

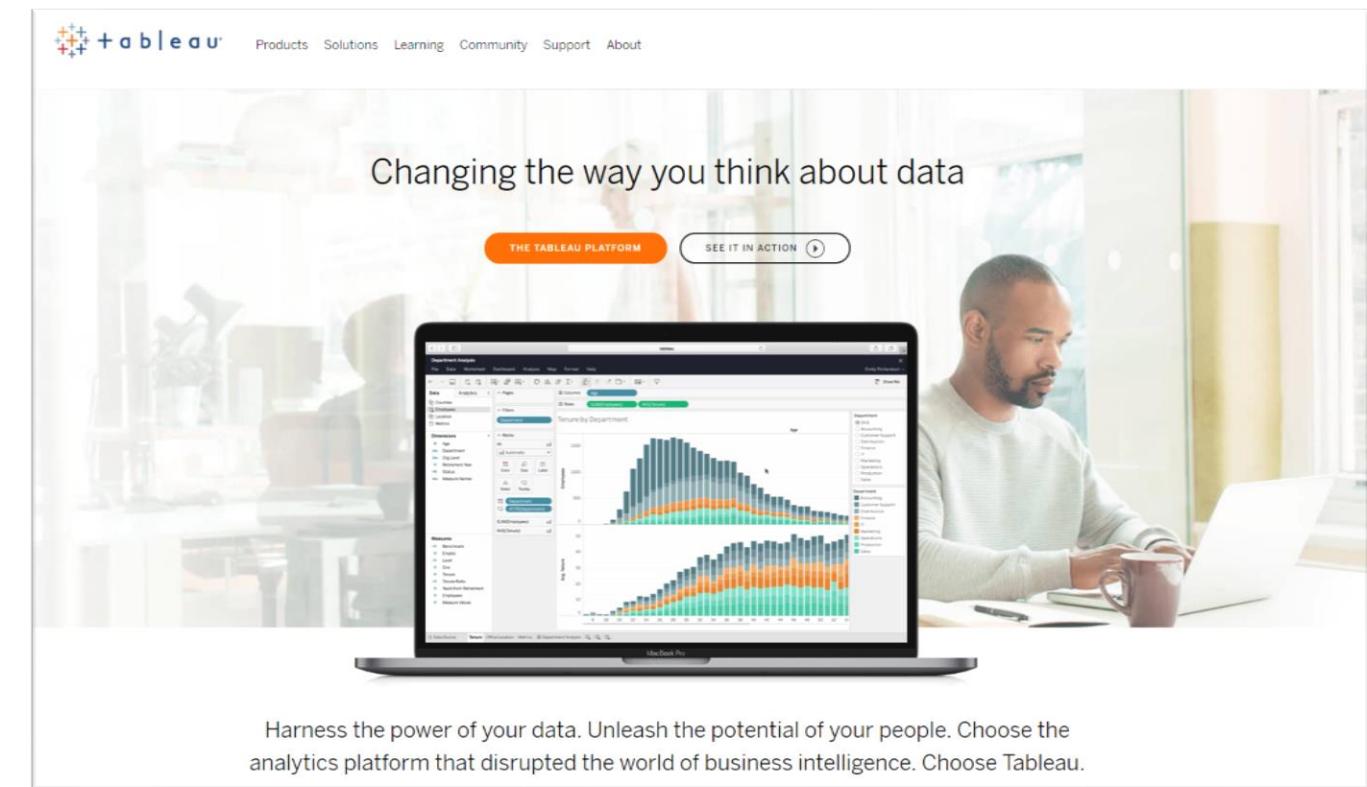
An Introduction to Analysis and Data Visualization using Tableau Software

Presentation Overview

- 01 What is Tableau Software?
- 02 Benefits for Teachers & Researchers
- 03 What is Data Visualization?
- 04 General Overview of Tableau
- 05 Use for Reporting - Examples
- 06 Use for Storytelling - Examples
- 07 Use for Analysis - Examples
- 08 Advanced Features - Example
- 09 Resources (Public, WMTUG, Books)

What is Tableau Software?

- Software company Founded in 2003 from Stanford research
- Intent is to bring 'data to the people' through easy to use data visualization software
- Would be classified as a hybrid business intelligence (BI) / analytics software company
- Used by many of the largest companies in the world and most large companies in West Michigan



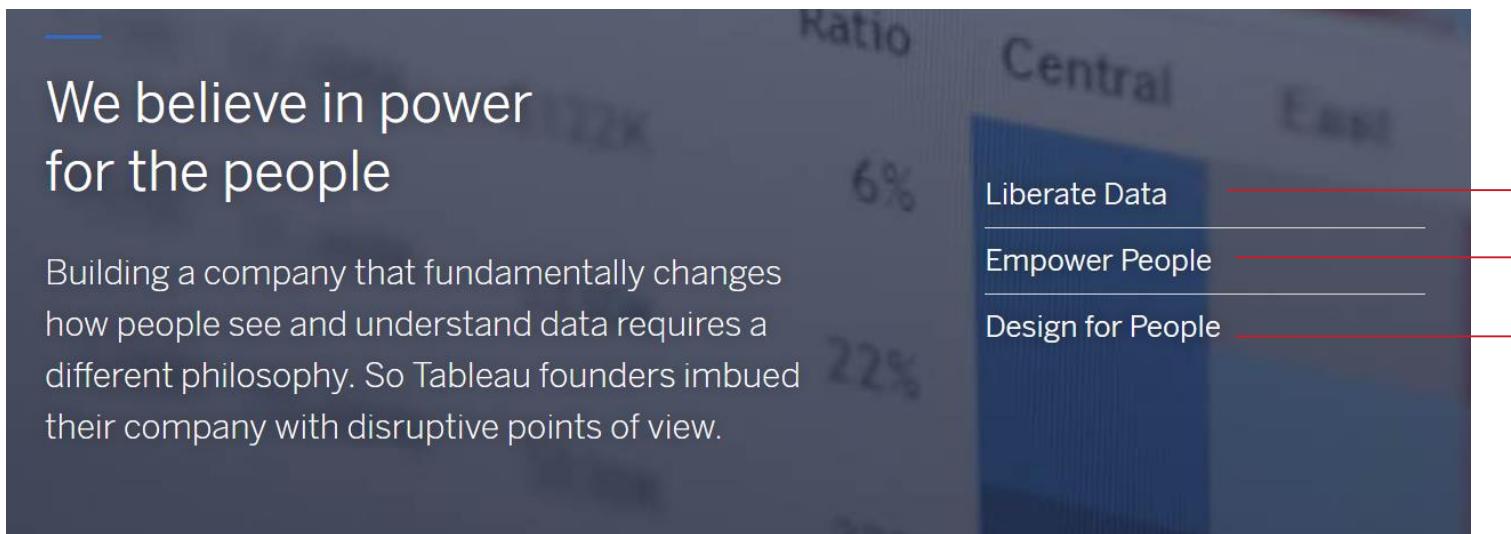
What is Tableau Software?

- Similar tools to Tableau include Microsoft Power BI, Qlik, Tibco Spotfire, and Looker – these are all data visualization tools

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms



What is Tableau Software?



Courtesy: www.Tableau.com

The main focus of Tableau software is for you to better understand your datasets, especially large datasets.

BI software in the past required highly technical IT skills and took a long time to build dashboards. Tableau has changed that paradigm.

Tableau invests a lot of research time into developing intuitive software. They approach software design from the human perspective.

Benefits for Researchers & Teachers

- Free course licenses for students
- Pre-built curriculum for teaching Tableau and data analysis
- Use of powerful 'big' data platform for large datasets
- Provides skills needed in industry (various professions)



Why Tableau?

In-demand skills

Data isn't going anywhere. Companies are hiring for analytical skills to tackle big data in every industry.

Student engagement

Spend less time teaching software and more time helping your students find deep analytical insights.

Easy to use

With a large library of on-demand tutorials and a dedicated support community, teaching data analysis is easier than ever.



Check out the instructor resource page for self-service learning options, ready-made curriculum materials, answers to frequently asked questions, and license support.

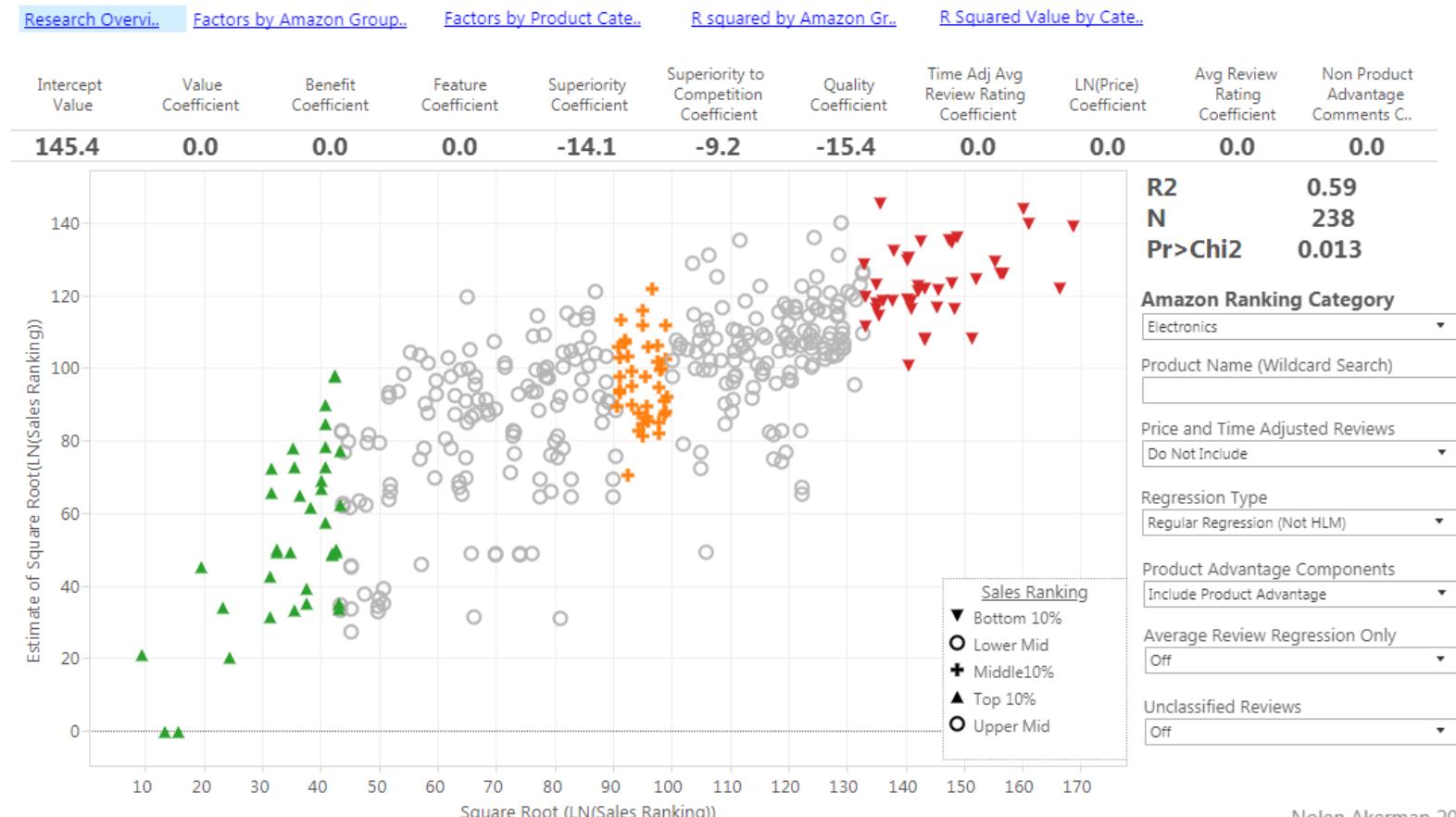
CHECK IT OUT →

<https://www.tableau.com/academic/teaching>

Benefits for Researchers

- Ability to handle ‘big’ data (hundreds of millions of rows) that Excel cannot
- Ability to share (link) your research articles to datasets and results through Tableau Public
- Access to online help forums & local users groups
- Ability to connect to “R” and Python for more advanced analytics and analysis

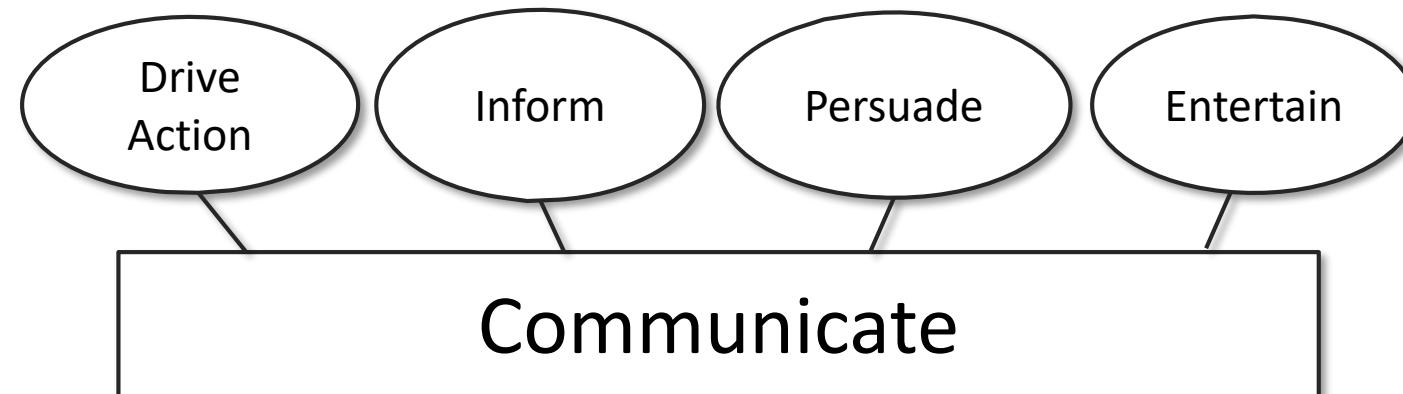
Modeling New Product Success from Component Measures of Product Advantage: *A Model Utilizing Automated Text Classification and Sentiment Analysis*



What is Data Visualization?

What is Data Visualization

What is the Purpose of Data Visualizations?



What guides the design process?
How do we judge success?

What is Data Visualization?

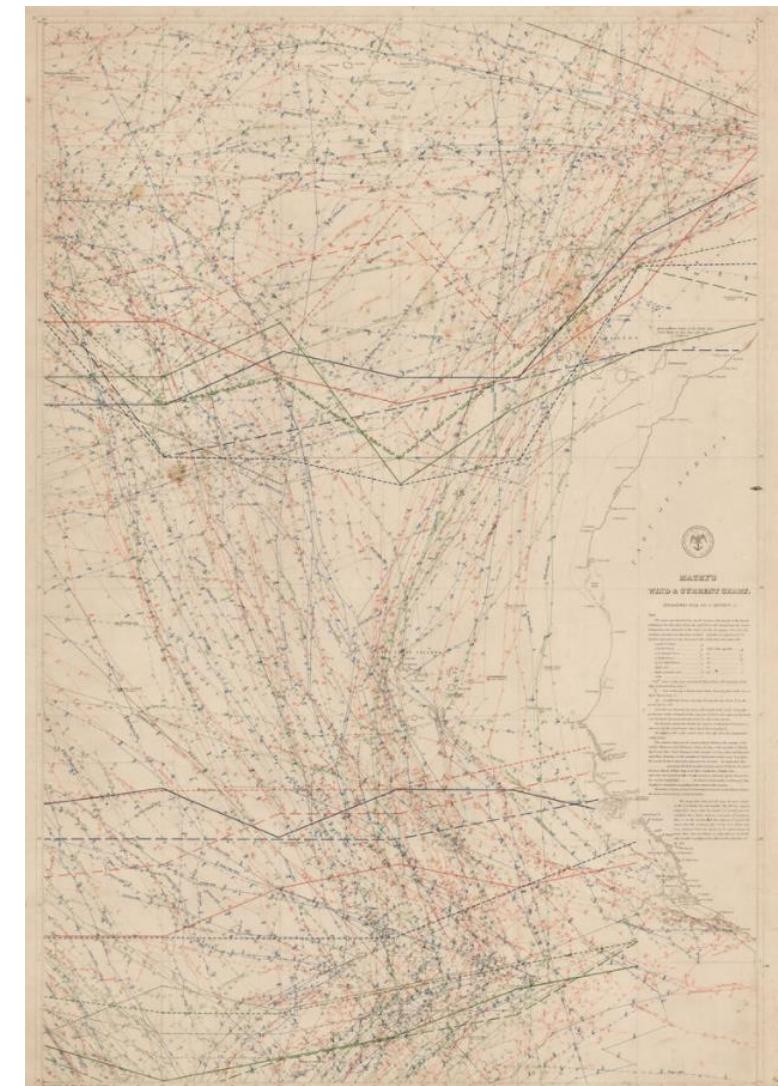


[https://en.wikipedia.org/wiki/
Matthew_Fontaine_Maury](https://en.wikipedia.org/wiki/Matthew_Fontaine_Maury)

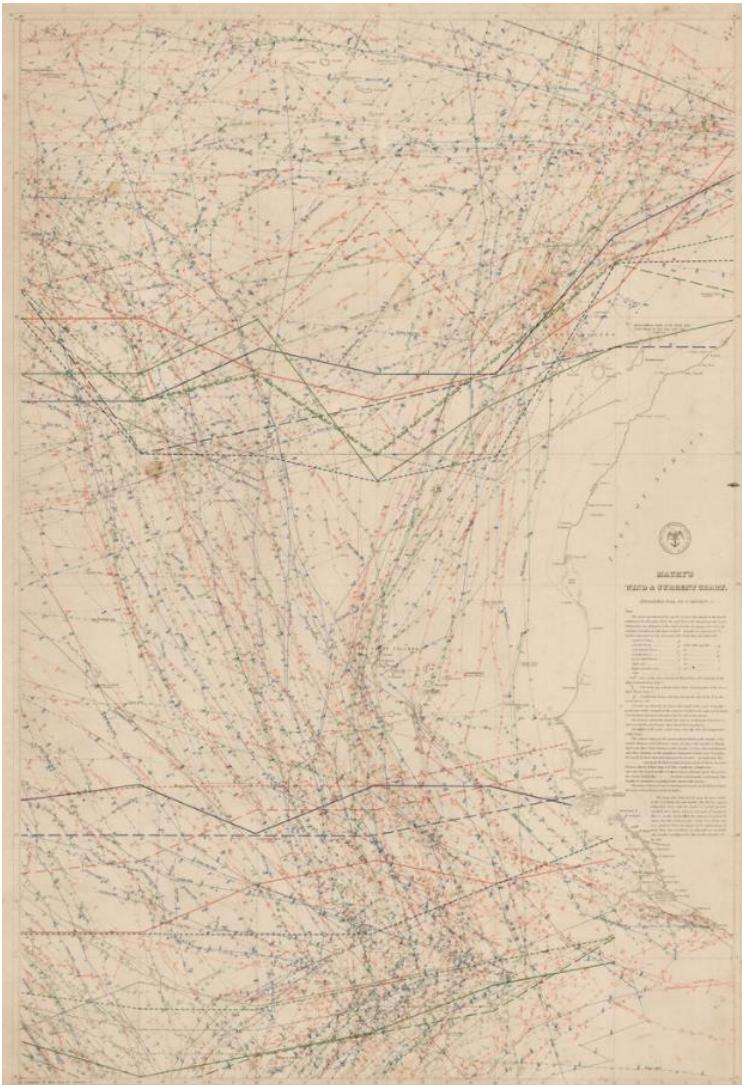
Matthew Fontaine Maury

- Unfit for duty due to a leg injury
- Sent to Depot of Charts and Instruments
- Vault of logs from every ship in US Navy
- Hundreds of thousands of observations available in written logs
- Manual 'data mining' with his team
- Standardized collection moving forward (form)

Ref. (The Clipper Ships – Time Life Books)
Ref. (Wind & Current Charts -1847)



What is Data Visualization?



Wind & Current Charts - 1847

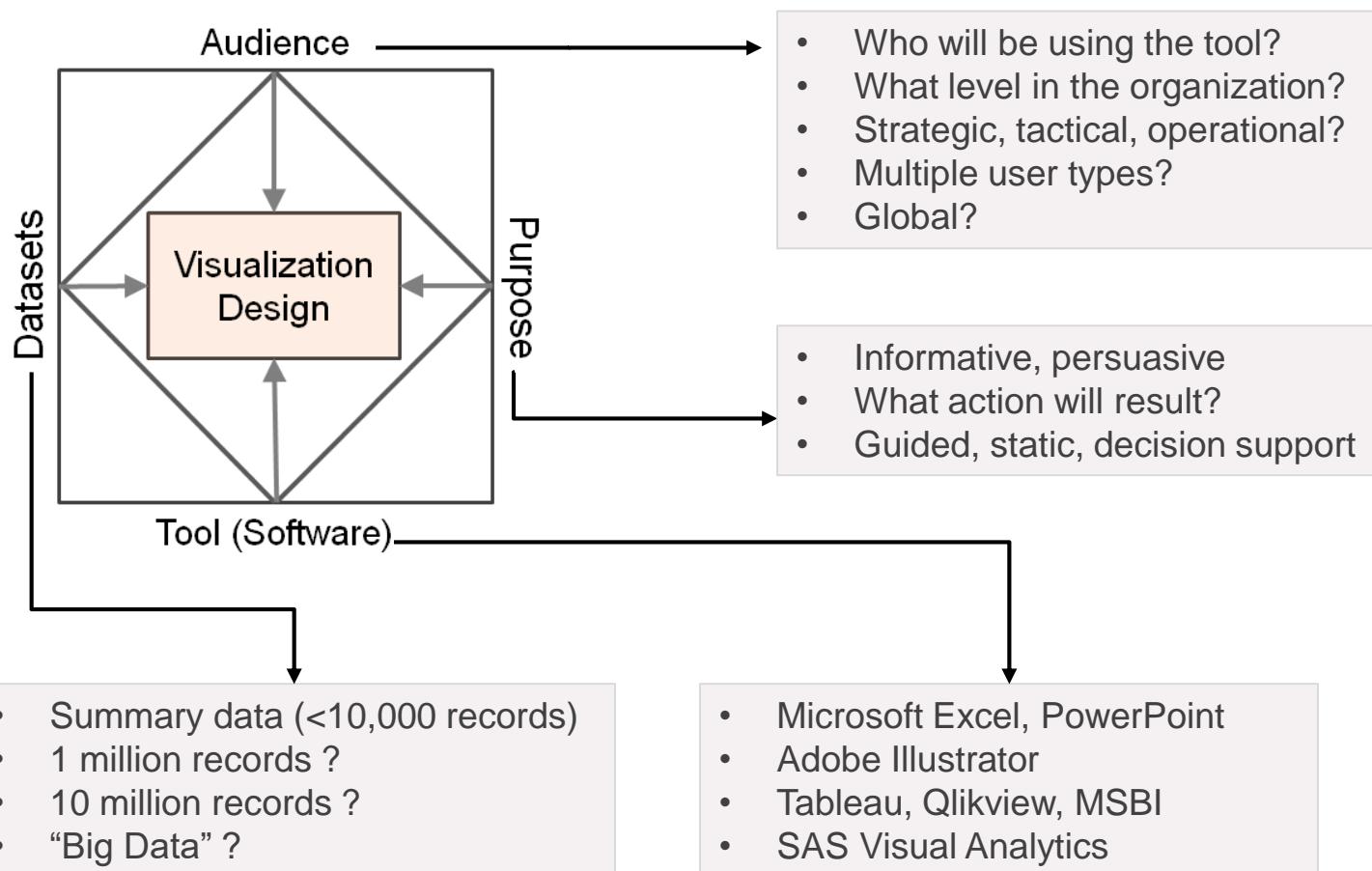
- Visualization of his team's findings
- Use of symbols and colors to highlight best routes
- Findings were counter-intuitive (heading west to go faster east)

Results

- Roundtrip from Virginia to Rio 75 days instead of 112 days
- Found the Gulf Stream's full shape
- Cut time from Cape Horn to California by a third
- Reduced ship lost due to storms

What is Data Visualization

A Basic Framework – Rhetoric for Data Visualization



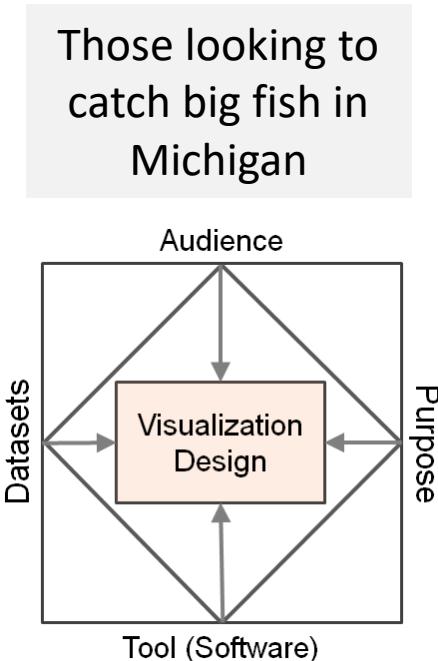
Methodology

1. Identify Purpose (Intended Use)
2. Consider Audience
3. Research
 - i. Identify Available Datasets
 - ii. Identify Data Elements
 - iii. Benchmark Designs
4. Design
 - i. Sketch
 - ii. Iterate
 - iii. Collect Feedback
5. Execute Design
 - i. Collect Feedback
6. Document – Deploy
7. Sustain

What is Data Visualization

Example – Decision Support

Michigan DNR
Database;
Public Use
Pictures



Tableau

Provide decision support to increase chances of catching big fish

Fishing Navigator Data References

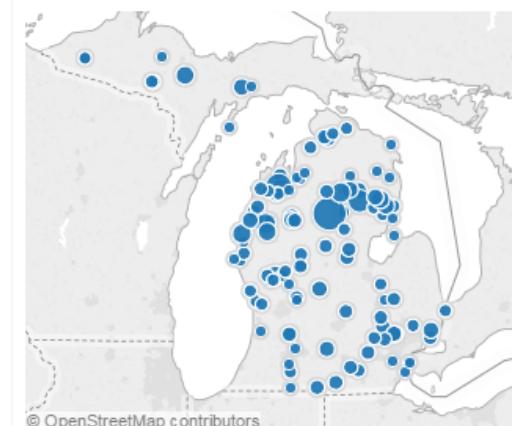
Fishin' in the Mitten: Catching Big Fish in Michigan

Michigan has a diverse selection of gamefish for fishing adventures year round. Select a species from the drop down menu to find out helpful hints for catching the big ones!

SELECT A FISH SPECIES

BLUEGILL

Michigan State Record(s)	Year	Waterbody	County	Inches	Pounds
	1983	Vaughn Lake	Alcona	13.75	2.75



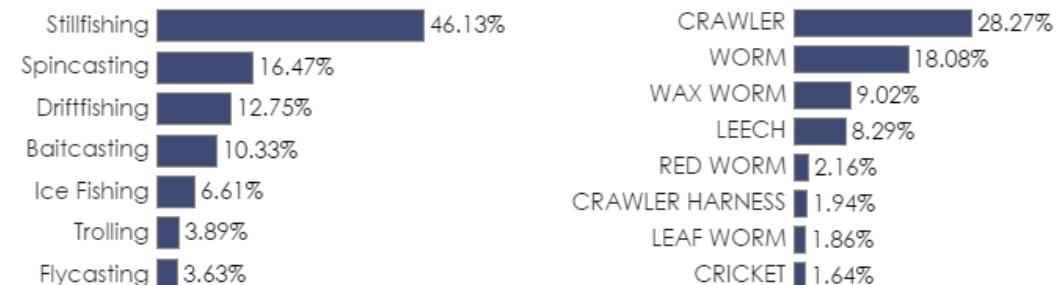
WHERE TO FIND BIG BLUEGILL



For big bluegill, look no further than Houghton Lake. Use a leech and be prepared to become a Master Angler!

Based on Master Angler records from 1993-2014, the map to the left shows the 'hot spots' for big **BLUEGILL**. The size of the circle represents the total Master Angler award level fish caught in that location. Zoom in, or mouse over a circle for additional information about that location. Click on a circle to filter the results below for just that body of water.

TOP FISHING METHODS AND BAITS FOR BIG BLUEGILLS

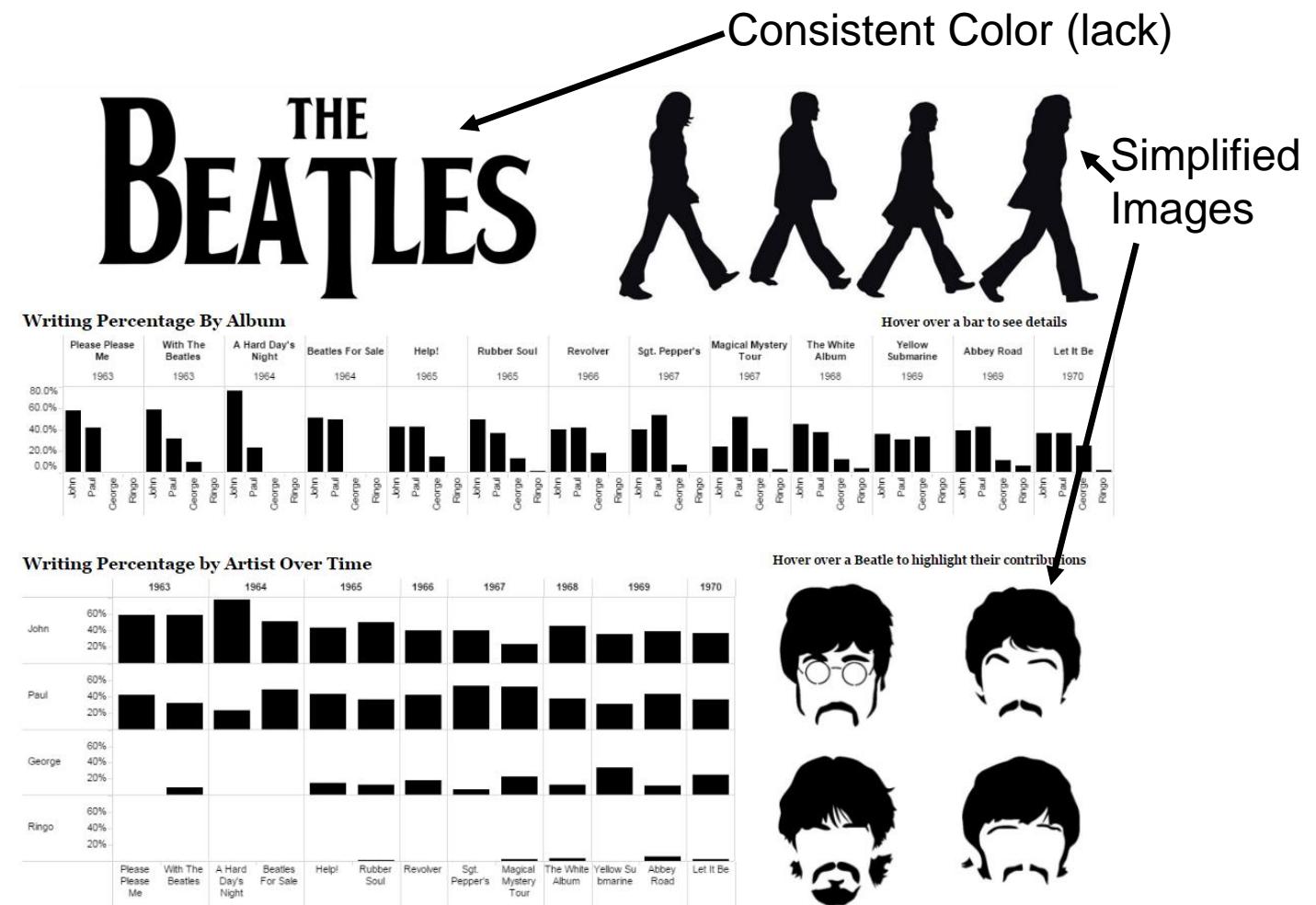


What is Data Visualization

Elements of Design - Unity

Unity is the application of methods that ensure that elements in the design appear to ‘go together’ - (color, font, & shape consistency)

Consistent
Font



<https://public.tableau.com/s/gallery/beatles-albums>

Author: Mike Moore

What is Data Visualization

Elements of Design - Hierarchy

Hierarchy is the application of design methods to indicate importance and ‘flow’ within the visual (size, placement)

Level 1

Level 2

Level 3

Level 4

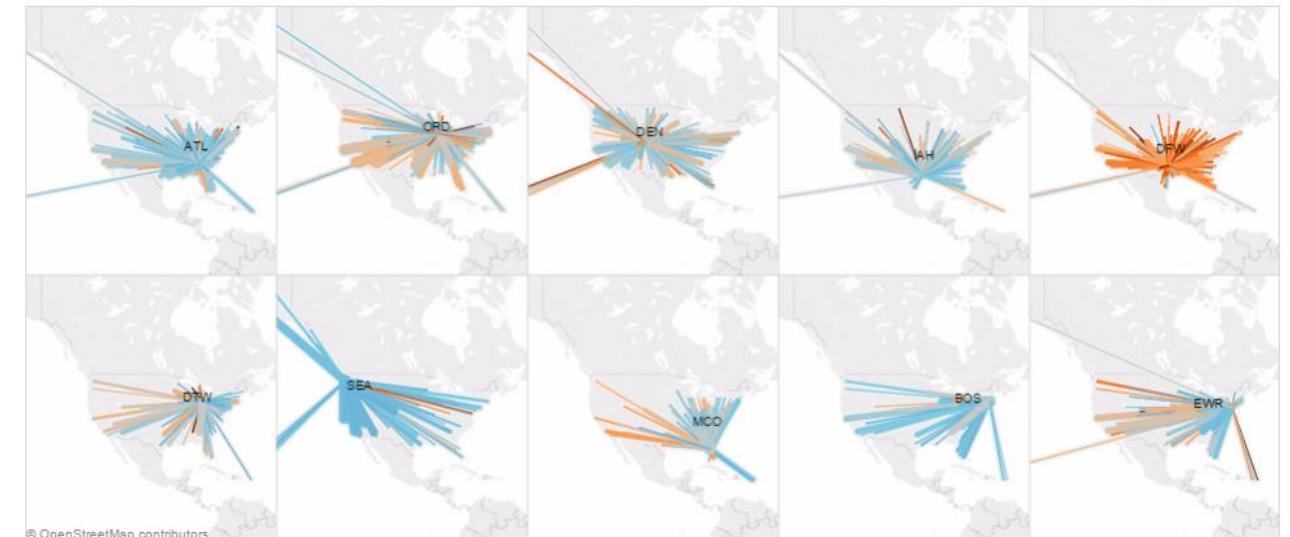
BLAME THE WEATHER

How does precipitation affect the Top 10 most traveled airports?

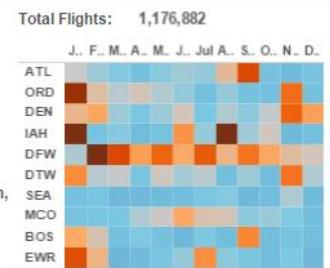


Weather Delay Minutes per Inch of Precipitation (WDMIP) is a metric used to see how precipitation impacts weather delays. Using this metric, Z scores can be calculated to see which flights and airports are more affected by precipitation. The higher the Z Score the more the flight or airport is affected by precipitation.

Use the heatmap on the right to control the dashboard. Click the airport row, month column, or individual cells to compare the affects of precipitation on weather delays.

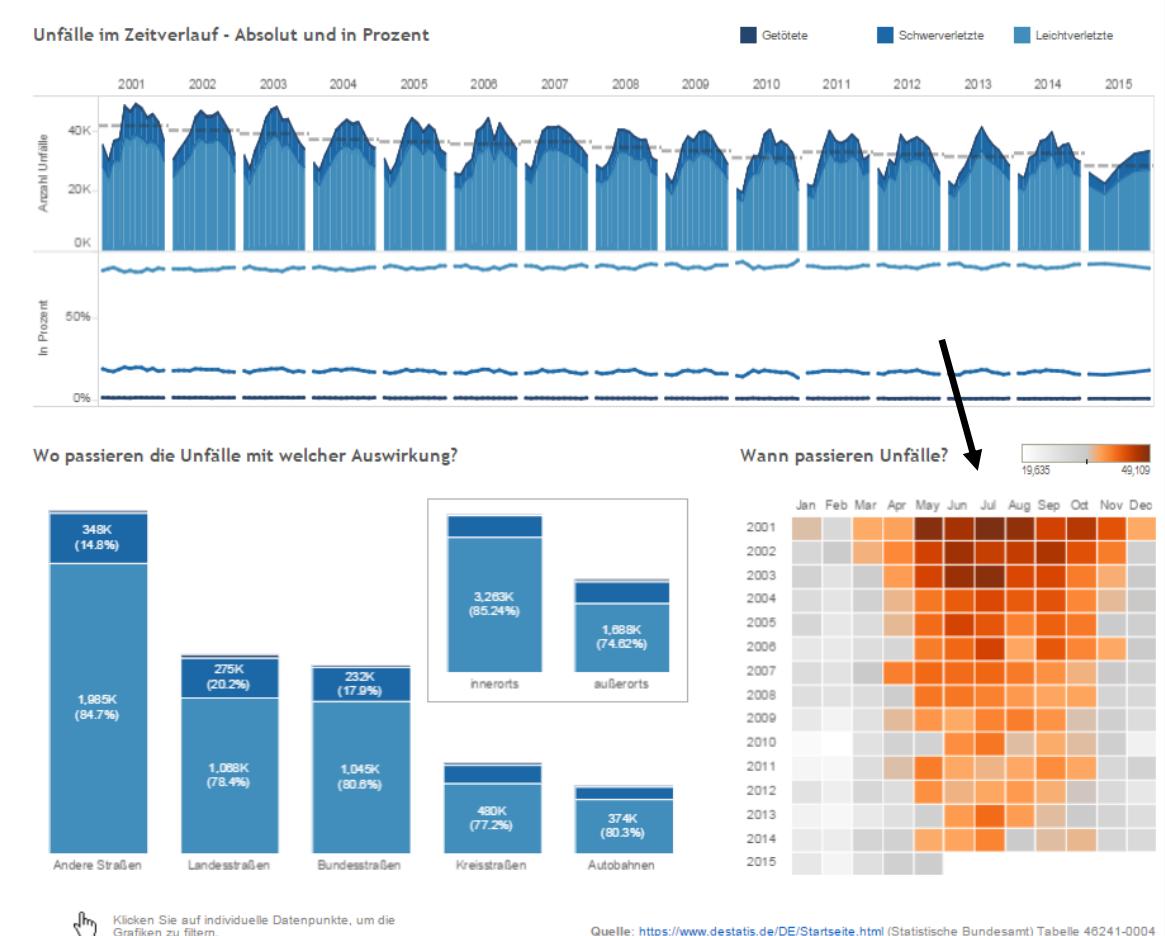


<https://public.tableau.com/s/gallery/blame-weather-us-flight-delayed-precipitation> Author: Matt Chambers



Elements of Design - Color

Use of color provides contrast for data points in opposition and brings attention to relevant elements within the visual.



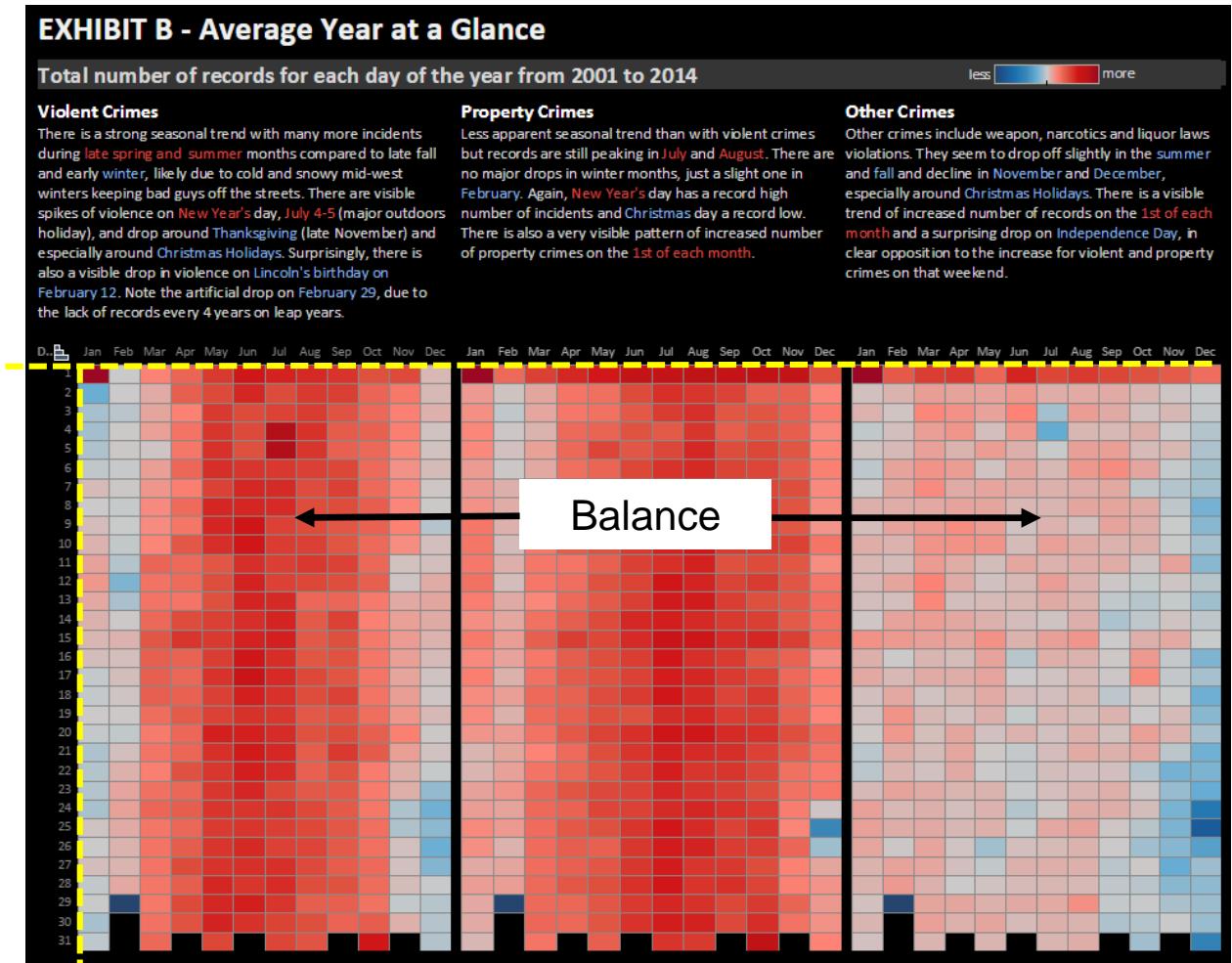
<https://public.tableau.com/s/gallery/road-accidents-germany>
Author: [Oliver Linder](#)

What is Data Visualization

Elements of Design – Balance & Alignment

Balance and alignment are used to create harmonious visuals that do not distract from the message being communicated.

Alignment



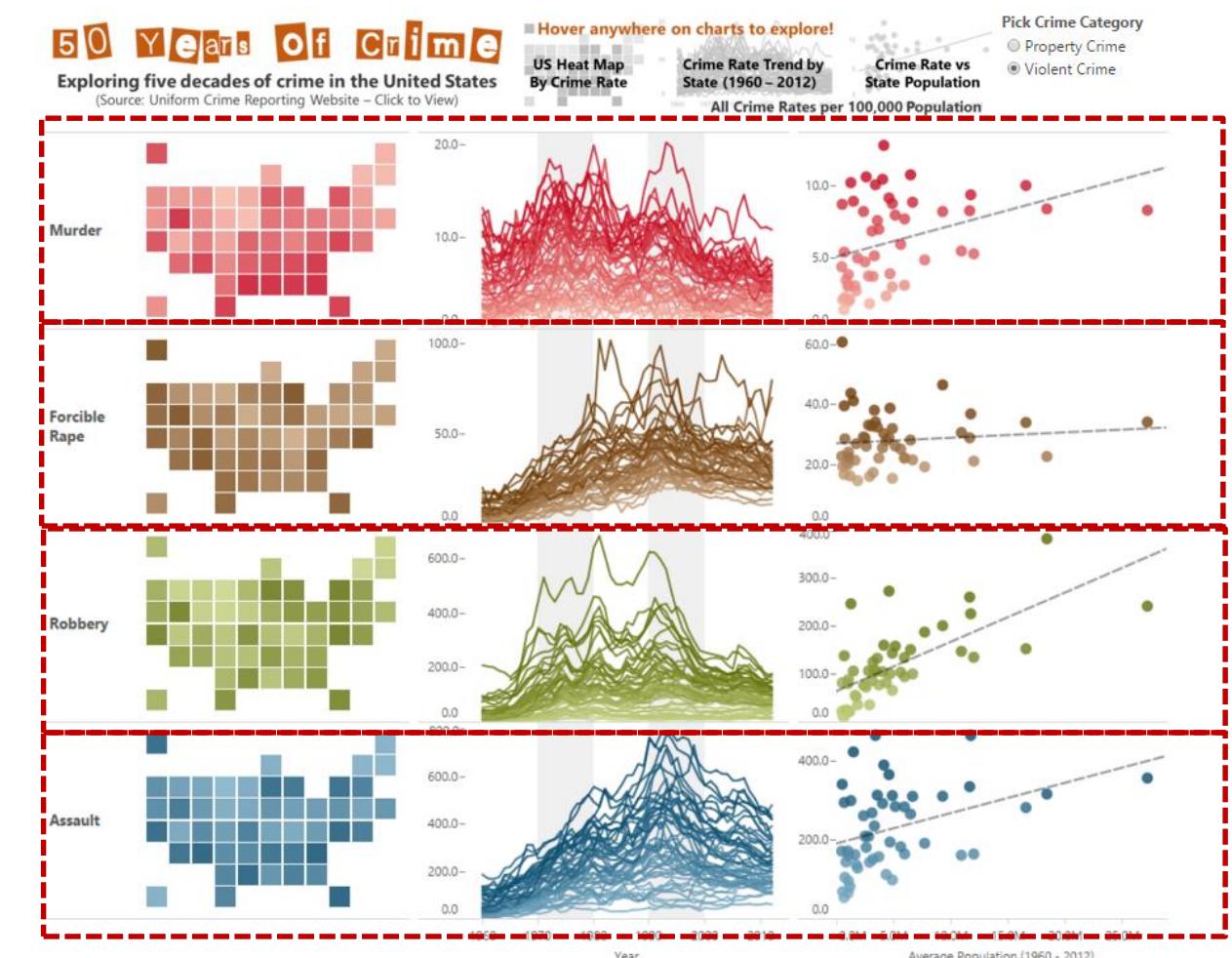
<https://public.tableau.com/s/gallery/chicago-crime-scene>

Author: [George Gorczynski](#)

What is Data Visualization

Elements of Design – Grouping / Spacing

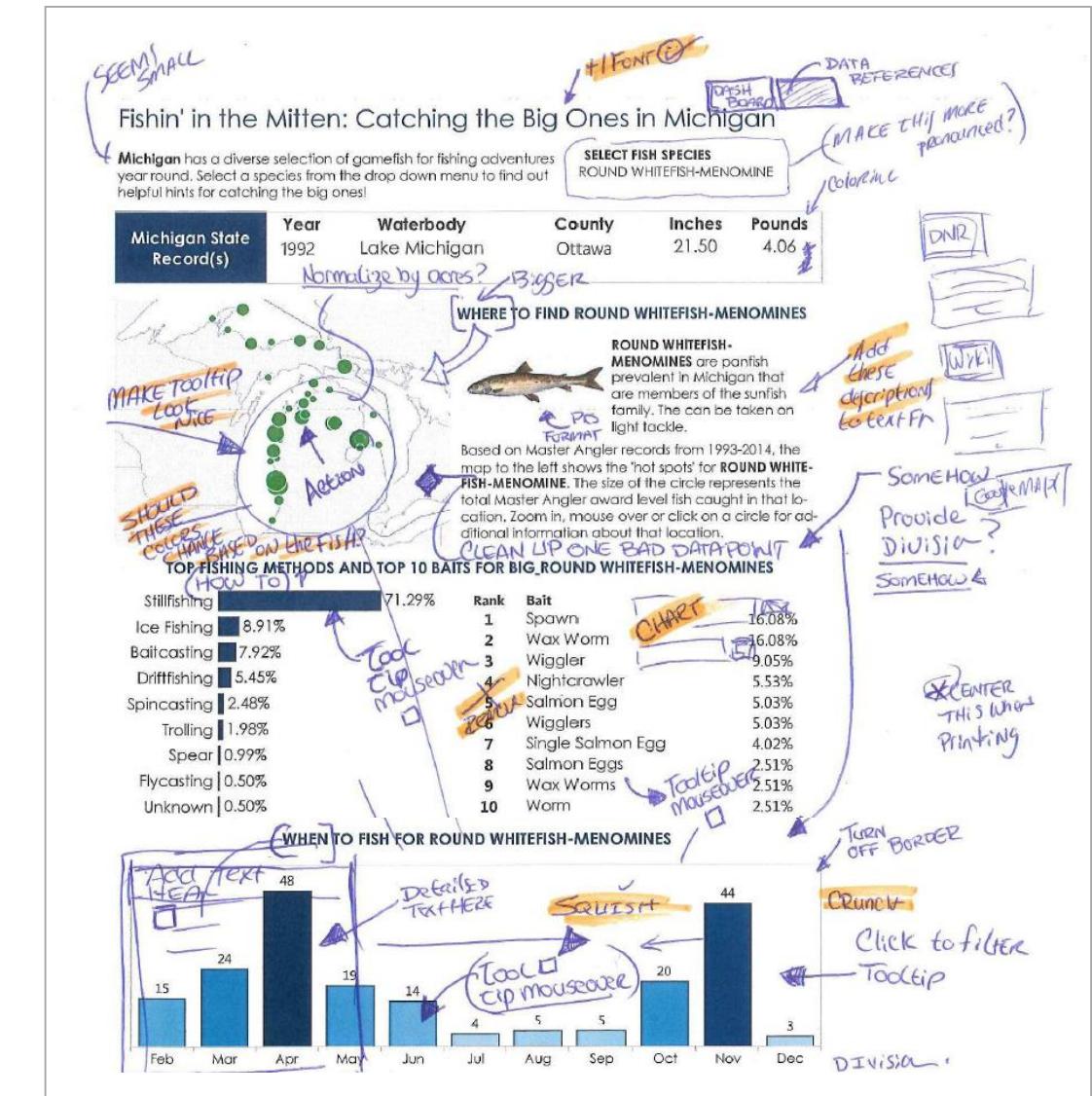
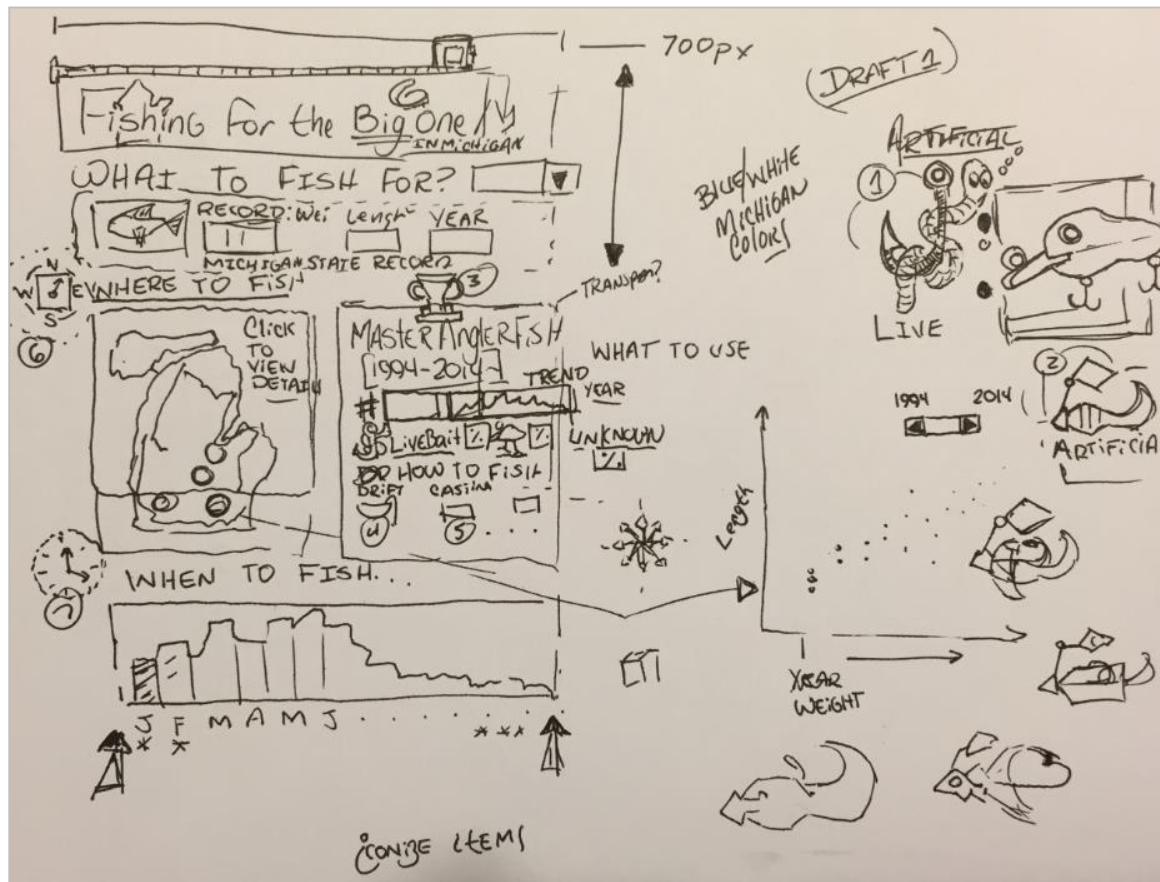
Grouping and spacing can be used to associate similar elements and provide a narrative or visual flow within the visualization.



<https://public.tableau.com/s/gallery/50-years-crime-us> Author: Shine Pulikathara

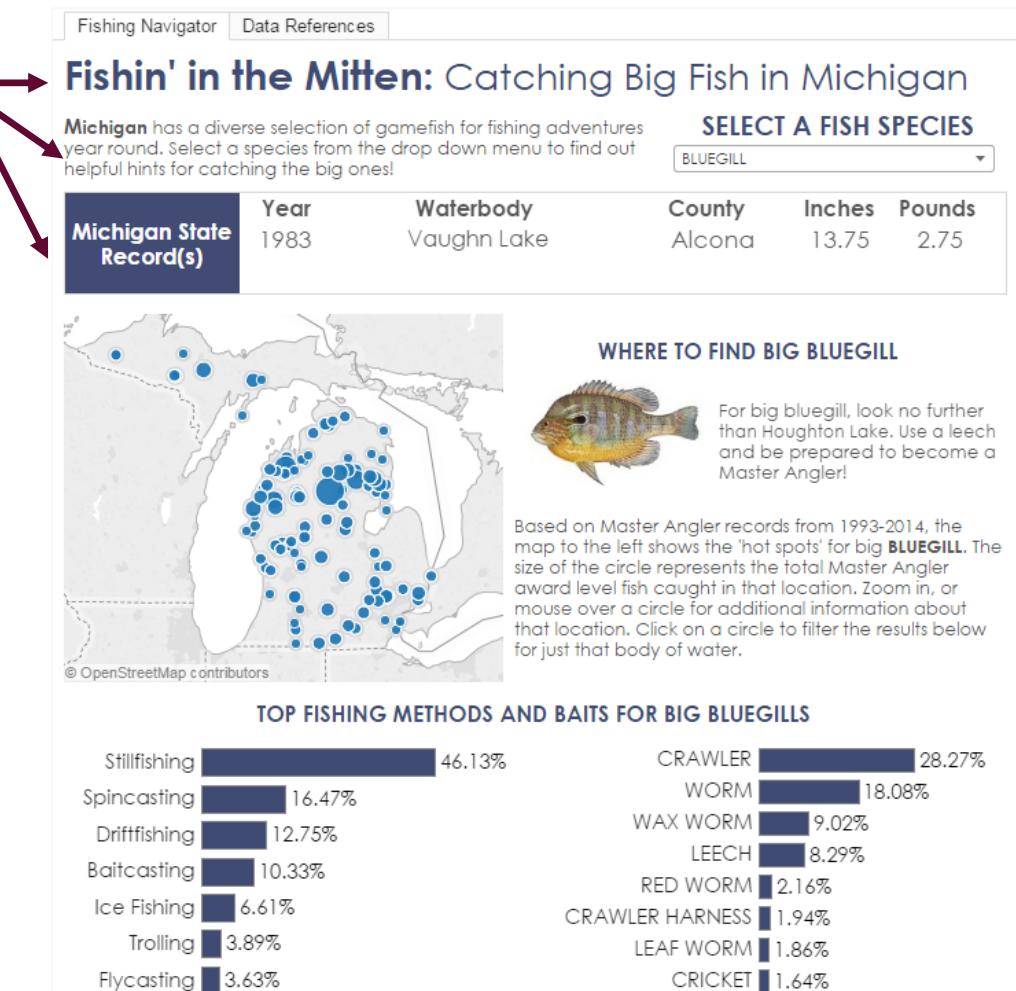
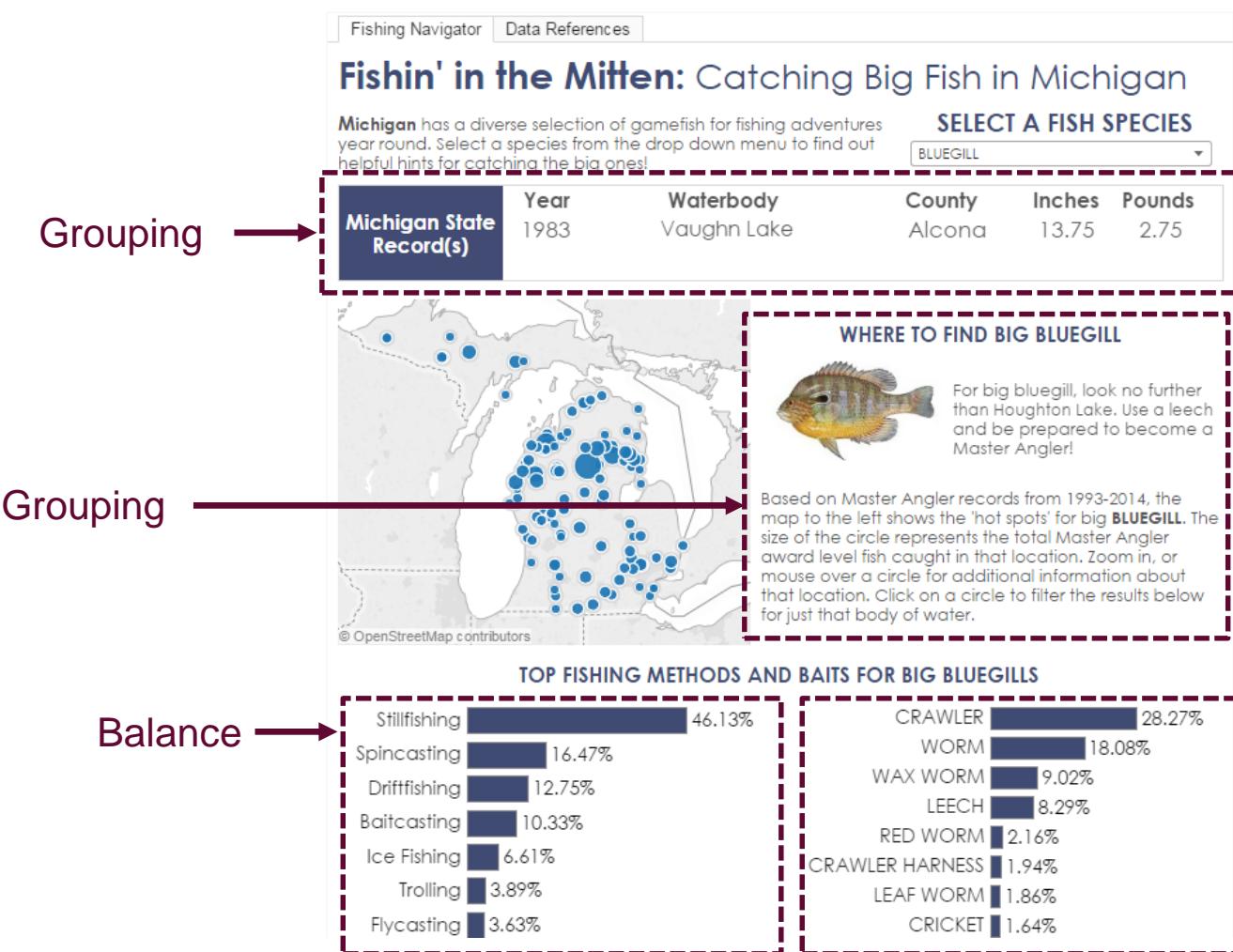
What is Data Visualization

The Iterative Design Process



What is Data Visualization

Detailed Example - Design



Now . . .

Back to Tableau

General Overview

Tableau – General Overview

- All worksheets & dashboards start with data
- Tableau connects to almost every type of data file imaginable
- You can join across different type of data sources!

Files
(Excel, CSV,
JSON, SAS...)

Servers
(Databases)

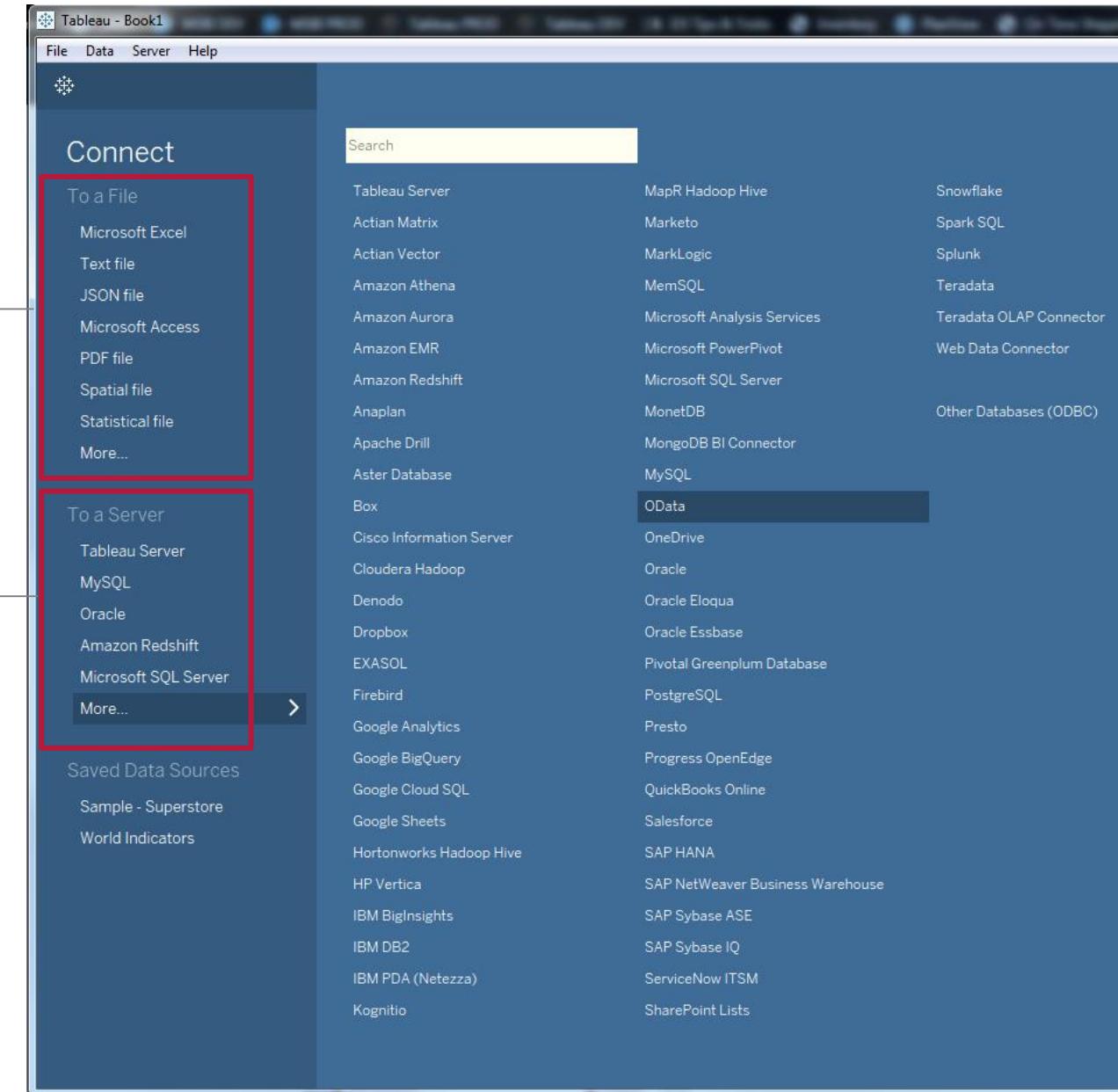


Tableau – General Overview – simple example

- A simple table with 15 rows of data in an Excel spreadsheet
- Build an interactive dashboard in under three minutes

A	B	C	D	E
1	Product Type	State	Zip Code	Sales
2	Chairs	Michigan	49012	\$ 78,847
3	Chairs	Michigan	49008	\$ 42,998
4	Chairs	Michigan	49014	\$ 39,554
5	Chairs	Michigan	49007	\$ 10,973
6	Chairs	Ohio	45891	\$ 9,558
7	Chairs	Ohio	45888	\$ 51,831
8	Chairs	Ohio	45871	\$ 34,972
9	Desks	Michigan	49012	\$ 30,838
10	Desks	Michigan	49008	\$ 71,298
11	Desks	Michigan	49014	\$ 11,558
12	Desks	Michigan	49007	\$ 74,435
13	Desks	Ohio	45891	\$ 52,503
14	Desks	Ohio	45888	\$ 45,530
15	Desks	Ohio	45871	\$ 84,076

Tableau - Book2

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Data Analytics Connect to Data Dimensions Measures

Pages Columns Rows

Marks Automatic Color Size Text Detail Tooltip

Sheet 1

Drop field here

Drop field here

Drop field here

Show Me

Connect to data

25

24

Tableau – General Overview

Calculated Fields

The screenshot shows the Tableau Calculated Field dialog. In the main text area, a calculated field named "zz_" is defined with the following DAX code:

```
if CONTAINS([Product Type], "CHAIR")
    then "SEATING"
    else "NOT SEATING" END
```

The calculated field is valid, as indicated by the message at the bottom of the dialog.

To the right of the main dialog is a floating search interface. It has a dropdown menu set to "All" and a search input field labeled "Enter search text". A list of functions is displayed, with "CONTAINS" highlighted. The details for the "CONTAINS" function are shown on the right side of the interface:

CONTAINS(string, substring)
Returns true if the string contains the substring.
Example:
CONTAINS("Calculation", "alcu") is true

Tableau – General Overview

Basic Analytics

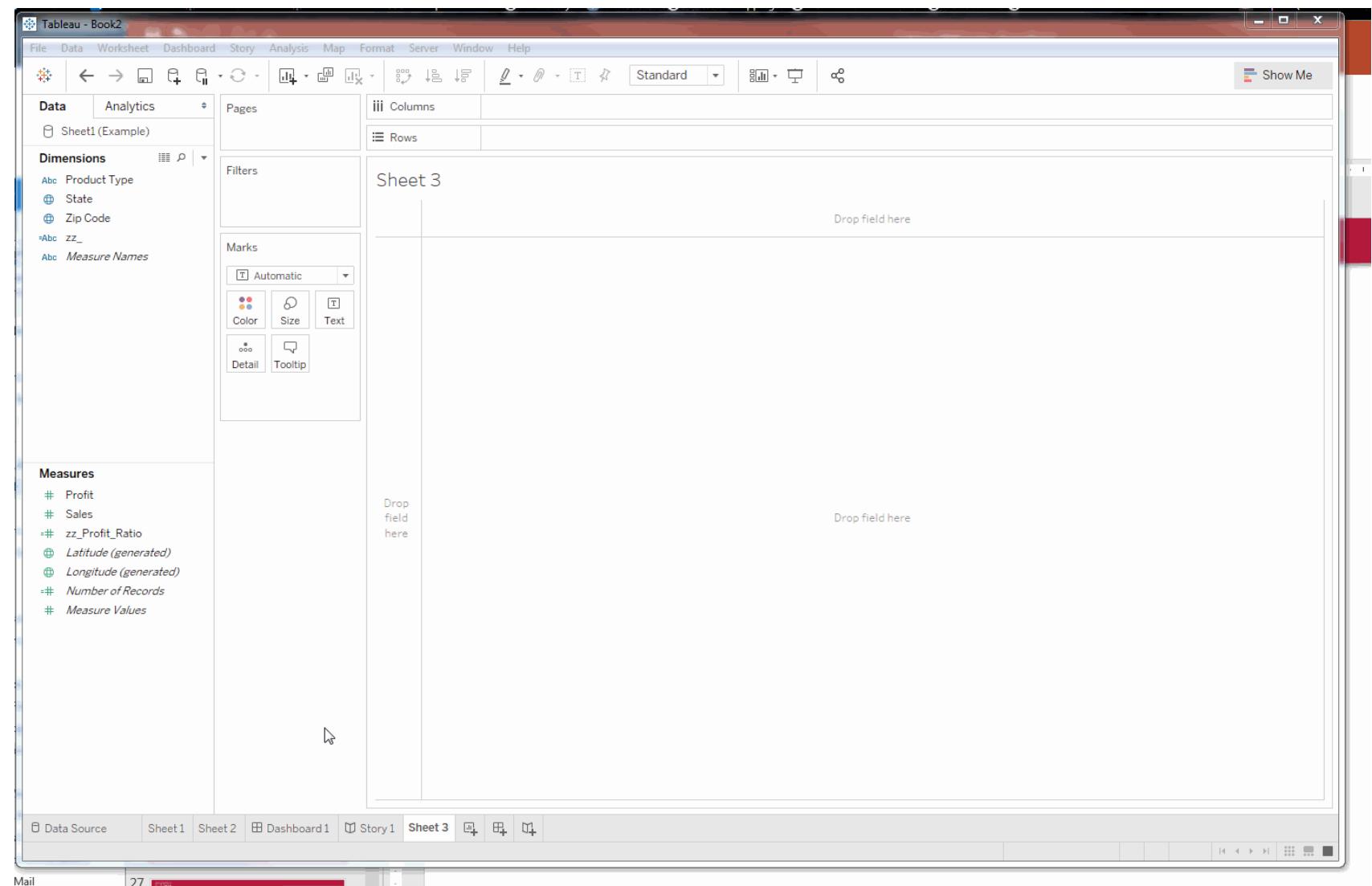


Tableau – General Overview: Bringing it all together

- Many different worksheets, text boxes, parameters, and filters come together to create a dashboard
- Multiple dashboards can be 'chained' together so that users are guided through multiple analytical paths

The dashboard is titled "Fishin' in the Mitten: Catching Big Fish in Michigan". It includes the following sections:

- Text Box**: Michigan has a diverse selection of gamefish for fishing adventures year round. Select a species from the drop down menu to find out helpful hints for catching the big ones!
- Text Box**: SELECT A FISH SPECIES (with a dropdown menu showing "BLUEGILL")
- Worksheet #3**: Michigan State Record(s) (with details: Year 1983, Waterbody Vaughn Lake, County Alcona, Inches 13.75, Pounds 2.75)
- Worksheet #2**: WHERE TO FIND BIG BLUEGILL (with a map of Michigan showing hot spots for big bluegill and a fish image)
- Text Box**: Based on Master Angler records from 1993-2014, the map to the left shows the 'hot spots' for big BLUEGILL. The size of the circle represents the total Master Angler award level fish caught in that location. Zoom in, or mouse over a circle for additional information about that location. Click on a circle to filter the results below for just that body of water.
- Worksheet #6**: TOP FISHING METHODS AND BAITS FOR BIG BLUEGILLS (with two bar charts)

Fishing Method/Bait	Percentage
Stillfishing	46.13%
Spincasting	16.47%
Driftfishing	12.75%
Baitcasting	10.33%
Ice Fishing	6.61%
Trolling	3.89%
Flycasting	3.63%

Fishing Method/Bait	Percentage
CRAWLER	28.27%
WORM	18.08%
WAX WORM	9.02%
LEECH	8.29%
RED WORM	2.16%
CRAWLER HARNESS	1.94%
LEAF WORM	1.86%
CRICKET	1.64%
- Worksheet #5**: (This section is labeled but does not have a visible chart or text box in the screenshot)
- Text Box**: (This section is labeled but does not have a visible chart or text box in the screenshot)
- Worksheet #4**: (This section is labeled but does not have a visible chart or text box in the screenshot)
- Worksheet #1**: (This section is labeled but does not have a visible chart or text box in the screenshot)

Use for Reporting

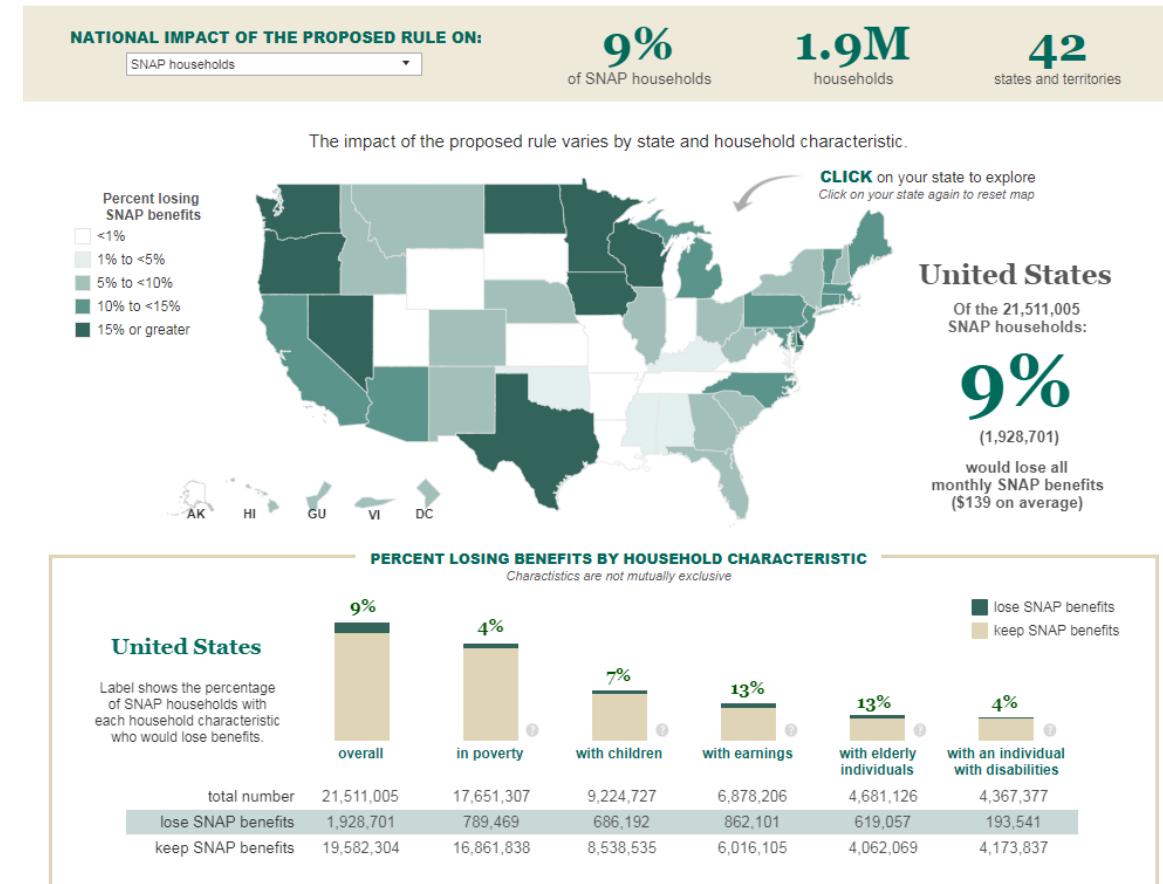
- Examples

Tableau – Reporting Example

- The results of detailed statistical analysis can be made available freely on Tableau Public where individuals can interact with data visualizations to view results – to supplement published research or publicly available reports
- Expands the audience for consuming research and provides a visual and interactive experience.

Impact of Proposed Policy Changes to SNAP Categorical Eligibility by State

On July 24, 2019, the United States Department of Agriculture (USDA) issued a proposed rule to eliminate SNAP categorical eligibility. USDA estimates that, under the proposed rule, 9 percent of current SNAP households would not otherwise meet SNAP's income and resource eligibility requirements and would therefore lose all of their SNAP benefits. This visualization explores the impact of the proposed rule by state and household demographic.

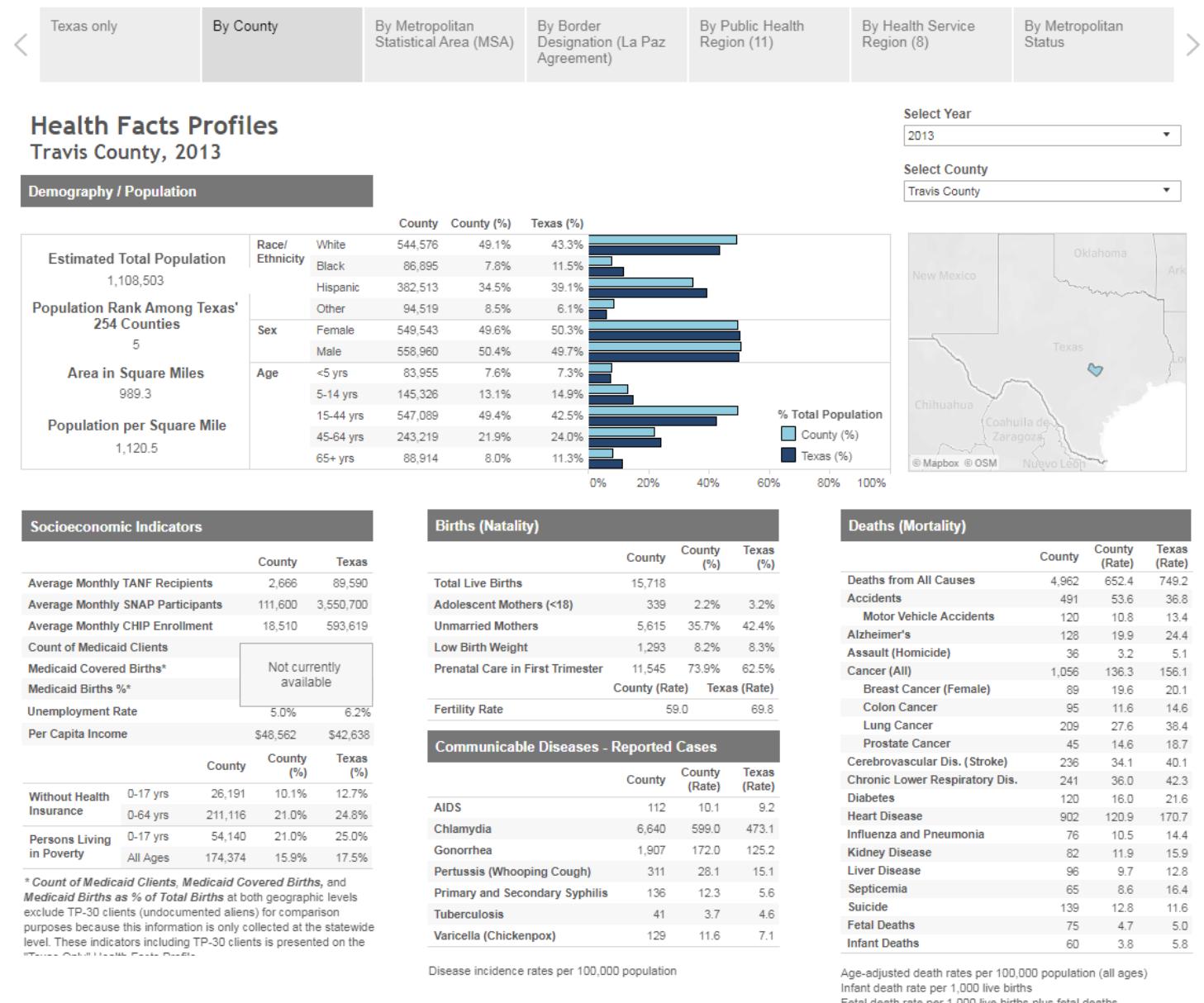


Source: Fiscal Year 2016 SNAP Quality Control sample. In FY 2016 there were 11 states that did not employ BBCE policies. Indiana implemented their BBCE policy in FY 2018. This visualization was funded by the Robert Wood Johnson Foundation. For detailed estimates or for more information about SNAP microsimulation modeling or Mathematica's work in this area, contact Senior Research Programmer Sarah Lauffer at SLauffer@mathematica-mpr.com or (206) 539-5792.

Tableau – Reporting Example

- Story Points – (a Tableau feature) provides a user experience similar to PowerPoint but with interactive data visualizations
- This allows for guided analytics where you create a general narrative and allow users to interact with visualizations to ‘deep dive’ into key points.

https://public.tableau.com/profile/texaschs#/vizhome/HFP_1/Story1



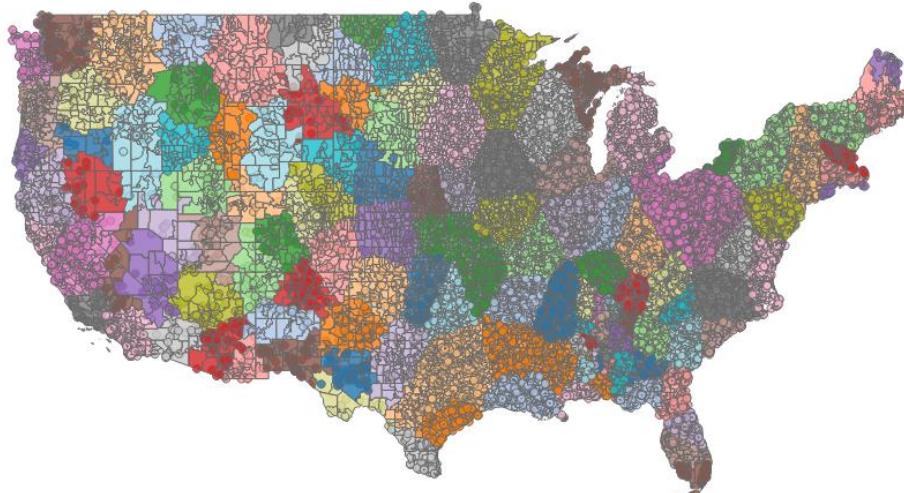
Use for Storytelling

- Examples

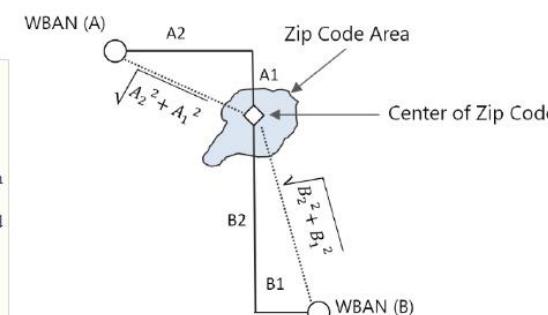
Tableau – Storytelling Example (Story Points)

Visualizing Weather Data in Tableau Software

< Introduction to the Methodology | Collecting the Data | Mapping WBAN to ZIP Codes | Visualizing Temperatures | Visualizing Precipitation | Polar Vortex >

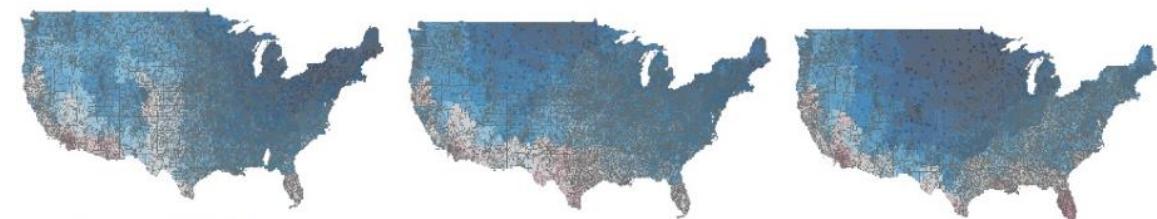


To map each ZIP code to the nearest WBAN, a simple pairwise comparison was used. For each ZIP code, a distance was calculated from the center of the ZIP code to the location of each WBAN utilizing the Latitude and Longitude data from the dataset. Each ZIP code was assigned to one WBAN sensor based on the minimum calculated distance in the pairwise matrix. The map above shows the ZIP codes assigned to each sensor.



Visualizing Weather Data in Tableau Software

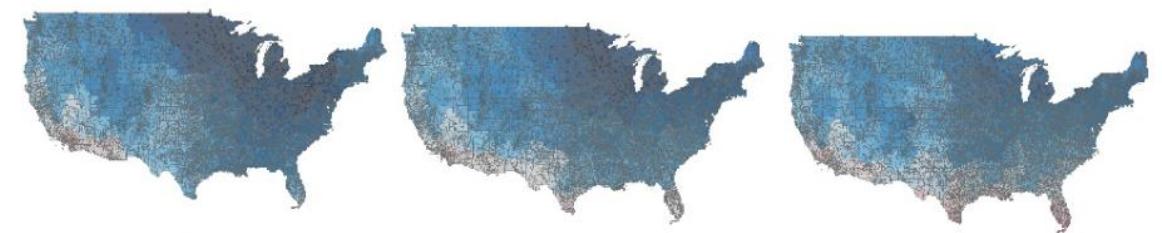
< Introduction to the Methodology | Collecting the Data | Mapping WBAN to ZIP Codes | Visualizing Temperatures | Visualizing Precipitation | Polar Vortex >



January 4th, 2014

January 5th, 2014

January 6th, 2014



January 7th, 2014

January 8th, 2014

January 9th, 2014

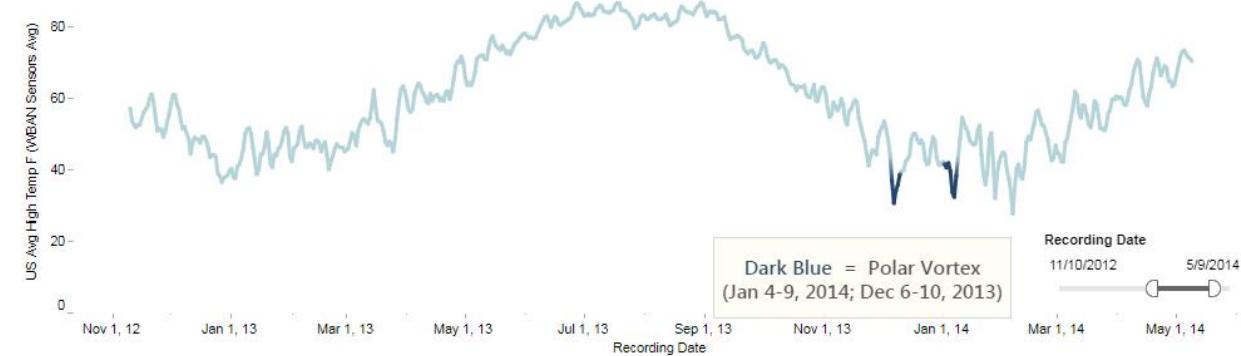
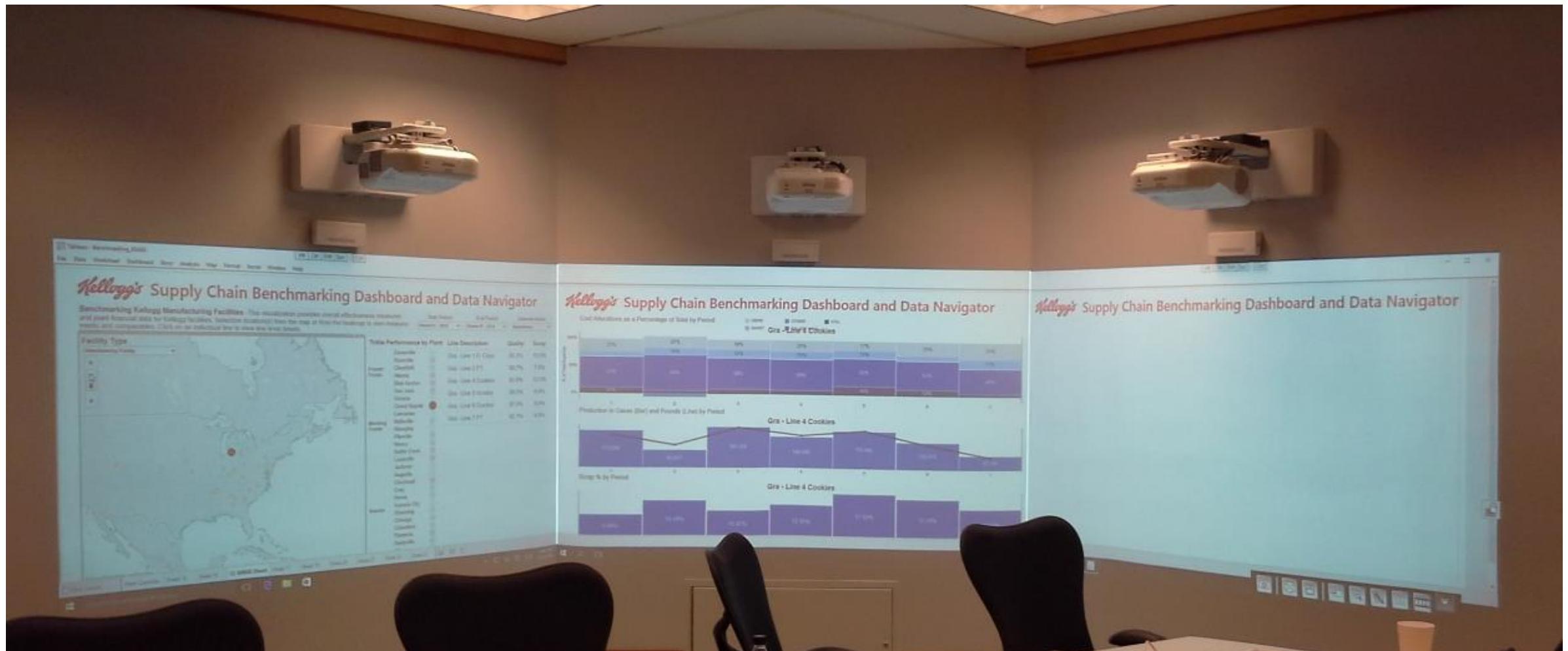


Tableau – Storytelling Example (K-MAX)

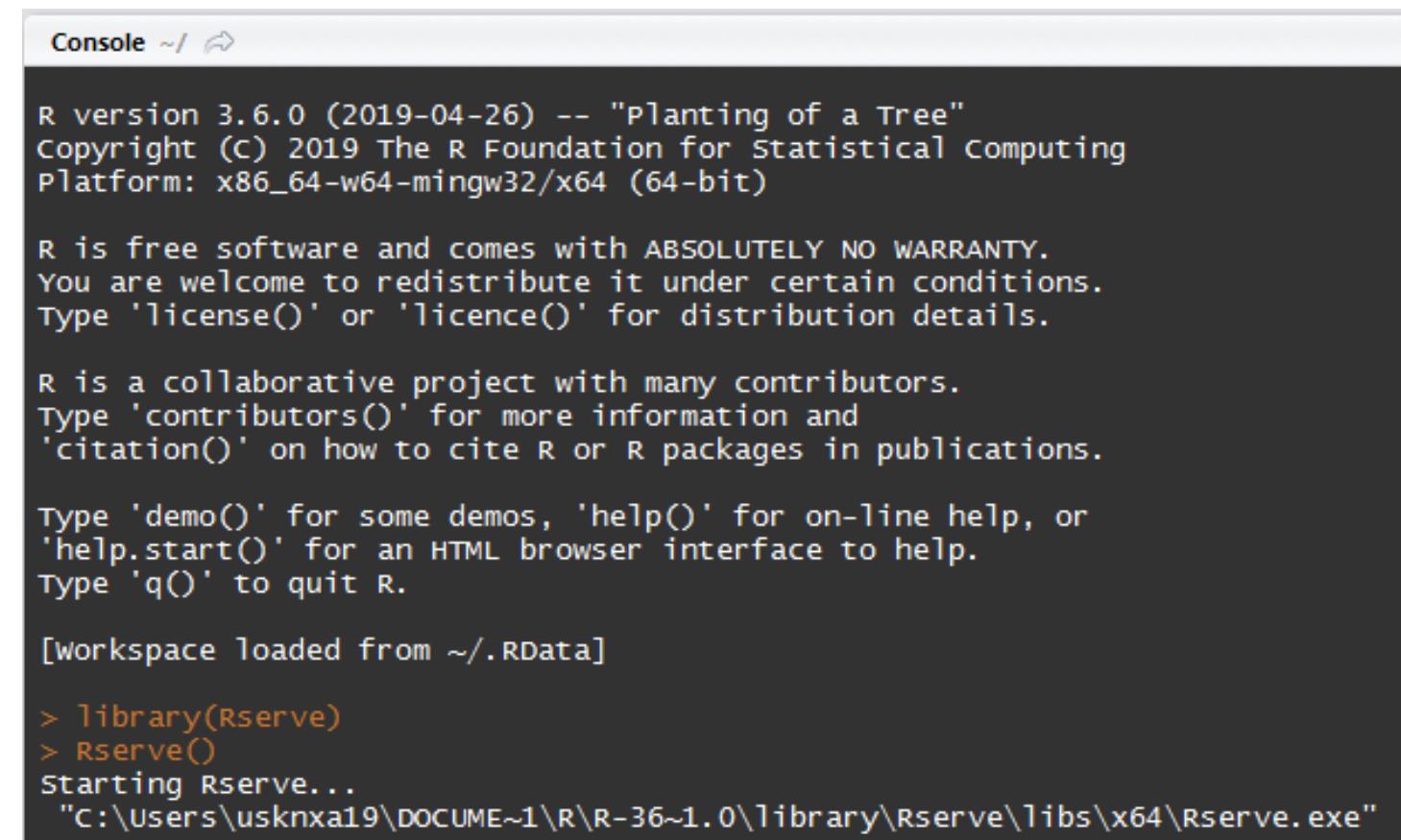


Advanced Features

- Examples

Advanced Features – Connecting Tableau to “R”

- Step #1
 - Install “R” or “R” Studio on your computer
 - Load the Rserve library package
 - Start Rserve



The image shows a screenshot of an R console window. The title bar says "Console". The window displays the standard R startup message, which includes the version number (R version 3.6.0), copyright information, and various usage instructions. At the bottom of the message, it says "[workspace loaded from ~/.RData]". Below the message, there are two lines of R code: "`> library(Rserve)`" and "`> Rserve()`". Following these commands, the text "Starting Rserve..." is displayed, along with the path to the Rserve executable: "`"C:\Users\usknxa19\DOCUMENTS\R\R-3.6.1\library\Rserve\libs\x64\Rserve.exe"`".

```
R version 3.6.0 (2019-04-26) -- "Planting of a Tree"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[workspace loaded from ~/.RData]

> library(Rserve)
> Rserve()
Starting Rserve...
"C:\Users\usknxa19\DOCUMENTS\R\R-3.6.1\library\Rserve\libs\x64\Rserve.exe"
```

Advanced Features – Connecting Tableau to “R”

- Step #2

- Connect Tableau to your Rserve instance

The screenshot shows the Tableau desktop application with a data source named "MLBGAMES". The interface includes a top menu bar with File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, and Help. Below the menu is a toolbar with various icons. The left side features a shelf with "Dimensions" (Gameday, Away, Home, Homeruns, ID, zz_date, Measure Names) and "Measures" (F1, zz_R_Script, Number of Records, Measure Values). A "Parameters" section contains a "Date Par" parameter. The main workspace displays a data table with columns: Gameday, Away, Awayruns, Home, and Homeruns. The data shows records for various teams on June 4, 2019, such as Atlanta (12 away runs), Pittsburgh (5 home runs), and so on. To the right of the workspace is a "Show Me" panel with various visualization options. At the bottom, there are tabs for Data Source, Sheet 1, Dashboard 1, etc., and a status bar indicating 15 marks and 15 rows by 1 column.

Gameday	Away	Awayruns	Home	Homeruns
6/4/2019	Atlanta	12	Pittsburgh	5
	Baltimore	12	Texas	11
	Boston	8	Kansas City	3
	Chi White S..	5	Washington	9
	Cincinnati	4	St. Louis	1
	Colorado	3	Chi Cubs	6
	Houston	11	Seattle	5
	LA Dodgers	9	Arizona	0
	Miami	16	Milwaukee	0
	Minnesota	2	Cleveland	5
	NY Yankees	3	Toronto	4
	Oakland	4	LA Angels	2
	Philadelphia	9	San Diego	6
	San Francis..	9	NY Mets	3
	Tampa Bay	6	Detroit	9

Advanced Features – Connecting Tableau to “R”

- Step #3
 - Write “R” script within a calculated field in Tableau

Note: This is also generally the same way to connect Tableau to Python in Anaconda – with a few small configuration differences.

Measures

```
=#  Away Score
=Abc Away Team
=#  Home Score
=Abc Home Team
=#  zz_Axis_1
=#  zz_date2
=#  zz_Total
=Abc zz_Winner_Color_AWAY
=Abc zz_Winner_Color_HOME
=#  Number of Records
#  Measure Values
```

Parameters

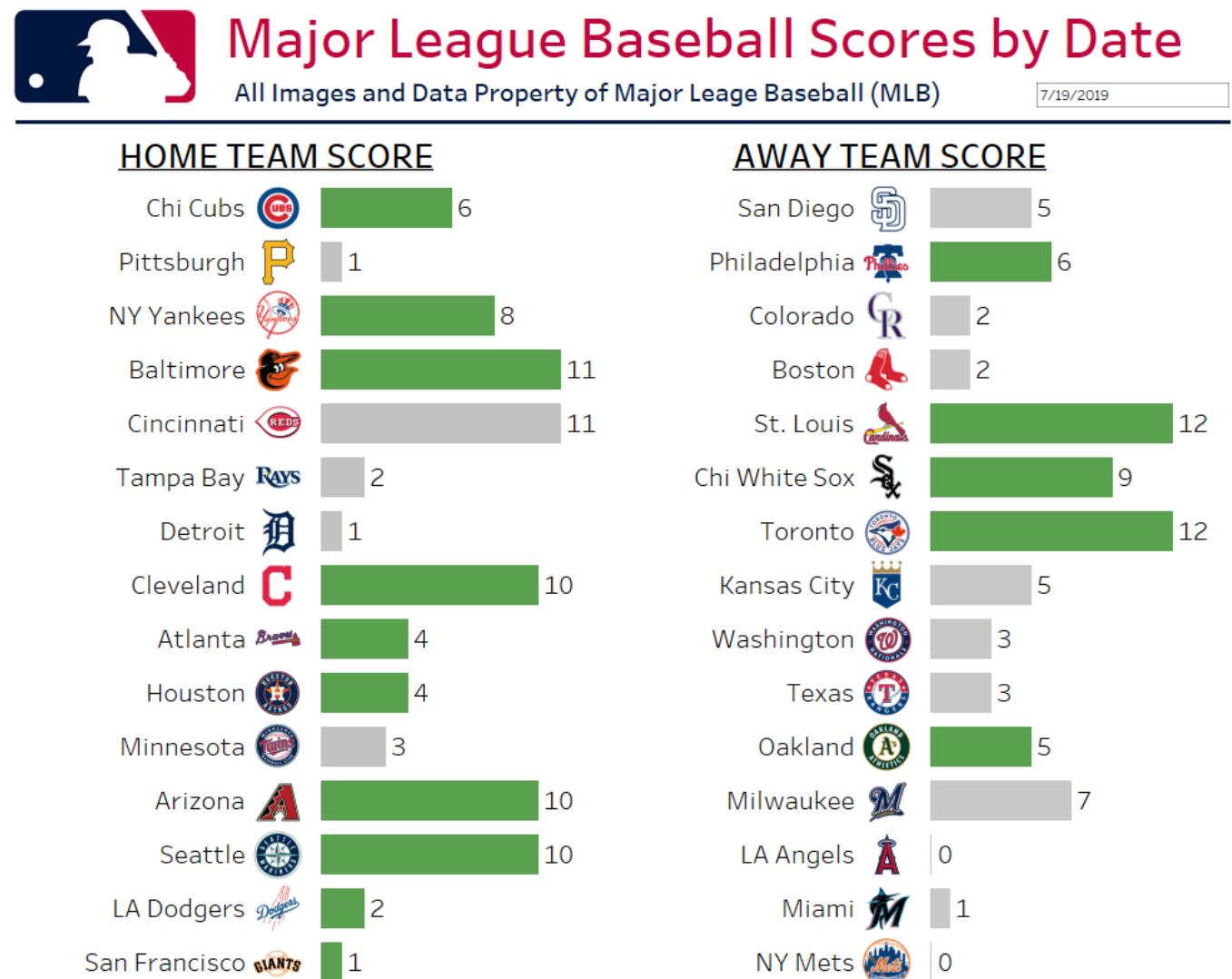
```
>Date Par
```



```
INT(SCRIPT_Str("library(xml2);
dater <- as.Date(Sys.Date()-.arg2);
year <- paste('year ', format(dater, '%Y'), '/', sep = '');
month <- paste('month ', format(dater, '%m'), '/', sep = '');
day <- paste('day ', format(dater, '%d'), '/', sep = '');
xmlFile <-
paste('http://gd2.mlb.com/components/game/mlb/', year,
month, day, 'miniscoreboard.xml', sep = '');
x <- read_xml(toString(xmlFile));
games=xml_children(x);
ns <- xml_ns(x);
awayruns <-xml_attr(games,'away_team_runs',ns);
awayrun sdf <- as.data.frame(awayruns);
awayrun sdf$ID <- seq.int(nrow(awayrun sdf));
toString(awayrun sdf[arg1, 1]);
",MAX([Idvalue]),max([zz_date])))")
```

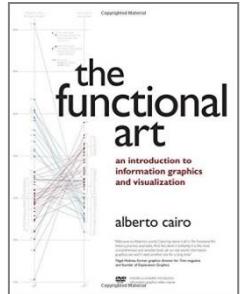
Advanced Features – Example

- Example that queries Major League Baseball's open API for statistics
- "R" script downloads data as an XML file, parses the data and returns the results to Tableau for visualization.

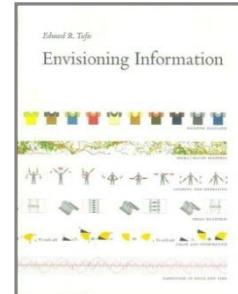


Available Resources

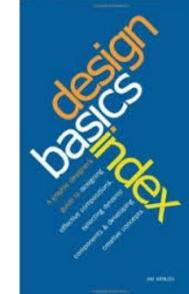
Books



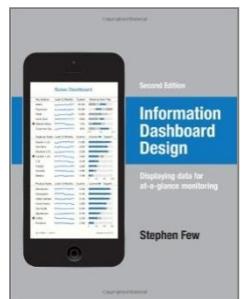
The Functional Art
Alberto Cairo



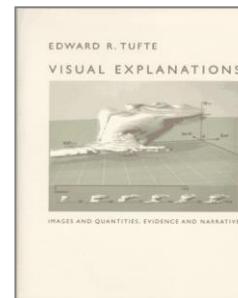
Envisioning Information
Edward Tufte



Design Basics Index
Jim Krause



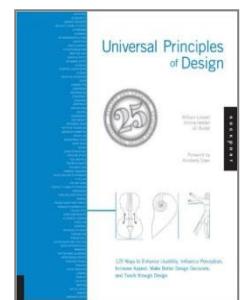
Information Dashboard Design
Stephen Few



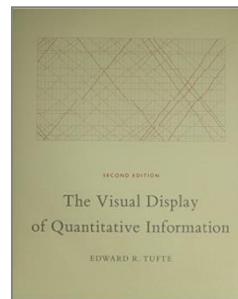
Visual Explanations
Edward Tufte



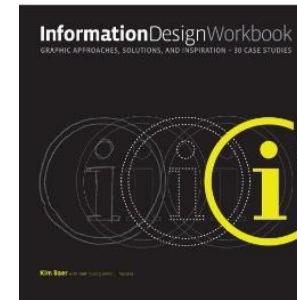
Beautiful Evidence
Edward Tufte



Universal Principles of Design
William Lidwell



The Visual Display of Quantitative Information
Edward Tufte



Information Design Workbook
Kim Baer

Tableau Public & Other Resources

<https://public.tableau.com/s/gallery>

- Daily inspiration through 'viz of the day'
- A place to upload your work to the cloud
- Open environment to share visualizations and data (don't post confidential data here ☺)

<http://www.visualnews.com/>

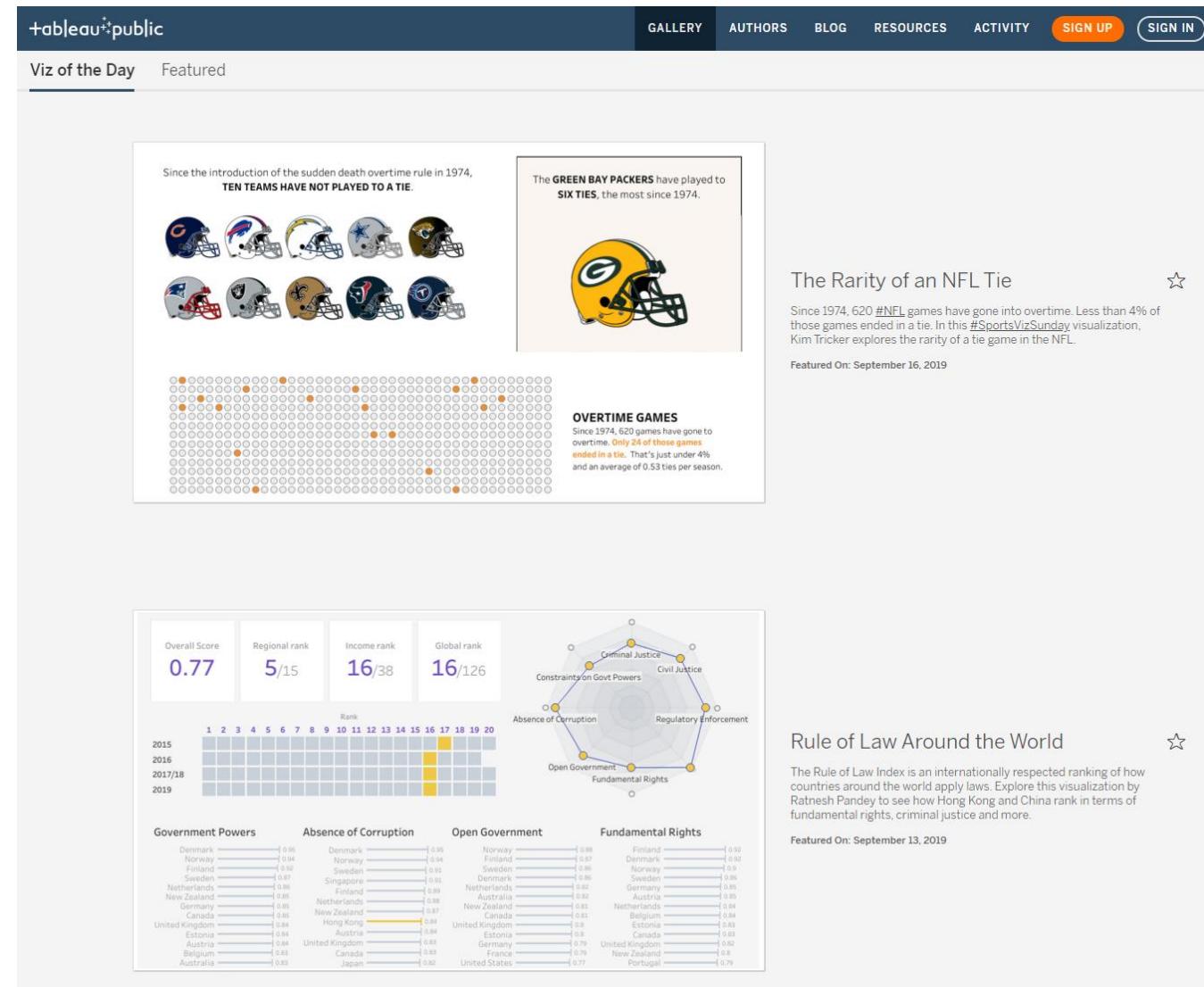
<http://www.flowingdata.com>

<http://www.thisiscoLOSSAL.com/>

<http://vizwiz.blogspot.com/>

<http://www.datavizdoneright.com/>

National Geographic Magazine
Bloomberg Businessweek



West Michigan Tableau Users Group (WMTUG)



<https://community.tableau.com/groups/west-michigan>

- Meet three to four times a year in Kalamazoo or Grand Rapids
- 100-150 participants
- Sharing tips, tricks, and case studies
- Develops a strong network with other analytics focused individuals



Tableau Conference

- 15,000 of your best data visualization friends in the same place
- One week of in-depth sessions on data visualization and Tableau software

