a file.

- *Base Map* is the background map we are using to give reference.
- *Reference yxdb* allows us to select an Alteryx database file with spatial objects defined so we can reuse previously created map objects.
- *Label Field* allows us to define which field in the Alteryx database file has the spatial object label names.

The map allows us to create spatial objects. The icons labeled 1 – 6 are:

- 1. create point objects
- 2. create line objects
- 3. create polygon objects
- 4. clear all created objects
- 5. show the base map or objects
- 6. zoom controls

9.6 Numeric Up Down

12.3\$	The <i>Numeric Up Down</i> question tool allows us to input a numeric value to be used in the data stream.			
Q Figure 9-9 – Numeric Un Down	Group	Input	Output	
Figure 9-9 - Numeric Op Down	Interface	None	None	
The bottom <i>Black Question</i> question anchor on top of as subsequent tool.	<i>1 Anchor</i> car nother tool to	n be conne pass the re	ected to a esult to the	

The *Numeric Up Down Configuration* window has six components.



Figure 9-10 - Numeric Up Down Configuration

- *Enter the text or question to be displayed* allows us to write the prompt the end user sees.
- *Minimum* allows us to set a lower threshold for the input.
- *Maximum* allows us to set an upper threshold for the input.
- *Increment* allows us to define the most granular level of change.
- *Default* is the value that will auto-populate.
- *Decimal Places* is the number of decimal places that are displayed.

9.7 Spatial Match



Figure 9-11 – Spatial Match

The *Spatial Match* allows us to use one spatial field to filter the data from another.

Group	Input	Output
Interface	Input T & Input F below	Output M & Output U below

Note: One of the two inputs needs to use a polygon. When we have multiple target spatial objects that match multiple universe spatial objects, the *Spatial Match* tool will replicate data in the output.

Input T: Any data stream with a spatial object field.

Input U: Any data stream with a spatial object field (optional depending on settings).

Output M: The target (T) input data stream limited by the records that matched in universe (U) input data stream.

Output U: The target (T) input data stream limited by the records that did not match (unmatched records) in universe (U) input data stream.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

The Spatial Match Configuration window has five components.

	rgets	(T Input)					
	Spatia	al Object F	ield: Centroi	d			
Ur	ivers	e					
(Us	e Records	from U Input	t			
(Us	e Records	from File or	Database:			
	Sp	atial Object	ct Field: Spa	tialObj			
					_		
Wh	ere I	arget Io	uches or inte	rsects	•	Universe.	
	Dutpu	it Intersecti	ion Object (Ir	ntersects On	y)		
_		1.20				2 AG 2	
Jp	tions	•	t	TIP: To	reo	rder multiple r	ows: sel
		Input	Field	Туре		Size	Renam
	122	Tamet	Latitude	Float	•	4	
		Torgos	1				
•		Target	Longitude	Float	•	4	
Þ	V V	Target Target	Longitude Centroid	Float SpatialObj	•	4 536870911	
Þ		Target Target Universe	Longitude Centroid Label	Float SpatialObj V_WString	•	4 536870911 1073741823	
		Target Target Universe Universe	Longitude Centroid Label SpatialObj	Float SpatialObj V_WString SpatialObj	* * *	4 536870911 1073741823 536870911	
		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown	Float SpatialObj V_WString SpatialObj Unknown	* * * *	4 536870911 1073741823 536870911 0	
•		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown	Float SpatialObj V_WString SpatialObj Unknown	* * *	4 536870911 1073741823 536870911 0	
•		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown	Float SpatialObj V_WString SpatialObj Unknown	• • •	4 536870911 1073741823 536870911 0	
•		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown rr	Float SpatialObj V_WString SpatialObj Unknown	* * *	4 536870911 1073741823 536870911 0 ✓ Univer	rse.
•		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown II touches Intersect	Float SpatialObj V_WString SpatialObj Unknown	• • •	4 536870911 1073741823 536870911 0 Univer	rse.
*		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown "" tunknown "" tunknown	Float SpatialObj V_WString SpatialObj Unknown	•	4 536870911 1073741823 536870911 0	rse.
~		Target Target Universe Universe	Longitude Centroid Label SpatialObj *Unknown II turkno III turkno III turkno I	Float SpatialObj V_WString SpatialObj Unknown	•	4 536870911 1073741823 536870911 0 Univer	rse.
		Target Target Universe Universe Universe	Longitude Centroid Label SpatialObj *Unknown I Unknown I Unknown I Unknown I Unknown I Unknown I Unknown I Unknown I I Unknown I I I I I I I I I I I I I I I I I I I	Float SpatialObj V_WString SpatialObj Unknown r or Intersects Rectangle C	• • •	4 536870911 1073741823 536870911 0 Univer	rse.

Figure 9-12 – Spatial Match Configuration

• *Targets (T Input)* allows us to select a spatial object field from the target input.

- *Universe* allows us to define if we are going to use a field in the universe input or connect to a spatial data file and then select the field.
- Where Target _____ Universe allows us to define the relationship that defines a match between the target and universe spatial fields, while the check box will include the intersection as an additional spatial object field.

The types of interaction are:

- *Intersects* means they share the same space in at least a partial manner.
- *Contains* means the entire universe spatial object is in the target spatial object.
- *Within* means the entire target spatial object is in the universe spatial object.
- *Touches* means the two spatial objects share a border but do not intersect.
- *Touches or Intersects* means the two spatial objects share a boarder or intersect.
- *Bounding Rectangle Overlaps* means the fields are matched if the smallest rectangle to contain the entire target spatial object intersects the smallest rectangle to contain the entire universe spatial object.
- *Custom DE-91M Relation* allows us to define customized relationships between the objects.
- The bottom of the tool has an embedded *Select* tool.

E .	The <i>Tr</i> polygon that is p	<i>ade Area</i> tool centered on a p rovided.	draws a point object
	Group	Input	Output
Figure 9-13 - Trade Area	Spatial	Any data stream with a point spatial field	None

Note: In order to use the drive radius data functionality we need to have the *TomTom* data package.

Lightning Bolt Anchor: An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros

The *Trade Area Configuration* window has four basic components.

Centre	oid 👻 🔽 Include in Output	
Rad	us, Doughnuts or Drivetime	
۲	Specific Value:	
	5.0	
0	From Field:	
	Store ID	Ĩ
Unit	3	
۲	Radius (Miles)	
0	Radius (Kilometers)	
0	Drivetime Minutes	
	Dataset: TomTom US - Most Recent Vintage	
E	minate Overlag (available for Specific Value Radius only)	

Figure 9-14 - Trade Area Configuration

- *Spatial Object Field of Point Source* shows us the list of spatial object fields in the incoming data and allows us to select the one we are interested in. Checking the *Include in Output* box maintains the point field in the output.
- *Radius, Doughnuts or Drivetime* allows us to set a numeric value for the distance between the point and the edge of the polygon.
- *Units* allows us to define what the unit of measurement Specific Value represents, and if it is Drivetime Minutes, which dataset to use to create the distance.

• *Eliminate Overlap* (available for Specific Value Radius only) allows us to create non-overlapping polygons.

9.9 Where's the Joe?



Macros are tools we develop so processes we need to reuse in the same workflow or others don't need to be replicated. This is useful for multiple reasons: The first is it will save us time in creating workflows, and the second is it makes maintenance much easier by allowing us to make a change to a single macro instead of each instance in all files that use the macro.

Let's start by building a macro that is given a data stream with latitude and longitude field. *Macros* are a special type of Analytic Application, so we will be taking the same approach in designing it as we do with apps.



Figure 9-15 - Mermaid Coffee Macro Input for Location

We are going to build this macro by starting with a *Macro Input* tool. We are going to create a sample data using the *Text Input* option and make sure that *Show Field Map* is checked. We are setting up the data this way because we need to make sure we have latitude and longitude fields in any data that uses the tool. In the event the incoming data doesn't have fields named *Latitude* and *Longitude* we are giving the person using this macro the ability to map the fields accordingly.

elect	: (3)	- Configu	uration				- ₽ X	Coordinates to Selected Trade Areas.yxmc* ×
Op	tions	+ 1 1	Ļ		TIP: To	reorder multipl	e rows: select, right-	
		Field	Туре		Size	Rename	Description	
•		Latitude	Float	•	4			Input
	V	Longitude	Float	•	4			
	7	*Unknown	Unknown	•	0		Dynamic or U	0
	Op ^r	options	elect (3) - Configu Options - 1 Field Fiel	elect (3) - Configuration Options	elect (3) - Configuration Options	elect (3) - Configuration Options If I TIP: To Field Type Size Field Roat If I Congitude Roat If I If I Congitude Roat If I If I Congitude Roat If I If I Congitude Roat If I Congitude	elect (3) - Configuration Options • TIP: To reorder multipl Field Type Size Rename V Letitude Float • 4 Congitude Float • 4 Longitude Float • 4 Vunknown Unknown • 0	elect (3) - Configuration

Figure 9-16 - Mermaid Coffee using Select for getting data

Now for the first time we pay attention to the **Unknown* option in the select tool. Because we are building a macro, we don't know what data will be coming into this tool; this option allows us to pass through any unknown fields in the data stream without modifying them.



Figure 9-17 - Mermaid Coffee Creating points using default settings

Since we have named the latitude and longitude fields in a way that the *Create Points* tool recognizes, it will auto populate with those fields in the appropriate places in the settings. We are going to use the default settings to create our points.



Figure 9-18 - Mermaid Coffee using Map Input and Spatial Match to filter incoming data



Figure 9-19 - Mermaid Coffee incoming data stream

In order to limit the incoming data, we are going to use a *Map Input* tool to create a spatial object that will be passed to the *Spatial Match* tool. The *Spatial Match* tool has the ability to compare different spatial objects to determine if they share space.

We are not going to configure the *Map Input* tool because we are going to be giving it spatial objects from an interface tool called *Map* once we have finished the basic workflow. The *Spatial Match* tool will use the configuration shown in the image to include only locations that are in/or touch the user defined Map Input region.



Figure 9-20 - Mermaid Coffee with points and limits

Now that we have limited the points to only those in the area of interest, we can create the *Trade Area* polygons for those points. We are going to make sure that the units are miles, but it doesn't matter that the *Specific Value* is 5 because we are going to use a *Numeric Up Down* to set the distance.



Figure 9-21 - Mermaid Coffee Trade Area Polygons

Now that we have created our trade areas, we need to create outputs.



Figure 9-22 - Mermaid Coffee - Using Report Map to create output

For the map output we need to create a map report object. Use any setting of choice; just make sure to include a base layer and the polygon layer from the data stream.



Figure 9-23 - Mermaid Coffee Macro Output

Now all that is left for the outputs is to name them and provide them with an identifying anchor character.



Figure 9-24 - Mermaid Coffee data stream after Macro Output

This macro now only needs the user settings so that the controls can be defined for the two tools we want to be able to change.



Figure 9-25 - Mermaid Coffee using Map Tool

	Specify display text
ase Map	
CloudMade - Most Recent Vi	ntage
Select/Draw	
I Feature	+
1 or More Features	Specify Base Map
Zoom To	
Oefault	
Custom: Edit Bounds	5
Reference File	
Reference Layer	
None	
File Browse Question (m	ust be connected)
Reference File	
File Path	
File Path	
Label Fie User ca	n create only polygon
File Path Label File User ca	n create only polygon
File Path Label File User ca Mode Draw	n create only polygon

Figure 9-26 - Mermaid Coffee configuration using Map Tool

We can use the *Map* tool to create a map in the interface, which allows the end user to select a specific region.

If we connect the *Map* tool directly to the *Map Input* tool, we see the *Action* tool will auto-populate with the appropriate settings.



Figure 9-27 - Mermaid Coffee using Numeric Up Down to set Number of miles a trade should cover

If we add a *Numeric Up Down* tool, we will have a numeric input so the user can set the number of miles a *Trade Area* should cover. If we connect this directly to the *Trade Area* tool, we can see the *Action* tool uses the special *Trade Area Radius* setting (so we don't need to configure the *Action* tool). We are giving a much wider set of selected radius here because we don't know how we may want this tool to be used in the future. Since we will be embedding this in another tool we can further limit the options for our current project.



Figure 9-28 - Mermaid Coffee data stream

Now that we have this macro, we can save it to use in other data streams. We are going to create a macro folder in *My Documents* called *Alteryx Macros* and create a sub-folder called *Training Macros*. Let us save this macro there, and we should modify the user settings so our macro is available in the *Tool Palette*.

Navigate to the Edit User Setting window in Tools > User Settings > Edit User Settings.



Figure 9-29 - Editing User Settings of Macro

If we go to the Macros Tab, we see the categories we defined and the folder location related to those categories. We have a *General Macros* category and a *Training Macros* category. If we click on the add symbol to the right of the window, it navigates to the *Training Macros*. Let's add it to this list with the title *Training Macros*.



If we take a look at the *Tool Palette*, we see at the far-right side a Category called *Training Macros*. The image shown here as well as description information can be edited if we modify the properties in the workflow *Configuration window* and *Interface designer*.



Figure 9-31 - Newly saved as Macro

orkflow - Configuration	• ₽×	Coor	dinates to Selected Trade Areas.yxmc* ×
Canvas Workflow Events Me	a Info		
Workflow Name	<u>^</u>		
Use File Name		I	
Custom: Coordinates to Sele	cted Trade Areas		Select an area of
Description			interest
		Inte	erface Designer - Properties 🛛 👋 🕹
		菖	Standard Icon
URL:		12:	Custom Icon
Display Text (optional):			
Tool Settings		1	Help: Use File or Hyperlink (relative paths OK)
Root Tool Name:	=	Å	
Tool Version:			Cutout fields change based on macro's configuration or data input
Taal In Database?			Confer news change based on mocro's configuration of data input
Tool Palette			
Category Name:			O O
Search Tags: Example: calcu	late, formula, math		
Author			
Name:			
Company:			
Copyright:			
	Remember as Default		
		1	
	*		

Figure 9-32 - Training Macro Workflow Configuration

Now that we have created the macro, we can build the app specific to the Mermaid Coffee Company. Open a new workflow and connect to the Mermaid Coffee Company Locations file. We can see that all fields are V_String. But we need Latitude and Longitude in a decimal format, which means we need to convert those fields to double so we will have no issue using our macro.

	-	Field	Туре		Size	Re
		Store ID	V_String	•	254	
	V	Name	V_String	•	254	
>	V	Brand	V_String	•	254	
	V	Store Number	V_String	•	254	
	V	Street Combined	V_String	•	254	
	V	Street 1	V_String	•	254	
	V	Street 2	V_String	•	254	
	V	Street 3	V_String	•	254	
	V	City	V_String	•	254	
	V	Country Subdivision	V_String	•	254	
	V	Country	V_String	•	254	
	V	Postal Code	V_String	•	254	
	V	Latitude	Double	•	8	
	V	Longitude	Double	•	8	
	V	Timezone	V_String	•	254	
		*Unknown	Unkno	•	0	

Figure 9-33 - Mermaid Coffee Select

Res	ults - Brov	vse (6)		- P	×
=	15 of 15 Fi	elds 🔻 🖌	.	+-	>>
Ä	Record #	Latitude	Longitude		*
	1	42.23518753	-83.68167114	East	
	2	42.08749008	-87.70139313	Cent	
	3	39.52075958	-119.7415619	Paci	
	4	39.14199448	-121.6445694	Paci	
	5	38.57354355	-89.93135834	Cent	
	6	30.00429535	-90.18016052	Cent	
	7	35.24547577	-120.6742096	Paci	
	8	39.78686523	-105.1101074	Mou	
	9	39.65342331	-104.8729248	Mou	
	10	41.17834473	-85.11387634	USE	
	11	41.38002777	-73.48099518	East	
	12	42.27028275	-88.97055817	Cent	
	13	41.7522049	-87.91249084	Cent	
	14	41.72386551	-88.37533569	Cent	
	15	39.56288528	-76.97272491	East	
	16	36.64411163	-93.21574402	Cent	
	17	40.7098999	-73.83049774	East	
	18	34.49869919	-84.93169403	East	-
	4		·		

Figure 9-34 - Mermaid Coffee Browse



Figure 9-35 - Mermaid Coffee US Location

We can see from the email that they are only interested in US locations, so we can remove all other countries. This means the data is ready for use in our macro.



Figure 9-36 - Mermaid Coffee Trade Area configuration to restrict to US

We must make sure the Latitude and Longitude fields match correctly but leave the other *Map* and *Trade Area Miles* as defaults for now. We will connect a *Map* tool and *Numeric Up Down* here to allow the end user to update these values.



Figure 9-37 - Mermaid Coffee US locations after setting Trade Area

Preview (Click to enlarge		
neview (click to enlarge		Sample D
Sottings D		Refre
Map Size (W x H):	8 by 6 Inches	
Resolution:	1x (96 dpi)	
Resolution: Scale:	1x (96 dpi) Miles	
Resolution: Scale: Reference Base Map:	1x (96 dpi) Miles CloudMade - Most Recent Vintage	
Resolution: Scale: Reference Base Map: Background Color:	1x (96 dpi) Miles CloudMade - Most Recent Vintage R=253, G=254, B:	
Resolution: Scale: Reference Base Map: Background Color: Map Drop Shadow:	1x (96 dpi) Miles CloudMade - Most Recent Vintage R=253, G=254, B- No	

Figure 9-38 - Mermaid Coffee Report Map Setting

Settings Data Layers Legend		Refree
Incoming Connections Imutation Dat		
- File Connections	a Connection Name:	#1
Spa	tial Field:	[All SpatialObjs]
Gro	uping Field:	Country Subdivis
The	matic Field:	Postal Code
Lab		[None]

Figure 9-39 - Mermaid Coffee Report Map Data

Settings Data Layers	Leg	end
⊕• ⊖ ↑ ↓ ⊡ •		
Points Layer	•	
Lines Layer		
Polygons Layer	•	#1

Figure 9-40 - Mermaid Coffee Report Map Layers

We know our macro created a universal map, so we can use that directly. However, we also need individual maps for each state for which we use another *Report Map* tool with the previously detailed settings to create each of the individual state maps.



Figure 9-41 - Mermaid Coffee data stream with Maps

Now that we have both sets of maps, we can combine them. If we want the overall map to precede the others, we need to adjust the order in the *Union* tool.

We have all of our maps in a single field. We can use the *Render* tool to create our report.



Figure 9-42 - Mermaid Coffee using Union Tool to get overall map

Render Cont	iguration		
Output Mode	e: Choose a	Specific Output File	
Output File:			
Group Da	ata Into Separa	ate Reports	
Field To	Group On:	Group	
Modify F	Filename By:	Appending Group To Filename	
Report Data			
Data Field	Layout		
Separator	Insert Section	n Breaks Between Records	
Header	Layout		
Footer	Layout		
	0 He	ader/Footer edge distance (points) der and footer on first page	
Report Style	-		
Paper Size	Letter (8.5"	x 11")	
Orientation	Portrait ()	Landscape	
Margins (in.	0.5 inch Ma	argins	

Figure 9-43 - Mermaid Coffee creating Report from Maps

Note that we are inserting section breaks between each record to make sure each map is on a page by itself.


>	Enter the text or question to be displayed:	1				
>	Select the Region of Interest					
•	Base Map					
	CloudMade - Most Recent Vintage 🔹					
•	Select/Draw 1 Feature 1 or More Features 	1273- 1	Numeric Un Down	(15) Configuration	_	п
	Zoom To		willenc op bown	(15) - Configuration	•	Τ.
	O Default	*	Enter the text or qu	uestion to be displayed:		
	Custom: Edit Bounds	3	Mile radius arrour	nd each store		
	Reference File	0	Minimum:	1.000		
	Peference I wer	0	Maximum:	20.000 🚔		
	None		Increment:	1.000		
	File Browse Question (must be connected)		Default:	5.000		
	Reference File		Desired Discours			
	File Dath		Decimal Places:	U 😴		
	Label Field	0				
	T					
	Mode					
	Draw					
	Draw Mode Properties User Can Draw: Points Lines V Polygons	9				

Figure 9-45 - Mermaid Coffee using Map and Numeric Up Down to get desired view

Now that we have the workflow developed, we can finish off this application by adding the two desired user inputs. We will define the *Map* input the same way we did in the macro. However, the *Numeric Up Down* will have a different range of possible values and default value so it reflects what the Mermaid Coffee Company would like to see.

The Where's The Joe? data stream should look like this when it's complete.



Figure 9-46 - Mermaid Coffee Final data stream when complete

9.10 Who Wants Coffee?

	To Alteryx Consultants							
	Subject	Who Wants Coffee?						
Hev								
,,								
The N	lermaid (Coffee company got back to me and really like what we gave them.						
They each a	want to a ip code.	dd another output to the app. In addition to the report, it should also return the data with total population joined onto They said if the zip code is invalid, just return the record without the population.						
Since censu	adding tl s data by	he total population is something that I know you will have to do regularly, you might as well make a macro that joins total vzip code onto the data stream that you give it.						
The d witho	ata comi ut a popi	ng into the macro needs a string field called "Zip" with Zip 5s in it. If the zip doesn't exist, it should return the record ulation.						
I'll ch	eck on yo	u in a little bit.						

г

Chapter 10 Meta-Morphosis

То	Alteryx Consultants						
Subject	All The Tools Along the way						
Attached	JSON.txt						
Hey,							
I have be higher-u	en asked to make sure you know how to use everything in a list of tools from the ps.						
There are to go ove	There are a few that you haven't gotten to know yet, so I have designed something for us to go over and then for you to work through.						
Attached, you will find a file called JSON.txt. We will be using it to create an application such that when you check a field, the JSON data will be converted into a relational structure. If you don't check the field, it will output a file identical to the original JSON file.							
Thanks							

10.1 Tools & Concepts

Tools	Concepts
Check Box	Alternative Data Prep
Detour	JSON Files
Detour End	Documentation
Dynamic Rename	
JSON Parse	
Message	

10.2 Check Box

	The <i>Check Box</i> interface tool allows us to create a check box question.					
	Group	Input	Output			
	Interface	None	None			
Figure 10-1 – Check Box						
Special Note: This Check Box returns a Boolean value.						

The bottom *Black Question Anchor* connects to question anchors on top of other tools to pass the results to the subsequent tool.

Properties Window:

The *Check Box Configuration* window has three components.



Figure 10.2 - Check Box Configuration

- *Enter the text or question to be displayed* is a text box that will be the question prompt.
- *Default Value* can either be checked or unchecked to determine what the default result will be.
- *Collapse Group When Deselected* allows us to hide the relevant group in the interface designer.

10.3 Detour



The	Detour	tool	allows	us	to	create
alter	native p	aths c	depende	ent o	n tł	ne user
inpu	t.					

· · · · · · · · · · · · · · · · · · ·			
	Group	Input	Output
>	Developer	Any data	Output L &
4		stream	R sections
			below

Note: Detour is useful if we are developing flexible tools that allow choices in data inputs or outputs. This tool should almost always be used with an action tool. Only the Left or Right stream will be active.

An Action tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output L: Nothing or the input data stream.

Output R: Nothing or the input data stream.

Properties Window:

The *Detour Configuration* window has only one component.



Figure 10-4 - Detour Configuration

- The *default* behavior for the *detour* tool is to always divert to the left (L) data stream.
- *Detour to the Right* allows us to reverse this default behavior.

5	The <i>Detour</i> optional pa	<i>r End</i> tool ths.	rejoins two
	Group	Input	Output
	Developer	Input L &	Any data
R		Input R	Stream
Figure 10.5 - Detour End		below	
0			

Note: The *Detour End* tool should be used after we have used a Detour tool in order to bring the data streams back together. Only the Left or Right stream can be active.

Input L: Nothing or any data stream.

Input R: Nothing or any data stream.

An Action tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Properties Window:

The Detour End Configuration window has no configuration.



Figure 10-6 - Detour End Configuration

10.5 Dynamic Rename

0 7	The <i>Dynamic Rename</i> tool allows us to systematically rename fields.					
	Group	Input	Output			
	Developer	Input L &	Any			
R		Input R	data			
		below	stream			
Figure 10.7 – Dynamic Rename						

Note: The *Dynamic Rename* tool is useful in properly identifying column headers in data that either have header rows before the column names or no column names.

Input L: Any data stream.

Input R: This input is optional, and its structure depends on the Rename Mode.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Application questions can be connected to the top *Black Question Anchor* to use answers in the tool.

Properties Window:

The *Dynamic Rename Configuration* window has three core components and eight different modes.

0	Oynamic Rename (15) - Configuration	- ₽ X
∢	Rename Mode: Formula	•
8 T T 1 8	Field_1 Select rename mode Dynamic or Unknown Fields	All Clear
•	Properties Variables Functions Saved Expressions	
	⊕ Current Field ⊕ Constants	
O	Expression:	

Figure 10-8 - Dynamic Rename Configuration

• *Rename mode* allows us to select the method used to rename the fields in the left (L) input.

If we are using one of the single input rename modes, we will see a field selection window.

- The *Properties* section changes, depending on the Rename Mode.
- *Formula Rename Mode* allows us to define an expression that will name our columns.



Figure 10-9a - Dynamic Rename (contd.)

- *Add Prefix/Suffix Rename Mode* allows us to define a string to add to the beginning or end of specific fields.
- *Remove Prefix/Suffix Mode* allows us to remove a common string from the beginning or end of specific fields.
- *Take Field Names from First Row of Data* allows us to replace the column headers with the first record of our data.
- *Take Field Descriptions from Right Input Rows* allows us to use a definition file to redefine the fields.

Oynamic Rename (15) - Config	uration	🝷 🕂 🗙 🙄 Dynamic Renar	me (15) - Configuration	▼ ₽	×
Dynamic Rename (15) - Config Rename Mode: Take Field Names f Wam if number of f Old Field Name from Column New Field Name from Column	In the second se	 Properties Properties Wam S) - Configuration 	we (15) - Configuration Take Reld Descriptions from Right Input Metadata if number of Field Names do not match	• 7	× ×
	Rename Mode: Take	Pield Names from Right Input Metadata			

Take Field Names from Right Input Rows allows us to use a definition file to rename the fields.

Take Field Names from Right Input Metadata allows us to use a data stream that we know matches to rename the fields.

Take Field Descriptions from Right Input Metadata allows us to use a data stream that we know matches to redefine the fields.

5	The JSON F JSON form tables.	Parse tool alle	ows us to read into relational
	Group	Input	Output
	Developer	JSON	See below
(cont)	-	formatted	
Figure 10-10 – JSON Parse		data	
		stream	
Note: For these unfor			NI stands (su

Note: For those unfamiliar with JSON: JSON stands for JavaScript Object Notation, which is a structured data format typically used in data transfer.

An *Action* tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Output: A data stream in relational data format.

Properties Window:

The *JSON Parse Configuration* window has three components.



Figure 10-11 - JSON Parse Configuration

- *JSON Field* is the field in the incoming data stream that has JSON data written with one complete record in the cell.
- *Include in Output,* when checked, will maintain the JSON field when it is output from this tool.
- *The Output Values* options allow us to decide if each of the data values should be included in a single string field or in a field defined by the JSON data type.

10.7 Message

Q F	The <i>Message</i> message in th	tool allows ne output w	us to create a indow.
	Group	Input	Output
	Developer	Any data stream	Any data stream
Figure 10-12 - Message			

Note: Message helps find out where something went wrong and log the settings during a run. It also provides updates about parts of our workflow that have completed.

Application questions can be connected to the top *Black Question Anchor* to use those answers in this tool.

An Action tool can be connected to the *Lightning Bolt Anchor* to modify how this tool works in apps and macros.

Properties Window:

The *Message Configuration* window has five basic components.

M	essage (17) - Configuration	- ₽ X
~	When To Send Message:	
3	Before First Record	
	Before Rows Where Expression is True	
→		
0	After Last Record	
EQ.	After All Down Stream Tools Have Completed	
0	Message Type:	
	Message	•
	Message Priority:	
	Normal	-
	Transient	
	Message Expression:	
	"This expression is written to the output window"	
0		
-		

Figure 10-13 - Message Configuration

- *When To Send Message* defines what the trigger for writing out the message is. It can be just before the first record, before every record where a Boolean expression is true, after the last record has passed through the tool, or when the data stream has terminated.
- *Message Type* defines what type of output should be written out and if it should stop the data flow (see Output Window in The Interface for more details).
- *Message Priority* determines what the message would do by default if it were in a macro. Normal will show the message if the tool is in the workflow. Medium will show

the message if the tool is directly in a macro used in the workflow. High will show the message no matter how many macros deep the message occurred.

- *Transient,* when checked, means that if the Message tool writes more than one message, it will replace the single message instead of creating a new one every time.
- *Message Expression* is an expression field that will output the result to the output window.

10.8 All the Tools Along the Way

То	Alteryx Consultants
Subject	All The Tools Along the way
Attached	ISON.txt
Неу,	
l have bee higher-up	en asked to make sure you know how to use everything in a list of tools from the is.
There are to go over	a few that you haven't gotten to know yet, so I have designed something for us r and then for you to work through.
Attached, such that structure.	you will find a file called JSON.txt. We will be using it to create an application when you check a field, the JSON data will be converted into a relational If you don't check the field, it will output a file identical to the original JSON file.
Thanks,	

The option for a tool to essentially just copy the file seems a little strange, but it will make a more sense when we start working on the follow-up assignment. We can write the file to a .csv instead of a .txt file at the end of the workflow. Let us start by connecting to the data.



Figure 10-14 - Open JSON File

	6.	
Space	Nor	e
		Selectio
-		
	Selv Space	Selection Space

Figure 10-15 – Read JSON File

_	_	Field	Туре		Size	Rename	Description
	V	Field_1	V_String	•	254		
	1	*Unknown	Unknown	-	0		Dynamic or Unkr

Figure 10-16 - JSON Select

1 of 1 Fields 🔻 🖌 🛛 Cell Viewer 👻 🏦 🖬 🖬 🖬 🛨					
Record #	ecord # Field_1				
1	[{"Category":"Predictive","Module":"AB Analysi				
2	{"Category":"Predictive","Module":"AB Control				
3	{"Category":"Predictive","Module":"AB Treatm				
4	{"Category":"Predictive","Module":"AB Trend",				
5	{"Category":"Interface","Module":"Action","In				
6	{"Category":"Interface","Module":"Action","In				
7	{"Category":"Interface","Module":"Action","In				
8	{"Category":"Interface","Module":"Action","In				
9	{"Category":"Demographic Analysis","Module"				
10	{"Category":"Demographic Analysis","Module"				

Figure 10-17 - JSON Browse

When we connect to this file, because it is *.txt*, Alteryx needs to know how it should read it. This is by design because file formats like JSON need to be read so that all data is in a single field and ideally one record per row. Because we want these settings, we can choose to have the file delimited with no delimiter so that nothing will break the field. If we take a look at the file, we can see there is no header row. This is because JSON

carries the header names in every record and thus does not need a devoted row.



Figure 10-18 - JSON Detour

Our next step will be to add a *Detour* tool. This tool allows us to use a question in order to decide which data stream to follow. The *Detour to the Right* option is so that while we are testing our workflow, we can properly check both sides. By design, the Detour tool detours to the left unless updated.



Figure 10-19 - JSON Detour to the Left

Let's start by designing the left (L) data stream. We are going to make this the side where we do not convert the data to a relational structure. We are going to use *Message* tools to record what will happen to the data. The Message tool allows us to write something to the output window. In this case, we are going to write *The data will remain in JSON format* so that it will be in the output log.



Figure 10-20 - JSON Message Configuration



Figure 10-22 - JSON Dynamic Rename

Now let us add a *Dynamic Rename* tool. This is because once we bring the data back together to write, we will need to write the header. Since we don't want a header for this option on the data stream, we will modify it such that the header is actually the first record of data. This will complete all of the data prep we need before rejoining the data streams to be written out.

JSON.btt Figure 10-23 - JSON Right data stream
Message (8) - Configuration
When To Send Message: Image: Second
After Last Record After All Down Stream Tools Have Completed
Message Type:
Message
Normal
Transient Message Expression: The data will be converted to a relational table [*]
 Type the message

Figure 10-24 - JSON Message configuration for Right Side data stream

We are going to start the right (R) side of the data stream by creating a parallel message to the left that reads *The data will be converted to a relational table*.



Figure 10-25 - JSON Right side data stream completed

We will introduce a *Formula* tool now and convert each row of the incoming data into an array by introducing '[' and ']' symbols at the start and end of each row.



Figure 10-25a - Formula Field Calculation

We then introduce Record ID tool to number each row which will eventually help us in grouping the data as we shall see. Now, we will bring the *JSON Parse* tool onto the canvas. This tool is designed to look at JSON data and convert it into relationally structured data. Unfortunately, if we just put the tool at the end of the data stream, it does not recognize any fields. This is because the *Detour* is diverting all information to the left, including the metadata. In order to configure the tools on this side of the *Detour*, we need to divert the tool to the right.

JSON Field Include in Output Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields Image: Second state type specific fields	
•	
Results - Browse (8) -	3
Record # Record ID JSON Name JSON ValueString	D
1 1 0.Category Predictive	-
2 1 0.Module AB Analysis	
3 1 0.In Book FALSE	
4 1 0.Chapter of Description	
5 1 0.Exercise Introduced	
6 1 0.Exercise Used	
7 2 0.Category Predictive	
8 2 0.Module AB Controls	
9 2 0.In Book FALSE	
10 2 0.Chapter of Description -	

Figure 10-26 - JSON Parse converting data to relationally structured data

Now that we have restructred the data, we should take a look at the result so we know what our next steps should be. We can see the JSON Parse tool has created two fields: *JSON_Name*, which has the record number and field name separated by a period, and *JSON_ValueString*, which has the values associated with each record and variable.

Note: If we wanted to convert data into JSON format using the *JSON Build* tool, this is the format it should be in.

Since we know we want a relational dataset that is easy to work with, we should make it tidy by creating a single column for each variable.

In order to do this, we first need to identify the unique column names by splitting them off of the record number. We can use the *Text To Columns* tool in order to create a separate column for both the record number and the column name so we can use the two of them to restructure the data.



Figure 10-27 - JSON - Splitting the columns and renaming them

In order to make the data easier to understand we should rename our fields.



Figure 10-28 - JSON - Relationally structured data

Now that we have our fields well named, we can pivot the data into a more usable format. We are going to group the data by the *Record ID*, use the *Field Names* as our columns, and use the value field as the data in the intersection. We need to set a methodology. However, since we are using a unique combination of header and grouping fields, we do not need to worry about clashing data.

C	ross Tab (6) - Configuration	-	中)
	Grouping Fields RecordID JSON Variable Value Row ID Column Name		
2	Header Field Column Name Data Field		•
	Value Methodologies (String Data Field)		•
	Concatenate First Last		
2	Separator: , Field Size: 2048		

Figure 10-29 - JSON - Data grouping by Record ID

7 of 7 Field	is 🕶 🕪 🛛 🕻	Cell Viewer 🔻	1 1 292 records displayed,	12 KB			🕛 • 🕞 •	÷
Record #	RecordID	Category	Chapter_of_Description	Exercise_Introduced	Exercise_Used	In_Book	Module	
1	1	Predictive				FALSE	AB Analysis	
2	2	Predictive				FALSE	AB Controls	
3	3	Predictive				FALSE	AB Treatments	
4	4	Predictive				FALSE	AB Trend	
5	5	Interface	Traffic Tickets In Baltimore	Lets Build An App For That!	A Coffee Fix	TRUE	Action	
6	6	Interface	Traffic Tickets In Baltimore	Lets Build An App For That!	A Coffee Fix Easy Mapping	TRUE	Action	
7	7	Interface	Traffic Tickets In Baltimore	Lets Build An App For That!	All The Tools Along The Way	TRUE	Action	٦,

Figure 10-30 - JSON - Grouped dataset

If we take a look at the data, we can see we have a tidy dataset, which means we can end this specialized portion of the workflow and bring the two data streams back together.



Figure 10-31 - JSON - Bringing both data streams together

Now we can connect both sides of the data stream to a *Detour End* tool to bring them back together so we can use a single output.



Figure 10-32 - JSON - Using Detour End Tool to bring together the data streams

Once we add the Output Data tool, we are almost done. As we can see, we have a fully operational data stream. The only problem we have is that we have no way to control if the file will be written as JSON or a relational table.

We are going to add a *Check Box* tool to the workflow so we can give the option of selecting between converting the data or not. If we connect the question directly to the *Detour* tool, the action will be created exactly how we want it to work, except that
we should take a look at the behavior of this action because it doesn't just change the direction, like implied by the action type.



Figure 10-33 – JSON – Adding Check Box to provide an option for conversion

Any time we are running the Detour tool as an application, the default path is to go to the left. When we have an action to update this behavior, it will change the result to divert to the right. We should name the question to reflect this behavior.



Figure 10-34 – JSON data stream on completion

10.9 Narrowing Down the Search

То	Alteryx Consultants	
Subject	Narrowing Down The Search	
Hey,		
This has come along nicely, but here's where it gets complicated; I need you to add a question that will filter the data streams to the specific tool you type into it.		
Thanks		

CHAPTER 11 NYC

Congratulations! You have made it through all of the basic training and are ready to take on a client of your own. Let me know how it goes when you're done.

То	Alteryx Consultants	
Subject	NYC	
Attached	andmarks.xisx; an NYC Subway Entrances.csv; an NYC Free Public WiFi 12052014.csv	
Hey,		
I just got off the phone with a new client, and based on what I've seen, you are ready to handle the engagement on your own.		
They are a travel agency and want an application built so that they can promote trips to NYC.		
They want to be able to select an area on a map, include three types of locations to select from, and choose to output a report with a map or file with the data. They provided the three attached files, which designate the three different options they should be able to select from.		
You know enough to build this. If you are having trouble getting started, make a list of the steps you need to do, and remember that looking through your previous work will usually help you solve your problems.		
Thanks.		

CHAPTER 12 Self-Guided Solutions

Readers are strongly encouraged not to use these solutions when solving the self-guided examples. There are many approaches to answering these problems, and figuring out how to do it on your own is the goal of these exercises. These screenshots of solutions are provided so that if you're stuck, you can refer to the logic of how the examples direct you to a solution.



12.1 The Games: The Almighty Dollar

Figure 12-1 - Almighty Dollar



12.2 Unisex Baby Names: What's In a Name?

Figure 12-2 – What's in a name?

12.3 The Direct Approach: Where, Oh Where Have My Three Files Gone?



Figure 12-3 - Where, Oh Where have the three files gone?

12.4 Cultural Musing: Culturally Divided



Figure 12-4 - Culturally Divided

12.5 The Sport Report: FOOTBALL!



Figure 12-5 - Football

12.6 Expensive Beauty Products: More Flags

Since the complete workflow is large, it is broken down into segments for better understanding.



Figure 12-6 - More Flags



Figure 12-6a - Incoming Data



Figure 12-6b - Regular Expressions



Figure 12-6c - Fuzzy Logic



Figure 12-6d - Join & Group



Figure 12-6e - Output

12.7 Applications Wanted: To Summarize or Not to Summarize: That is the Question.

Since the complete workflow is large, it is broken down into segments for better understanding.

4		5		
	Interface Actions for Data Stream	•	Interface Actions for Tool Container Activity	۲
	3			
	Bas	sic data s	tream 🛞	

Figure 12-7 - To Summarize or not to Summarize



Figure 12-7a - Interface Actions for Data Stream



Figure 12-7b - Interface Actions for Tool Container Activity



Figure 12-7c - Basic data stream

12.8 Where's The Joe? Zip 5 Total Population Macro and Workflow



Figure 12-8b - Who wants coffee? - Workflow



12.9 Metamorphosis: Narrowing Down the Search

Figure 12-9 - Metamorphosis

12.10 Capstone: NYC

Since the entire workflow is large, it is broken down into segments for better understanding.



Figure 12-10 - NYC



Figure 12-10a - Action tools for mapping



Figure 12-10b - Gather data, Create Special Points and Filter by Special Match



Figure 12-10c - Action tools for Data processing and Detour



Figure 12-10d - Filtering, Mapping and Layout Configuration

Appendices

Appendix A – File Types

.aws	Allocate Workspace	
.cydb	Calgary database	
.pcxml	XML data output type that allows the end user to select from multiple file types for a final output when using Alteryx Gallery	
.yxdb	Alteryx database	
.yxft	Alteryx Header File	
.yxlc	Alteryx License File	
.yxmc	Alteryx Macro File	
.yxmd	Alteryx Workflow File	
.yxwv	.yxwv Alteryx Analytic App Values File which allows predefined selections to be made.	
.yxwz	Alteryx Analytic App File	
.yxzp	Packaged Alteryx File	

Appendix B – Hot Keys

Ctrl+Alt+B	Show/Hide Toolbar
Ctrl+Alt+D	Show/Hide Interface Designer
Ctrl+Alt+O	Show/Hide Output
Ctrl+Alt+P	Show/Hide Properties
Ctrl+Alt+T	Show/Hide Tool Palette
Ctrl+Alt+V	Show/Hide Overview
Ctrl+C	Сору
Ctrl+F	Fine
Ctrl+F4	Close workflow
Ctrl+N	New workflow
Ctrl+O	Open workflow
Ctrl+R	Run/Cancel Workflow
Ctrl+S	Save workflow
	Add a Browse tool after all selected
Ctrl+Shift+B	tools that a Browse tool can be
	connected to
Ctrl+V	Paste
Ctrl+X	Cut
Ctrl+Y	Redo
Ctrl+Z	Undo
Ctrl++	Vertically align selected tools
Ctrl+-	Horizontally align selected tools
Del (Delete)	Deletes selected part of workflow
F1	Open Help menu
F5	Refresh Configuration

Appendix C - Downloads/Content

US 2010 Census SF1 http://downloads.alteryx.com/data.html

USGS North America Map http://downloads.alteryx.com/data.html

Visual Analytics Kit (Tableau) http://alteryx.com/kit

Visual Analytics Kit (Qlik) http://pages.alteryx.com/VisualAnalyticsKitforQlik_Reg-LP.html

Alteryx Gallery https://gallery.alteryx.com

Appendix D – Field Types

Туре	Class	Description
		A field with two values, True and
Bool	Boolean	False
		A field with integers between 0 and
Byte	Numeric	255
		A field with integers between -
Int16	Numeric	32,768 and 32,767
		A field with integers between
Int32	Numeric	-2,147,483,648 and 2,147,483,647
		A field with integers between
		-9,223,372,036,854,775,808 and
Int64	Numeric	9,223,372,036,854,775,807
		A field with a specific width
Fixed		including the number of decimal
Decimal	Numeric	places
		A field with up to 7 digits of
		accuracy between -3.4^38 and
Float	Numeric	3.4^38
		A field with up to 15 digits of
		accuracy between -1.7^308 and
Double	Numeric	1.7^308
String	String	A field with up to 8192 characters
		A field with up to 8192 characters
WString	String	that will accept Unicode characters
		A field that will adjust the size
		depending on the length of the
V_String	String	strings in it

		A field that will adjust the size	
		depending on the length of the	
		strings in it and will accept	
V_WString	String	Unicode characters	
		A field in the format "yyyy-mm-	
Date	String	dd"	
Time	String	A field in the format "hh:mm:ss"	
		A field in the format "yyyy-mm-dd	
DateTime	String	hh:mm:ss"	
Blob	Blob	A field with image or sound files	
		A field with points, lines, polylines,	
SpatialObj	Blob	or polygons.	

Appendix E- Properties Window



Figure 12-11 - Properties Window

Appendix F - Boolean Expressions

Expression	Meaning
A = B	Expression A equals expression B
A != B	A does not equal B
A > B	A is greater than B
A >= B	A is greater than or equal to B
A < B	A is less than b
A <= B	A is less than or equal to B
А	A is True
NOT A	A is False
A IN B	A is in the set of B things
A NOT IN B	A is not in the set of B things
A OR B	Expression A is True or expression B is True or both expressions are True
A AND B	Expression A is True and expression B is True
A OR (B AND C)	Expression A is True or expression B and expression C are both True or expression A and expression B and C are True
A AND (B OR C)	Expression A is True and expression B, expression C, or expressions B and C are True

Basic Boolean expressions can be combined to create more complex logic by replacing A, B or C with any basic Boolean Expression.

Appendix G – Data Components

day - Spelled out weekday
any opened out weeking
dd - 2 digit day of month
dy - Abbreviated day of week
hh – 2 digit hour
mm – 2 digit minute
MM - 2 digit month of year
Mon - Abbreviated month of year
Non noneviace month of year
Month - Spelled out month
ss – 2 digit second
0.1:
yy – 2 digit year
yyyy – 4 digit year

Figure 12-12 – Date Components
Appendix H – Date/Time Units

Keywords
Years
Months
Days
Hours
Minutes
Seconds

Appendix I – RegEx Cheat Sheet

	Any Character
\$	End of a line
0	Marked Group
(?:)	Unmarked Group
*	Repeat the previous thing 0 or more times
+	Repeat the previous thing 1 or more times
[]	A set to be used for optional lists of single characters
[^]	A set to be used for optional lists of single characters to be excluded
[[:alpha:]]	Any letter
\setminus	Escape the following symbol
\<	Beginning of a word
\>	End of a word
\d	Digit
$\setminus 1$	Lowercase letter
∖n	New line character
$\setminus s$	Space
\t	Tab character
\u	Uppercase letter
$\setminus w$	Word characters
∖W	Non-word characters
^	Beginning of a new line or "not" depending on context
	Or

Appendix J - Action Tool Sets

These tools do not receive connections from an Action tool.

		Folder	Numeric Up
Action	Date	Browse	Down
Checkmar			
k	Drop Down	List Box	Radio Button
	Error		
Comment	Message	Macro Input	Text Box
		Macro	
Condition	Explorer Box	Output	Tree
Control			
Parameter	File Browse	Мар	

All other tools have at least these basic methods

Delete Tool from Workflow Update Field List Update Raw XML with Formula Update Select with Reverse Field Map Update Value Update Value

These tools only have the basic methods.

	Data Stream		Report:
AB Analysis	In	Join In-DB	Comparison
	Data Stream		Report:
AB Controls	Out	Join Multiple	Detail
AB	DataSift		Report:
Treatments	Connector	JSON Build	Rank

=

			Run
AB Trend	Date Filter	JSON Parse	Command
		K-Centroids	
Allocate	Date Time	Cluster	Running
Append	Now	Analysis	Total
Allocate		K-Centroids	
Input	DateTime	Diagnostics	Sample
Allocate	Decision		Sample In-
Metainfo	Tree	Layout	DB
Allocate			
Report	Detail Fields	Lift Chart	Scatterplot
Amazon S3		Linear	
Download	Detour End	Regression	Score
Amazon S3		Logistic	
Upload	Distance	Regression	Select In-DB
	Distribution		Select
API Output	Analysis	Make Columns	Records
Append	Dynamic		
Cluster	Input	Make Grid	Smooth
Append	Dynamic		
Fields	Rename	Make Group	Spatial Info
	Dynamic	Map Legend	Spatial
ARIMA	Replace	Builder	Match
	Dynamic	Map Legend	Spatial
Arrange	Select	Splitter	Process
Association		Marketo	Spearman
Analysis	Email	Append	Correlation
			Spline
Auto Field	ETS	Marketo Input	Model
Base 64		Marketo	
Encoder	Field Info	Output	Stepwise
Behavior	Field		Street
Metainfo	Summary	MB Inspect	Geocoder

Blob Convert	Filter	MB Rules	Summarize
			Summarize
Blob Output	Filter In-DB	Message	In-DB
		U	Support
Block Until		Multi-Field	Vector
Done	Find Nearest	Binning	Machine
Boosted	Find Nearest	Multi-Field	
Model	Neighbors	Formula	Table
			Tableau
		Multi-Row	Workbook
Browse	Find Replace	Formula	Macro
Browse Data		Naïve Bayes	
In-DB	Footer	Classifier	Test
			Test of
Buffer	Forest Model	Nested Test	Means
Business		Neural	Text To
Match (US)	Formula	Network	Columns
Calgary	Formula In-	Non Overlap	
Input	DB	Drivetime	Throttle
	Foursquare		
Calgary Join	Search	Overlay	Tile
Calgary	Frequency	Oversample	
Loader	Table	Field	Transpose
Canada			
Geocoder	Fuzzy Match	Parse Address	TS Compare
			TS
	Gamma	Pearson	Covariate
CASS	Regression	Correlation	Forecast
Charting	Generalize	Plot of Means	TS Filler
	Generate		
Cluster Code	Rows	Poly-Build	TS Forecast
Compare			
Behavior	Gnip Search	Poly-Split	TS Plot

Connect In-	Google	Principal	Twitter
DB	Analytics	Components	Search
ConsumerVi		1	
ew Matching	HDFS Input	Profile Input	Union
Contingency	HDFS		Union In-
Table	Output	Profile Output	DB
Count			
Records	Header	R	Unique
Count		Random %	US
Regression	Heat Map	Sample	Geocoder
		-	US Zip 9
Create Points	Heat Plot	Record ID	Coder
Create			
Profile	Histogram	RegEx	Violin Plot
	Household		
Create	File		Weighted
Samples	Matching	Render	Average
Cross Count	Image	Report Map	Write In-DB
Cross Count			
Append	Imputation	Report Text	XML Parse
Cross Tab	Join		

The following tools have the basic tools plus

Update Password

	MongoDB	Salesforce	SharePoint
Download	Input	Input	List Input
	MongoDB	Salesforce	SharePoint
	Output	Output	List Output

The following tools have the basic tools plus

Update Input Data Tool

|--|

The following tools have the basic tools plus unique methods

Detour	Update Detour Direction from Condition
	Update Detour Direction from Question
Directory	Update Directory with Folder Browse
Map Input	Update Map Input with Map Question
Output Data	Update Output Data Tool
Select	Update Select with multi-select ListBox
Sort	Update Sort Order
Text Input	Update Cell
Tool	Enable/Disable Container From Condition
Container	Enable/Disable Container From Question
Trade Area	Update Trade Area Radius

Appendix K – Directory Tool Data

FullPath	The direct path to open the file.
Directory	The direct path to the folder the file is
	in.
FileName	The actual name and extension of the
	file.
ShortFileName	A short file name provided for
	compatibility with legacy software.
CreationTime	This indicates the first time the file was
	created.
LastAccessTime	This indicates the last time the file was
	opened.
LastWriteTime	This indicates the last time the file was
	saved.
Size	This is the total size of the file in bytes.
AttributeArchive	This indicates if the file is archived.
AttributeCompressed	This indicates if the file is compressed.
AttributeEncrypted	This indicates if the file is encrypted.
AttributeHidden	This indicates if the file is hidden.
AttributeNormal	This indicates if the file has normal
	attributes.
AttributeOffline	This indicates if the data is unavailable.
AttributeReadonly	This indicates if the file is read only.
AttributeResparsePoint	This indicates that the file is associated
	with a resparse point.
AttributeSparseFile	This indicates that the file is sparse.
AttributeSystem	This indicates that the file is a system
	file.
AttributeTemporary	This indicates that the file is a
1 5	temporary file.

Credited Original Data Sources

The Games

- Medals:
- Access through Tableau Public and assumed to be in the public domain https://public.tableau.com/profile/a.m.5517#!/vizhome/OlympicMedals2000-2012_0/Dashboard2 published by A.M.
- Nobel Laureates:
- Access through Tableau Public and assumed to be in the public domain: http://tableausoftware.co.nz/public/gallery/nobelprizes published by Ross Perez
- GDP per Capita:
- Accessed through GitHub https://github.com/cschin/ipython_d3_mashup/blob/master/ipython_13_vis_example/ gm_data/GDPpercapitaconstant2000US.csv published by Jason Chin

Unisex Baby Names

- Unisex Names:
- Access through Tableau Public and assumed to be in the public domain http://tableausoftware.co.nz/public/gallery/unisex-baby-names_published by Steve Ruble

The Direct Approach

- Created by Author
- Special Thanks to http://watchout4snakes.com/ for use of their random word generator in creating the names of the folders in the exercise Where, Oh Where Have The Three Files Gone?

Cultural Musings

- GDP per Capita:
- List of U.S. states by GDP per capita. (2015, March 28). In Wikipedia, The Free Encyclopedia. Retrieved 00:01, May 12, 2015, fromhttp://en.wikipedia.org/w/index.php?title=List_of_U.S._states_by_GDP_per_capita &oldid=653869588
- Museum Data:
- Data retrieved from Data.Gov (http://data.gov) Direct Link: http://catalog.data.gov/dataset/museums-universe-data-file-mudf-fy-2014-3rdquarter/resource/d4b36e0f-1d9d-4c2a-8882-138c15763717
- State Map:
- Created by Author

Sport Report

- Major US Sport Teams:
- o The author created this data using Wikipedia pages and tools
- Primary Wikipedia Article
- List of professional sports teams in the United States and Canada. (2015, March 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 14:34, January 7, 2015, from

 $http://en.wikipedia.org/w/index.php?title=List_of_professional_sports_teams_in_the_United_States_and_Canada&oldid=650502141$

- The Geocoded locations of each of the sites were accessed by using the coordinates link embedded in their web pages with utilizes a Wikimedia Tool called GeoHack which can be accessed directly at http://tools.wmflabs.org/geohack/
- Turner Field. (2015, March 22). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:02, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Turner_Field&oldid=653079052
- Marlins Park. (2015, March 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:03, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Marlins_Park&oldid=649550406
- Citi Field. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:03, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Citi_Field&oldid=653482419
- Citizens Bank Park. (2015, March 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:03, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Citizens_Bank_Park&oldid=649435153
- Nationals Park. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:03, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Nationals_Park&oldid=655186324
- Wrigley Field. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:04, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Wrigley_Field&oldid=655275853
- Great American Ball Park. (2015, April 7). In Wikipedia, The Free Encyclopedia. Retrieved 15:04, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Great_American_Ball_Park&oldid=6552856 31

• Miller Park (Milwaukee). (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:04, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Miller_Park_(Milwaukee)&oldid=655121188

- PNC Park. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:05, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=PNC_Park&oldid=654838712
- Busch Stadium. (2015, March 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:05, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Busch_Stadium&oldid=649550165
- Chase Field. (2015, April 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:05, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Chase_Field&oldid=655057408
- Coors Field. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:09, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Coors_Field&oldid=654538824
- Dodger Stadium. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:09, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Dodger_Stadium&oldid=654869845
- Petco Park. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:10, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Petco_Park&oldid=654538750
- AT&T Park. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:10, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=AT%26T_Park&oldid=654538692

• Oriole Park at Camden Yards. (2015, March 18). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:11, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Oriole_Park_at_Camden_Yards&oldid=651 898983

- Fenway Park. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:11, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Fenway_Park&oldid=653504103
- Yankee Stadium. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:11, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Yankee_Stadium&oldid=654863343
- Tropicana Field. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:11, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Tropicana_Field&oldid=655160166
- Rogers Centre. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:12, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Rogers_Centre&oldid=654913445
- U.S. Cellular Field. (2015, March 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:12, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=U.S._Cellular_Field&oldid=649856024
- Progressive Field. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:12, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Progressive_Field&oldid=653499953
- Comerica Park. (2015, March 17). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:13, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Comerica_Park&oldid=651832995
- Kauffman Stadium. (2015, March 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:13, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Kauffman_Stadium&oldid=650377832
- Target Field. (2015, February 19). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:13, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Target_Field&oldid=647933510
- Minute Maid Park. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:13, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Minute_Maid_Park&oldid=655303381
- Angel Stadium of Anaheim. (2015, April 1). In Wikipedia, The Free Encyclopedia. Retrieved 15:13, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Angel_Stadium_of_Anaheim&oldid=654460 841

- •O.co Coliseum. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:13, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=O.co_Coliseum&oldid=653612043
- Safeco Field. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:14, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Safeco_Field&oldid=655148228
- Globe Life Park in Arlington. (2015, March 22). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:14, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Globe_Life_Park_in_Arlington&oldid=6529 59003

- TD Garden. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:14, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=TD_Garden&oldid=653403579
- Barclays Center. (2015, April 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:14, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Barclays_Center&oldid=654631943
- Madison Square Garden. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:14, April 7, 2015,
- $from \ http://en.wikipedia.org/w/index.php?title=Madison_Square_Garden\&oldid=654828742$
- Wells Fargo Center (Philadelphia). (2015, March 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 7, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Wells_Fargo_Center_(Philadelphia)&oldid= 649738895

- Air Canada Centre. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Air_Canada_Centre&oldid=655130412
- United Center. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=United_Center&oldid=655268263
- Quicken Loans Arena. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 7, 2015,
- from http://en.wikipedia.org/w/index.php?title=Quicken_Loans_Arena&oldid=655272955
- The Palace of Auburn Hills. (2015, February 22). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=The_Palace_of_Auburn_Hills&oldid=64827 2487

- Bankers Life Fieldhouse. (2014, December 29). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015,
- $from \ http://en.wikipedia.org/w/index.php?title=Bankers_Life_Fieldhouse\&oldid=640160898$
- BMO Harris Bradley Center. (2015, February 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=BMO_Harris_Bradley_Center&oldid=64564 6883

- Philips Arena. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Philips_Arena&oldid=654538478
- Time Warner Cable Arena. (2015, March 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Time_Warner_Cable_Arena&oldid=6533736 73

• American Airlines Arena. (2015, February 10). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:16, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=American_Airlines_Arena&oldid=64642698 3

- Amway Center. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:17, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Amway_Center&oldid=655268871
- Verizon Center. (2015, March 14). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:17, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Verizon_Center&oldid=651299892
- Pepsi Center. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:17, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Pepsi_Center&oldid=653560944
- Target Center. (2015, April 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:17, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Target_Center&oldid=655045980
- Chesapeake Energy Arena. (2015, March 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Chesapeake_Energy_Arena&oldid=6500533 51

• Moda Center. (2015, March 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Moda_Center&oldid=652902770

• EnergySolutions Arena. (2015, March 30). In Wikipedia, The Free Encyclopedia. Retrieved 15:18, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=EnergySolutions_Arena&oldid=654209905

- Oracle Arena. (2015, February 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Oracle_Arena&oldid=648884284
- Staples Center. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, April 7, 2015, from http://en.wikipedia.org/w/index.php?title=Staples_Center&oldid=654943633
- Talking Stick Resort Arena. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:19, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Talking_Stick_Resort_Arena&oldid=654788 366

•Sleep Train Arena. (2015, February 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:19, April 7, 2015,

from http://en.wikipedia.org/w/index.php?title=Sleep_Train_Arena&oldid=648040550 •AT&T Center. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:40, April 8,

- 2015, from http://en.wikipedia.org/w/index.php?title=AT%26T_Center&oldid=654184289 •Smoothie King Center. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:40,
- Smoothie King Center. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:40, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Smoothie_King_Center&oldid=653411690
- FedExForum. (2015, March 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:40, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=FedExForum&oldid=649655797
- Toyota Center. (2015, February 28). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:41, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Toyota_Center&oldid=649249468
- American Airlines Center. (2015, March 29). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:41, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=American_Airlines_Center&oldid=65400062

•Sleep Train Arena. (2015, February 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:19, April 7, 2015,

 $from\ http://en.wikipedia.org/w/index.php?title=Sleep_Train_Arena&oldid=648040550$

- Philips Arena. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Philips_Arena&oldid=654538478
- Allstate Arena. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Allstate_Arena&oldid=654835689
- Mohegan Sun Arena. (2015, March 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Mohegan_Sun_Arena&oldid=650179511
- Bankers Life Fieldhouse. (2014, December 29). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015,

 $from\ http://en.wikipedia.org/w/index.php?title=Bankers_Life_Fieldhouse\&oldid=640160898$

•Madison Square Garden. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015,

 $from \ http://en.wikipedia.org/w/index.php?title=Madison_Square_Garden\&oldid=654828742$

• Verizon Center. (2015, March 14). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Verizon_Center&oldid=651299892

- Staples Center. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:43, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Staples_Center&oldid=654943633
- Target Center. (2015, April 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Target_Center&oldid=655045980
- Talking Stick Resort Arena. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Talking_Stick_Resort_Arena&oldid=654788 366

- AT&T Center. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=AT%26T_Center&oldid=654184289
- KeyArena. (2015, March 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=KeyArena&oldid=653237504
- BOK Center. (2015, March 31). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=BOK_Center&oldid=654259841
- •Santander Arena. (2015, February 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Santander_Arena&oldid=648485678
- Odeum Expo Center. (2015, February 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Odeum_Expo_Center&oldid=648219741
- The Gardens Ice House. (2014, December 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=The_Gardens_Ice_House&oldid=636303525

- Tiger Arena. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Tiger_Arena&oldid=653415784
- Dow Event Center. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Dow_Event_Center&oldid=654204684

• York City Ice Arena. (2014, October 10). In Wikipedia, The Free Encyclopedia. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=York_City_Ice_Arena&oldid=629018943

• Talking Stick Resort Arena. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Talking_Stick_Resort_Arena&oldid=654788 366

- Honda Center. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Honda_Center&oldid=654822150
- Alamodome. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Alamodome&oldid=654207806
- Moda Center. (2015, March 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Moda_Center&oldid=652902770
- SAP Center at San Jose. (2015, April 2). In Wikipedia, The Free Encyclopedia. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=SAP_Center_at_San_Jose&oldid=654649938

• Spokane Veterans Memorial Arena. (2015, March 29). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, fromhttp://en.wikipedia.org/w/index.php?title=Spokane_Veterans_Memorial_Arena&oldid =654081587

• Jacksonville Veterans Memorial Arena. (2015, March 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Jacksonville_Veterans_Memorial_Arena&ol did=652941155

- CFE Arena. (2015, February 12). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=CFE_Arena&oldid=646752542
- Amalie Arena. (2015, February 12). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:44, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Amalie_Arena&oldid=646850271
- Quicken Loans Arena. (2015, April 7). In Wikipedia, The Free Encyclopedia. Retrieved 13:44, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Quicken_Loans_Arena&oldid=655272955

• Smoothie King Center. (2015, March 25). In Wikipedia, The Free Encyclopedia. Retrieved 13:45, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Smoothie_King_Center&oldid=653411690

• Wells Fargo Center (Philadelphia). (2015, March 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Wells_Fargo_Center_(Philadelphia)&oldid= 649738895

• Tim Hortons Field. (2015, March 17). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Tim_Hortons_Field&oldid=651817295

• Percival Molson Memorial Stadium. (2015, January 16). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Percival_Molson_Memorial_Stadium&oldid =642808624

• TD Place Stadium. (2015, February 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=TD_Place_Stadium&oldid=648816970

- Rogers Centre. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Rogers_Centre&oldid=654913445
- BC Place. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=BC_Place&oldid=653675532
- McMahon Stadium. (2015, February 28). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=McMahon_Stadium&ol Gillette Stadium. (2015, April 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Gillette_Stadium&oldid=654585813did=649 207077

• Commonwealth Stadium (Edmonton). (2015, March 4). In Wikipedia, The Free Encyclopedia. Retrieved 13:45, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Commonwealth_Stadium_(Edmonton)&oldi d=649791012

• Mosaic Stadium at Taylor Field. (2015, February 15). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Mosaic_Stadium_at_Taylor_Field&oldid=64 7219959

- Investors Group Field. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Investors_Group_Field&oldid=655386188
- Ralph Wilson Stadium. (2015, April 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Ralph_Wilson_Stadium&oldid=655091171 •Sun Life Stadium. (2015, March 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8,
- 2015, from http://en.wikipedia.org/w/index.php?title=Sun_Life_Stadium&oldid=652918431 •MetLife Stadium. (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8,
- 2015, from http://en.wikipedia.org/w/index.php?title=MetLife_Stadium&oldid=655501155
- •M&T Bank Stadium. (2014, December 16). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:45, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=M%26T_Bank_Stadium&oldid=638418974 • Paul Brown Stadium. (2014, December 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:46, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Paul_Brown_Stadium&oldid=637052857
- FirstEnergy Stadium (Cleveland). (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:46, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=FirstEnergy_Stadium_(Cleveland)&oldid=6 54531854
- Heinz Field. (2015, March 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:47, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Heinz_Field&oldid=653482405
- NRG Stadium. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:47, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=NRG_Stadium&oldid=655407096
- Lucas Oil Stadium. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:47, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Lucas_Oil_Stadium&oldid=655376758
- EverBank Field. (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:47, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=EverBank_Field&oldid=653161581
- LP Field. (2014, November 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:47, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=LP_Field&oldid=634894166
- Sports Authority Field at Mile High. (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015,
- fromhttp://en.wikipedia.org/w/index.php?title=Sports_Authority_Field_at_Mile_High&oldi d=653126660
- Arrowhead Stadium. (2015, March 29). In Wikipedia, The Free Encyclopedia. Retrieved 13:48, April 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Arrowhead_Stadium&oldid=654006833 •O.co Coliseum. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8,
- 2015, from http://en.wikipedia.org/w/index.php?title=O.co_Coliseum&oldid=653612043
 •Qualcomm Stadium. (2015, March 27). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April
- Qualcomm Stadium. (2015, March 27). In *Vvikipeala, The Free Encyclopeala*. Retrieved 13:48, A 8, 2015,
- from http://en.wikipedia.org/w/index.php?title=Qualcomm_Stadium&oldid=653741585
- AT&T Stadium. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=AT%26T_Stadium&oldid=654731734

- MetLife Stadium. (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=MetLife_Stadium&oldid=655501155
- Lincoln Financial Field. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015,
- $from \ http://en.wikipedia.org/w/index.php?title=Lincoln_Financial_Field\&oldid=655256338$
- FedExField. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=FedExField&oldid=654429456
- Soldier Field. (2015, March 12). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Soldier_Field&oldid=651010515
- Ford Field. (2015, March 28). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Ford_Field&oldid=653890809
- Lambeau Field. (2015, March 10). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Lambeau_Field&oldid=650772069
- TCF Bank Stadium. (2015, March 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=TCF_Bank_Stadium&oldid=649686714
- Georgia Dome. (2015, March 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Georgia_Dome&oldid=653308324
- Bank of America Stadium. (2015, March 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Bank_of_America_Stadium&oldid=6533742 05

- Mercedes-Benz Superdome. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:48, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Mercedes-Benz_Superdome&oldid=654152346
- Raymond James Stadium. (2015, February 28). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:49, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Raymond_James_Stadium&oldid=64920243 5

• University of Phoenix Stadium. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:49, April 8, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=University_of_Phoenix_Stadium&oldid=653 559579

• Edward Jones Dome. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:49, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=Edward_Jones_Dome&oldid=654167126 • Levi's Stadium. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:49, April 8,

- Levi's Stadium. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Refrieved 13:49, April 8, 2015, from http://en.wikipedia.org/w/index.php?title=Levi%27s_Stadium&oldid=655210318
- CenturyLink Field. (2015, March 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 13:49, April 8, 2015,

from http://en.wikipedia.org/w/index.php?title=CenturyLink_Field&oldid=652787982

- Toyota Park. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:30, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Toyota_Park&oldid=654161546
- Mapfre Stadium. (2015, March 19). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Mapfre_Stadium&oldid=652025679
- Robert F. Kennedy Memorial Stadium. (2015, March 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=Robert_F._Kennedy_Memorial_Stadium&ol did=652754664

- Saputo Stadium. (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Saputo_Stadium&oldid=653113798
- Gillette Stadium. (2015, April 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Gillette_Stadium&oldid=654585813
- Yankee Stadium. (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Yankee_Stadium&oldid=655513009
- Red Bull Arena (New Jersey). (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Red_Bull_Arena_(New_Jersey)&oldid=6531 49464

• Orlando Citrus Bowl Stadium. (2015, March 21). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Orlando_Citrus_Bowl_Stadium&oldid=652 829636

- PPL Park. (2015, March 18). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=PPL_Park&oldid=651972956
- BMO Field. (2015, March 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=BMO_Field&oldid=653369346
- Dick's Sporting Goods Park. (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,

 $from \ http://en.wikipedia.org/w/index.php?title=Dick\%27s_Sporting_Goods_Park\&oldid=65~5566313$

• Toyota Stadium (Texas). (2014, December 16). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,

 $from \ http://en.wikipedia.org/w/index.php?title=Toyota_Stadium_(Texas)\&oldid=638418651$

- BBVA Compass Stadium. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=BBVA_Compass_Stadium&oldid=65523563 6
- StubHub Center. (2015, March 12). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:31, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=StubHub_Center&oldid=651093792
- Providence Park. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Providence_Park&oldid=654711739
- Rio Tinto Stadium. (2015, March 31). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Rio_Tinto_Stadium&oldid=654304261 • Avaya Stadium. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9,

- 2015, from http://en.wikipedia.org/w/index.php?title=Avaya_Stadium&oldid=655230355
- CenturyLink Field. (2015, March 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=CenturyLink_Field&oldid=652787982

• Sporting Park. (2015, March 13). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Sporting_Park&oldid=651202621

- BC Place. (2015, March 26). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=BC_Place&oldid=653675532
- Harvard Stadium. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Harvard_Stadium&oldid=654508027
- Village of Lisle-Benedictine University Sports Complex. (2014, September 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:32, April 9, 2015, fromhttp://en.wikipedia.org/w/index.php?title=Village_of_Lisle-
- Benedictine_University_Sports_Complex&oldid=624371085
- BBVA Compass Stadium. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=BBVA_Compass_Stadium&oldid=65523563
- Providence Park. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Providence_Park&oldid=654711739
- Memorial Stadium (Seattle). (2015, March 22). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Memorial_Stadium_(Seattle)&oldid=652979 450

- Yurcak Field. (2014, October 11). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Yurcak_Field&oldid=629218625
- Maryland SoccerPlex. (2014, October 17). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Maryland_SoccerPlex&oldid=629989133

• Sahlen's Stadium. (2015, March 31). In Wikipedia, The Free Encyclopedia. Retrieved 16:33, April 9, 2015,

 $from \ http://en.wikipedia.org/w/index.php?title=Sahlen\%27s_Stadium\&oldid=654403301$

• Atlanta Silverbacks Park. (2014, November 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

 $from\ http://en.wikipedia.org/w/index.php?title=Atlanta_Silverbacks_Park\&oldid=632159843$

• WakeMed Soccer Park. (2015, February 18). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

 $from\ http://en.wikipedia.org/w/index.php?title=WakeMed_Soccer_Park\&oldid=647633122$

- Clarke Stadium. (2015, April 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Clarke_Stadium&oldid=654678732
- Lockhart Stadium. (2014, November 4). In Wikipedia, The Free Encyclopedia. Retrieved 16:33, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Lockhart_Stadium&oldid=632454890

• Baseball Grounds of Jacksonville. (2015, March 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

 $from http://en.wikipedia.org/w/index.php?title=Baseball_Grounds_of_Jacksonville&oldid=6~49550076$

• IU Michael A. Carroll Track & Soccer Stadium. (2014, December 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:33, April 9, 2015,

fromhttp://en.wikipedia.org/w/index.php?title=IU_Michael_A._Carroll_Track_%26_Soccer_Stadium&oldid=636779965

- National Sports Center. (2014, December 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=National_Sports_Center&oldid=636289847
- James M. Shuart Stadium. (2015, January 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=James_M._Shuart_Stadium&oldid=6406885 11
- Toyota Field. (2015, April 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Toyota_Field&oldid=654448149
- Al Lang Stadium. (2015, March 20). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Al_Lang_Stadium&oldid=652776072
- Blackbaud Stadium. (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- $from \ http://en.wikipedia.org/w/index.php?title=Blackbaud_Stadium\&oldid=653227162$
- American Legion Memorial Stadium. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- fromhttp://en.wikipedia.org/w/index.php?title=American_Legion_Memorial_Stadium&oldi d=654180309
- Skyline Sports Complex. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Skyline_Sports_Complex&oldid=654929230 • Louisville Slugger Field. (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Louisville_Slugger_Field&oldid=655428425 •Saputo Stadium. (2015, March 23). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9,
- 2015, from http://en.wikipedia.org/w/index.php?title=Saputo_Stadium&oldid=653113798 •Highmark Stadium. (2015, April 4). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9,
- 2015, from http://en.wikipedia.org/w/index.php?title=Highmark_Stadium&oldid=654950422 •City Stadium (Richmond). (2015, February 19). In *Wikipedia, The Free Encyclopedia*. Retrieved
 - 16:34, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=City_Stadium_(Richmond)&oldid=64780196
- •Sahlen's Stadium. (2015, March 31). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Sahlen%27s_Stadium&oldid=654403301
- St. Louis Soccer Park. (2015, April 3). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:34, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=St._Louis_Soccer_Park&oldid=654795723
- Ontario Soccer Centre. (2015, March 2). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:35, April 9, 2015,
- $from \ http://en.wikipedia.org/w/index.php?title=Ontario_Soccer_Centre\&oldid=649486610$
- Legion Stadium. (2014, December 16). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:35, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Legion_Stadium&oldid=638361469
- •Scottsdale Stadium. (2015, March 19). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Scottsdale_Stadium&oldid=652123780

- House Park. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=House_Park&oldid=654152211
- Sand Creek Stadium. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Sand_Creek_Stadium&oldid=654238924
- StubHub Center. (2015, March 12). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=StubHub_Center&oldid=651093792
- Bishop McGuinness Catholic High School (Oklahoma). (2015, April 8). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015,
- fromhttp://en.wikipedia.org/w/index.php?title=Bishop_McGuinness_Catholic_High_School _(Oklahoma)&oldid=655587468
- Anteater Stadium. (2014, December 16). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Anteater_Stadium&oldid=638362970
- Merlo Field. (2015, March 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Merlo_Field&oldid=654210299
- Rio Tinto Stadium. (2015, March 31). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015,
- from http://en.wikipedia.org/w/index.php?title=Rio_Tinto_Stadium&oldid=654304261
- Bonney Field. (2015, April 6). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Bonney_Field&oldid=655241105
- Starfire Sports. (2015, March 22). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=Starfire_Sports&oldid=653021031
- ONEOK Field. (2015, January 30). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:38, April 9, 2015, from http://en.wikipedia.org/w/index.php?title=ONEOK_Field&oldid=644849140
- Thunderbird Stadium. (2015, April 7). In *Wikipedia, The Free Encyclopedia*. Retrieved 16:39, April 9, 2015,

from http://en.wikipedia.org/w/index.php?title=Thunderbird_Stadium&oldid=655408047

Expensive Beauty Products

- Created by Author

Applications Wanted

- All Recorded Traffic Tickets:
- Data provided by the Baltimore Department of Finance on Open Baltimore (https://data.baltimorecity.gov/) Direct Link: https://data.baltimorecity.gov/Transportation/All-Time/ks7u-tsjz (accessed on

Where's the Joe?

- Mermaid Coffee Company: Data modified to exclude "Starbucks" from everywhere in the data.
- Open Data https://opendata.socrata.com/Business/All-Starbucks-Locations-in-the-World/xy4y-c4mk provided by Starbucks Corporation

Meta-morphosis

- Created by Author

NYC

- Landmarks:
- o Created By Author
- NYC Subway Entrances:
- NYC Open Data https://data.cityofnewyork.us/Transportation/Subway-Entrances/drexxx56 provided by Metropolitan Transportation Authority (MTA)
- NYC_Free_Public_WiFi_12052014:
- NYC Open Data https://data.cityofnewyork.us/Social-Services/NYC-Wi-Fi-Hotspot-Locations/a9we-mtpn provided by Department of Information Technology & Telecommunications (DoITT)

Content:

The Alteryx Help documentation was used extensively in the research for this manual. Alteryx Help can be accessed through the Help Menu, Properties Window, or with the Hot Key F1 in Alteryx or can be accessed directly at http://help.alteryx.com/.

USEREADY

USEReady is a leading provider of advanced business intelligence, data management, and network security solutions. Our mission is to help businesses succeed by fast-tracking their business performance through our industry, domain and technology expertise, agile global delivery practices, and our customer fanaticism. For more information, visit: www.useready.com.



- Laura Sellers. Vice President. Product Management, Alteryx



Authored By

Abhilash Ramanathan

Rahul Upadhye

Anand Vadul

Ashish Tergaonkar

Jagdeesh Narayanan

Rima Upadhye

Vijai Narasimha

Vijay Gowtham

Published by



www.useready.com contact@useready.com

©2016 USEReady, All Rights reserved.