Institut Informatika & Bisnis DARMAJAYA
Yayasan Alfian Husin

SAINS DATA
Institut Informatika & Bisnis Darmajaya

Kampus Merdeka INDONESIA JAYA

MERDEKA BELAJAR

DATA SCIENCE DARMAJAYA
"YOUR BEST FUTURE IN DATA"

MEETING: [17]

An Introduction to Analysis and Data Visualization using TABLEAU SOFTWARE

BY: HENDRA KURNIAWAN

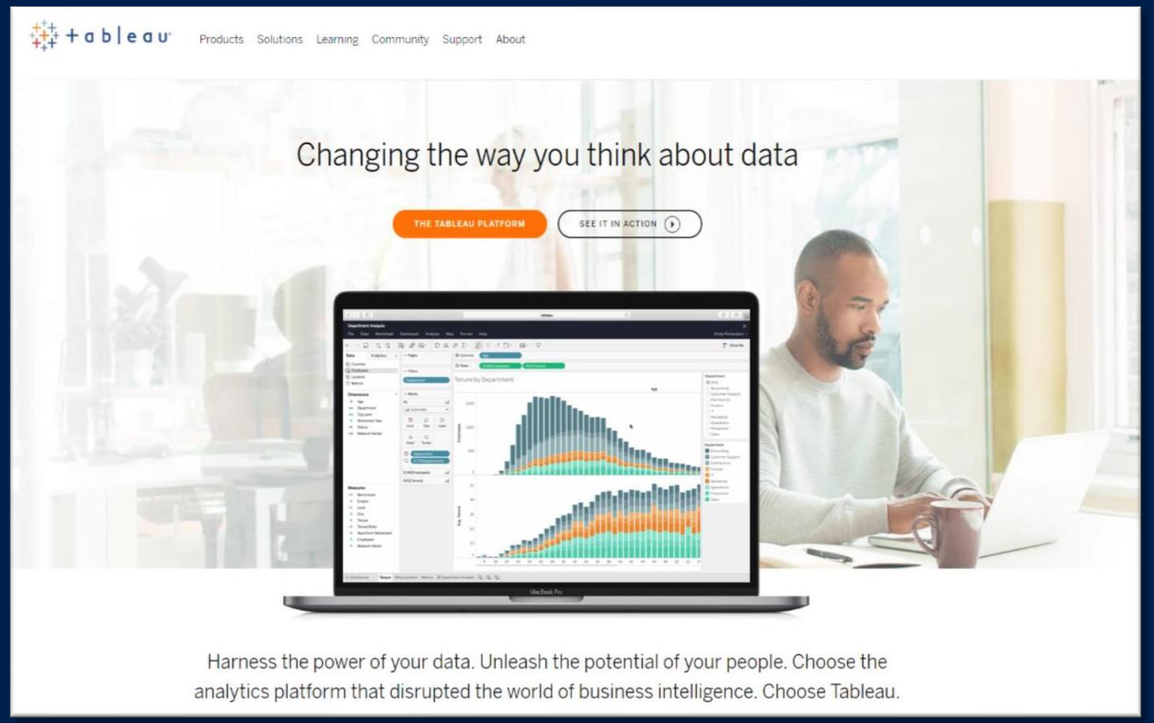
[12-09-2023] | [KODE MK: BDG23407 | MATAKULIAH: DATA SCIENCE AND BUSINESS ANALYTICS | SKS: 2/2 | VERSI: 01

ANALYTICS AND DATA VISUALIZATION USING TABLEAU

1. What is Tableau?
2. **What is Data Visualization?**
3. General Overview of Tableau
4. Use for Reporting - Examples
5. Use for Storytelling - Examples
6. Use for Analysis – Examples
7. Advanced Feature – Examples

What is Tableau Software?

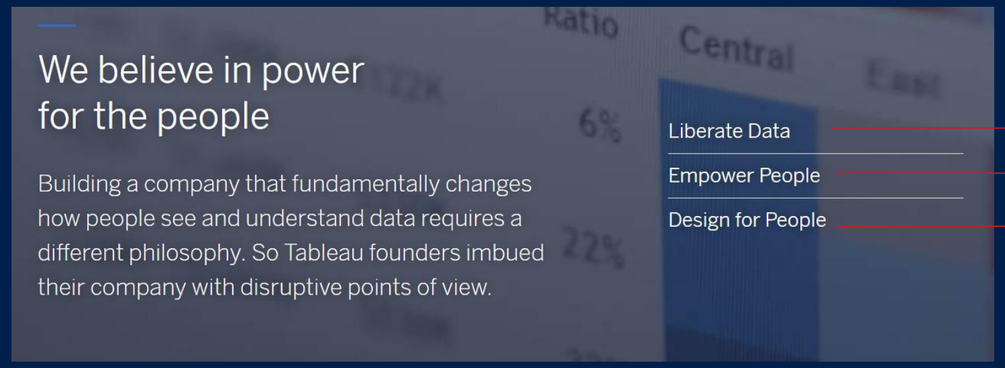
- Software company Founded in 2003 from Stanford research
- Intent is to bring 'data to the people' through easy to use data visualization software
- Would be classified as a hybrid business intelligence (BI) / analytics software company
- Used by many of the largest companies in the world and most large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

What is Tableau Software?



Courtesy: www.Tableau.com

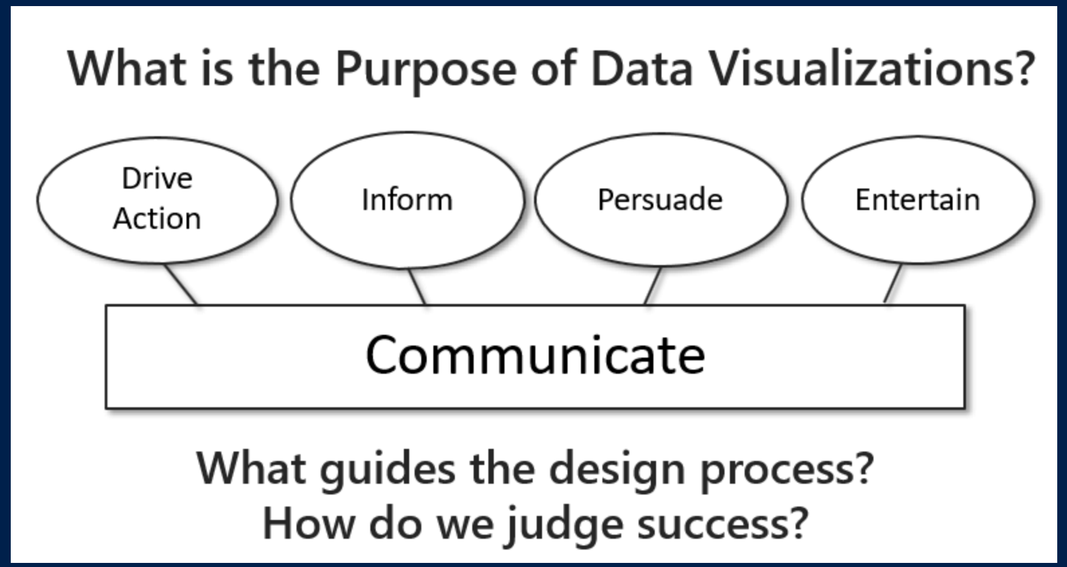
The main focus of Tableau software is for you to better understand your datasets, especially large datasets.

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.

What is Data Visualization?

What is Data Visualization?

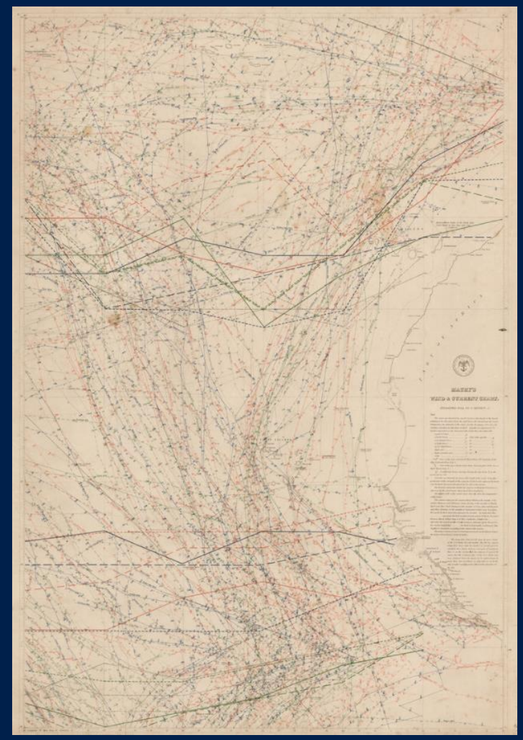


What is Data Visualization?

https://en.wikipedia.org/wiki/Matthew_Fontaine_Maury

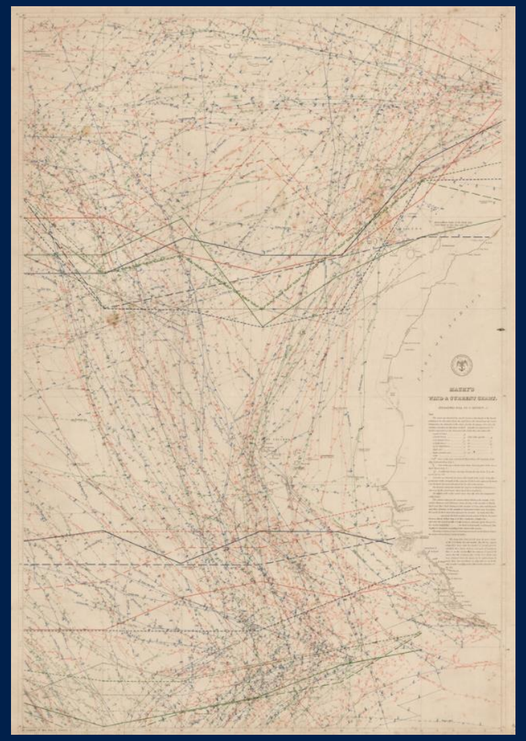
Matthew Fontaine Maury

- 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

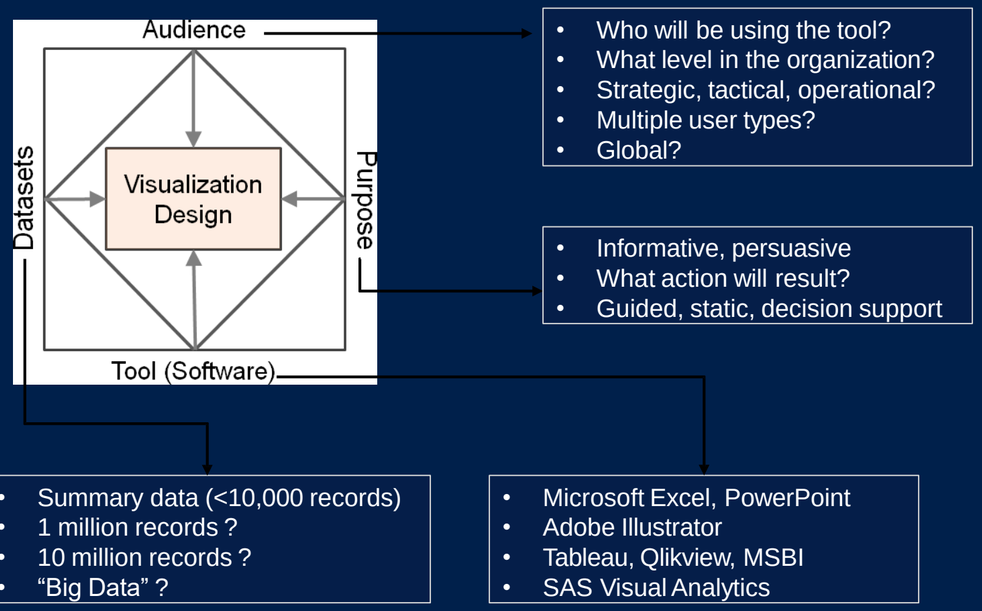
- Visualization of his team's findings
- Use of symbols and colors to highlight best routes
- Findings were counter-intuitive (heading west to go faster east)

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

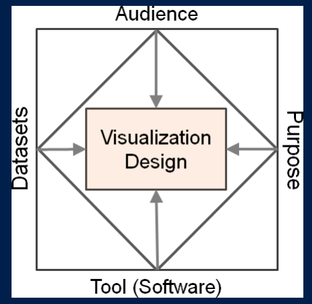
1. Identify Purpose (Intended Use)
2. Consider Audience
3. Research
   i. Identify Available Datasets
   ii. Identify Data Elements
   iii. Benchmark Designs
4. Design
   i. Sketch
   ii. Iterate
   iii. Collect Feedback
5. Execute Design
   i. Collect Feedback
6. Document – Deploy
7. Sustain

What is Data Visualization?

Example – Decision Support

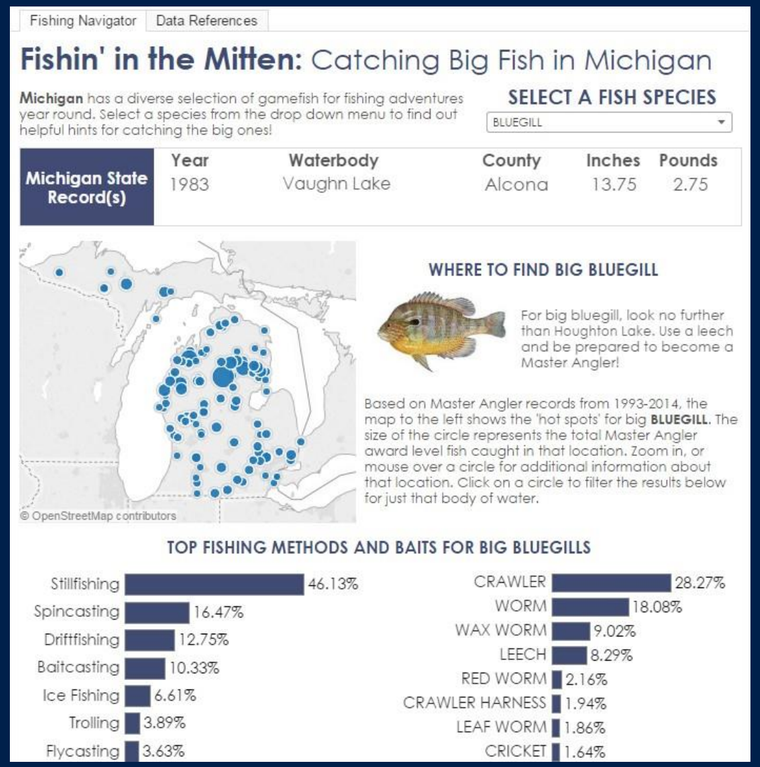
Those looking to catch big fish in Michigan

Michigan DNR Database; Public Use Pictures



Provide decision support to increase chances of catching big fish

Tableau



What is Data Visualization?

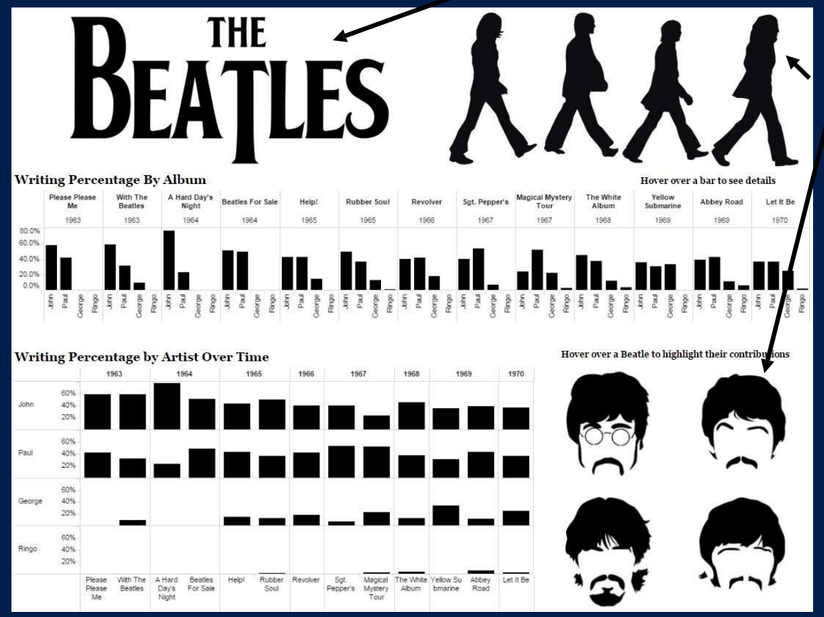
Elements of Design - Unity

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

Consistent Color (lack)

Simplified Images

Consistent Font



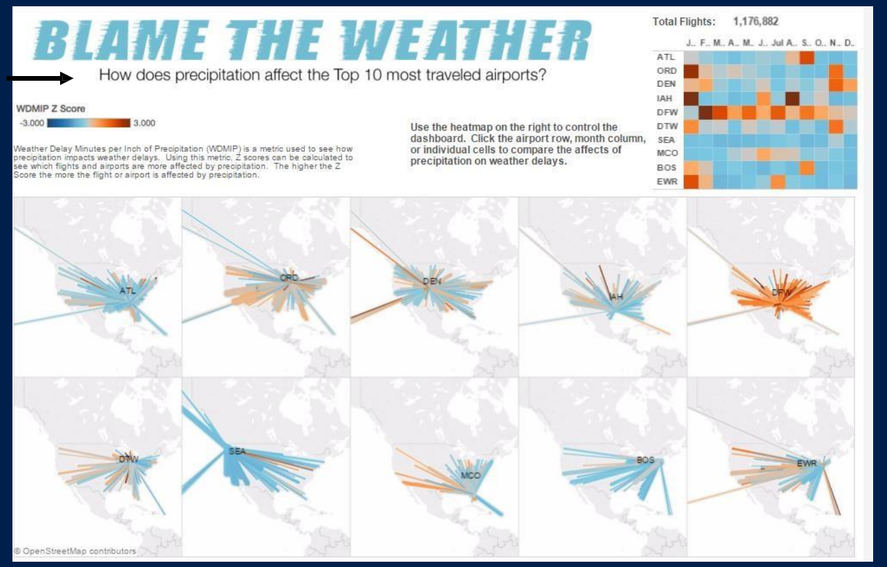
https://public.tableau.com/s/gallery/beatles-albums Author: Mike Moore

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



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.

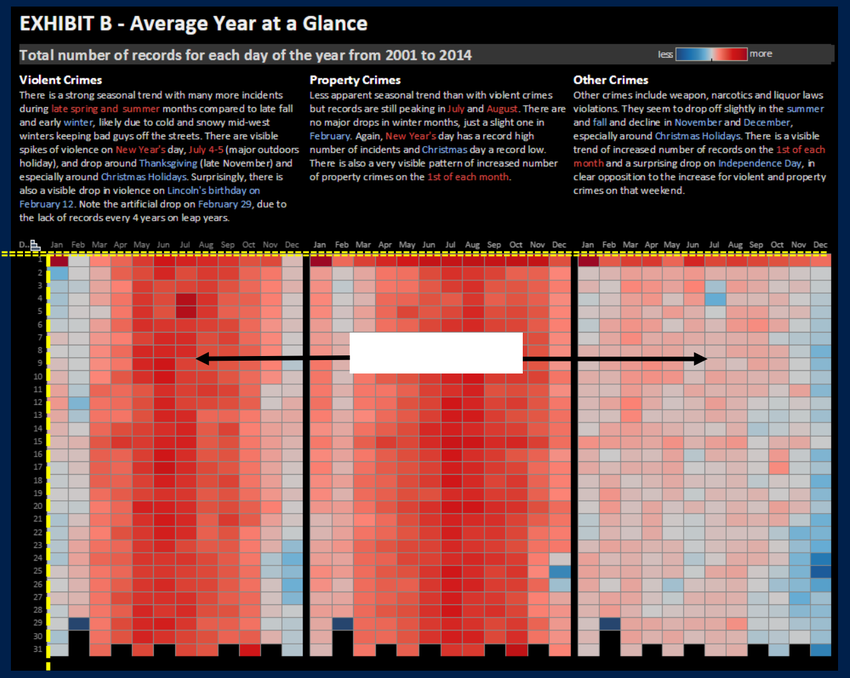
https://public.tableau.com/s/gallery/road-accidents-germany
Author: Oliver Linder

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

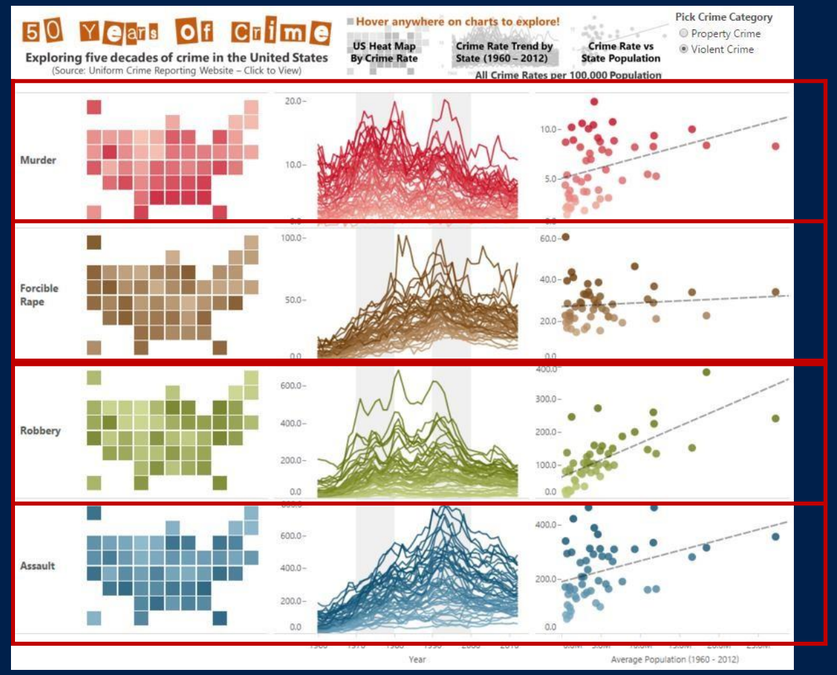


https://public.tableau.com/s/gallery/chicago-crime-scene Author: George Gorczynski

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

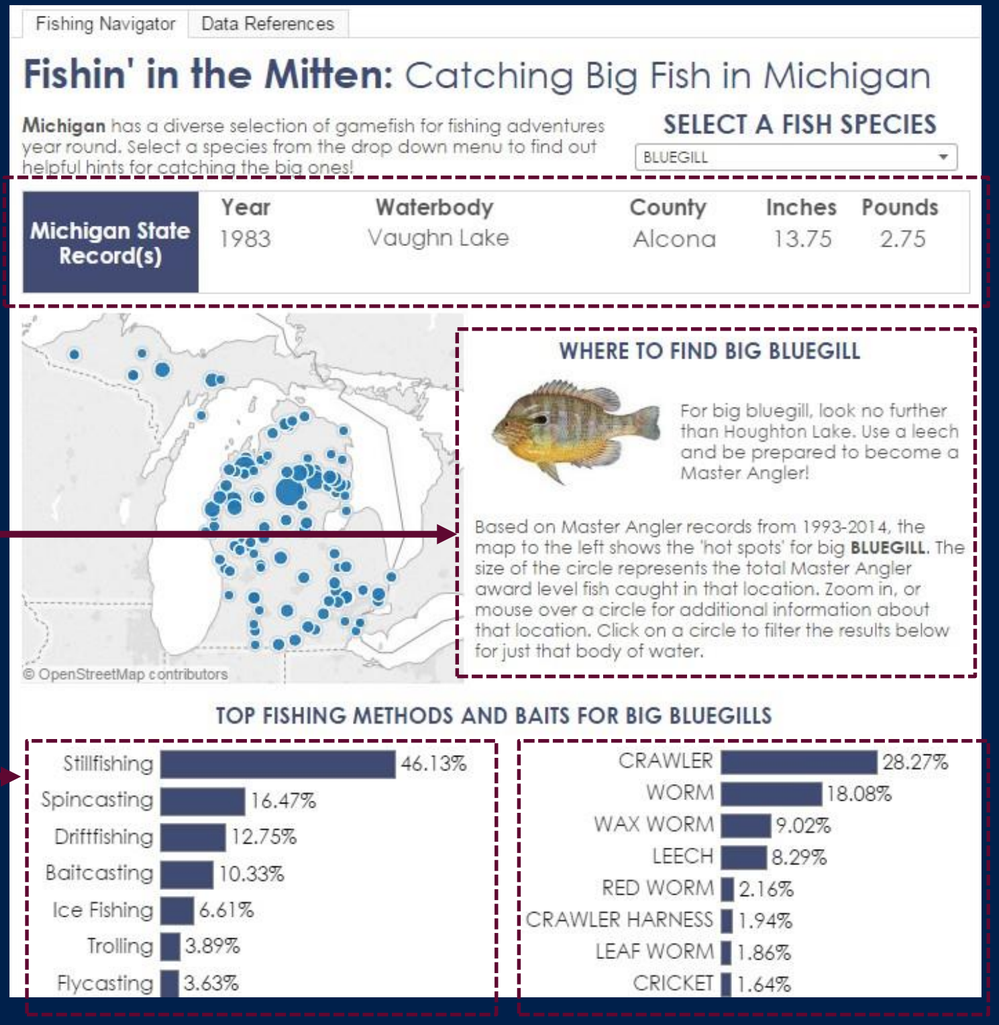
Detailed Example - Design

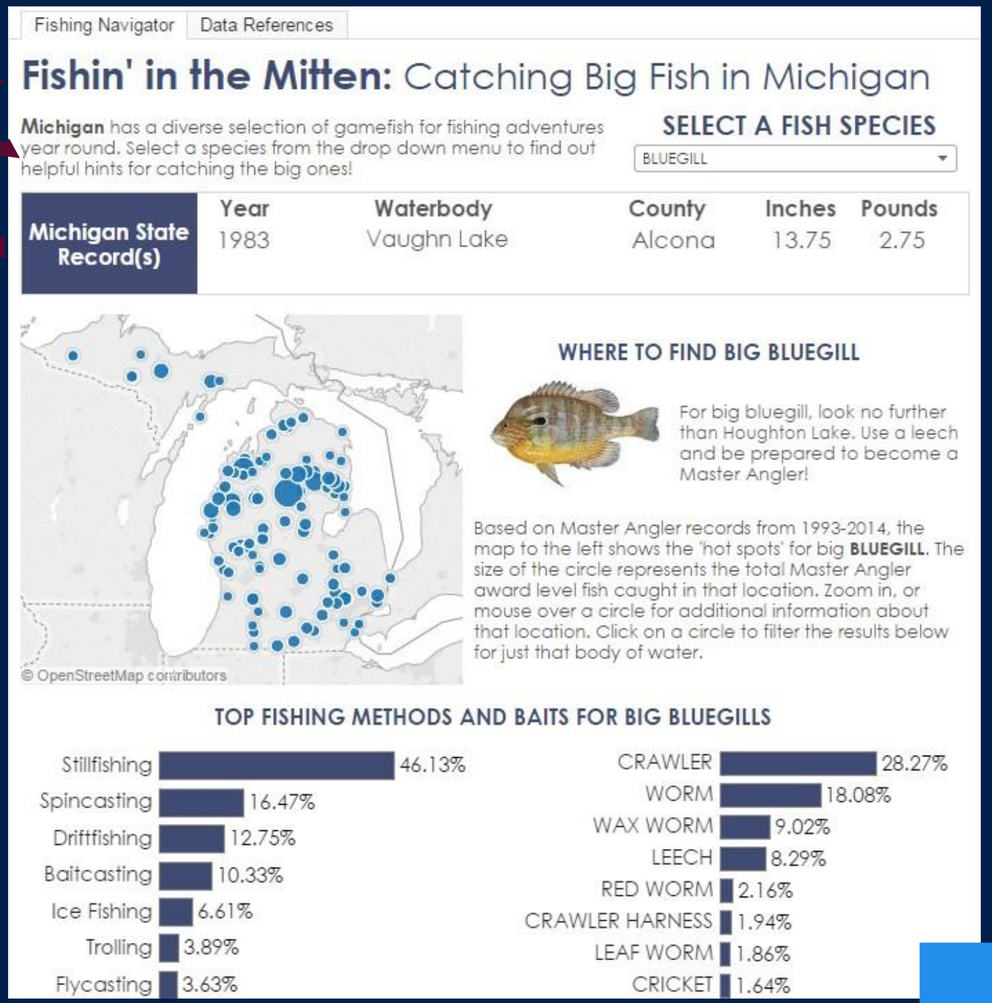
Hierarchy

Grouping

Grouping

Balance





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)

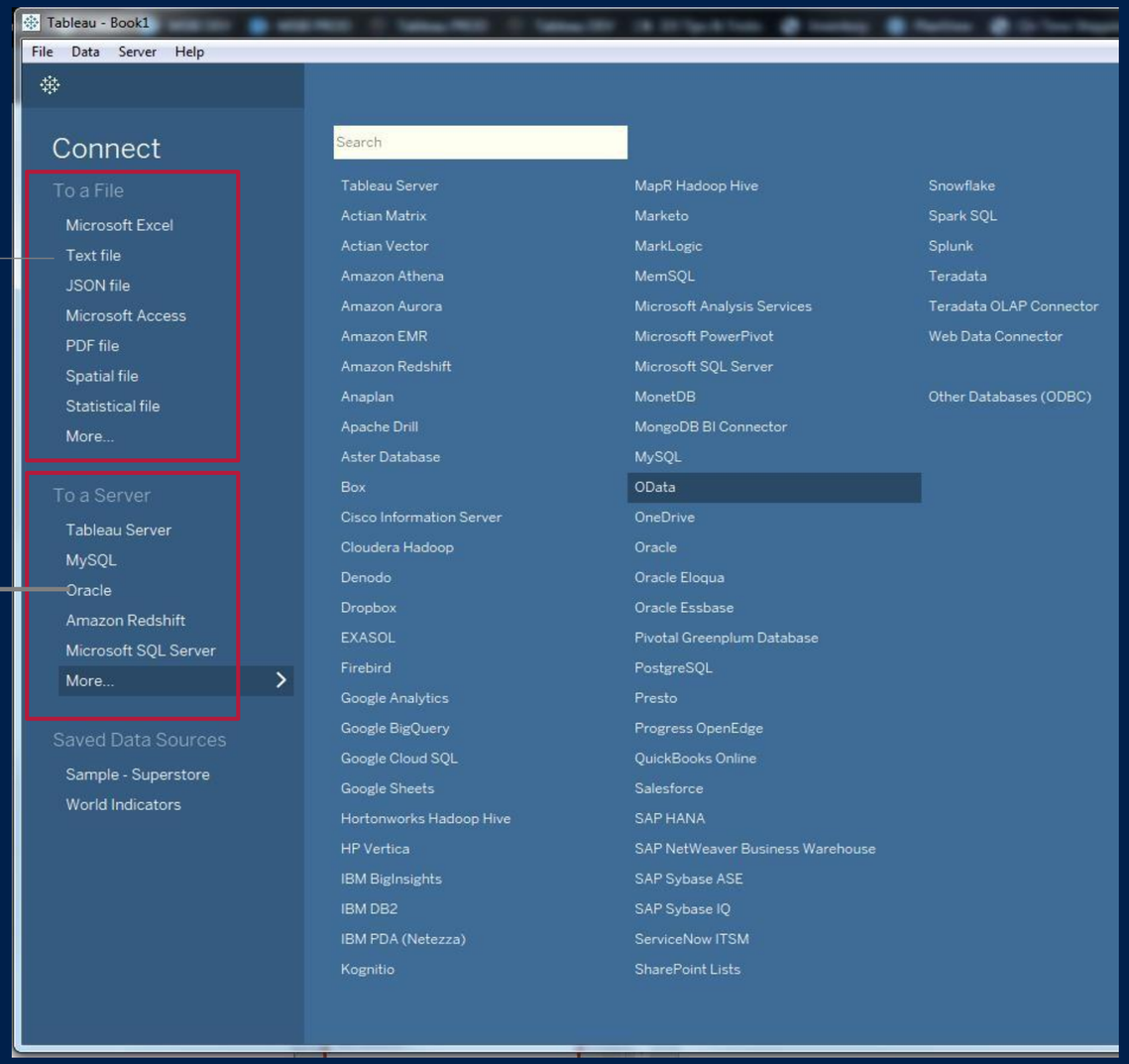


Tableau – General Overview – simple example

- 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**	**Profit**
2	Chairs	Michigan	49012	$ 78,847	$ 13,462
3	Chairs	Michigan	49008	$ 42,998	$ 2,578
4	Chairs	Michigan	49014	$ 39,554	$ 3,998
5	Chairs	Michigan	49007	$ 10,973	$ 2,112
6	Chairs	Ohio	45891	$ 9,558	$ 940
7	Chairs	Ohio	45888	$ 51,831	$ 6,443
8	Chairs	Ohio	45871	$ 34,972	$ 5,504
9	Desks	Michigan	49012	$ 30,838	$ 1,134
10	Desks	Michigan	49008	$ 71,298	$ 5,720
11	Desks	Michigan	49014	$ 11,558	$ 1,756
12	Desks	Michigan	49007	$ 74,435	$ 31
13	Desks	Ohio	45891	$ 52,503	$ 10,610
14	Desks	Ohio	45888	$ 45,530	$ 397
15	Desks	Ohio	45871	$ 84,076	$ 6,996

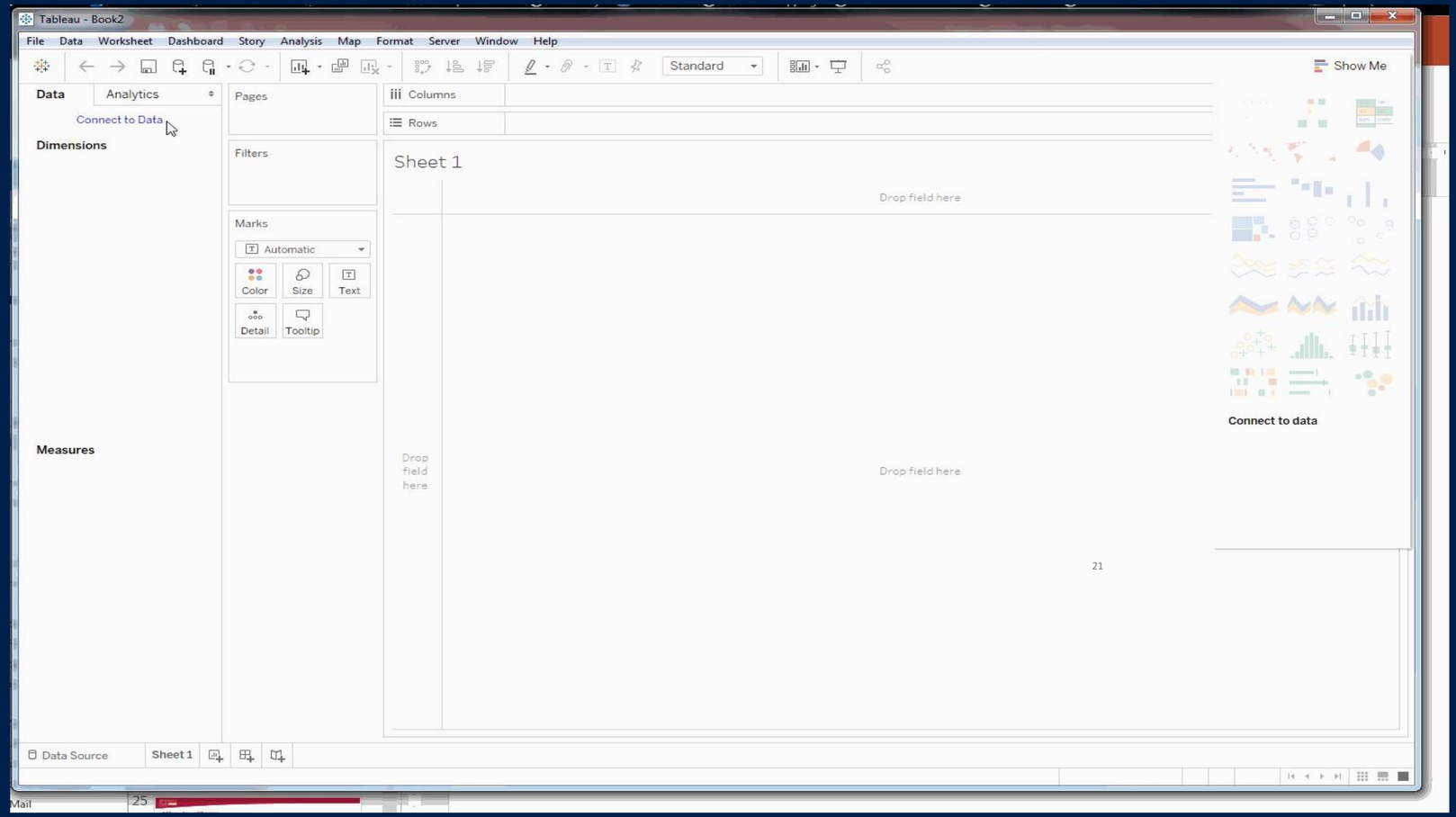


Tableau - Book2
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Connect to Data
Dimensions
Measures
Pages
Columns
Rows
Filters
Marks
Automatic
Color
Size
Text
Detail
Tooltip
Sheet 1
Drop field here
Drop field here
Drop field here
Connect to data
21
Data Source
Sheet 1

Tableau – General Overview
Calculated Fields

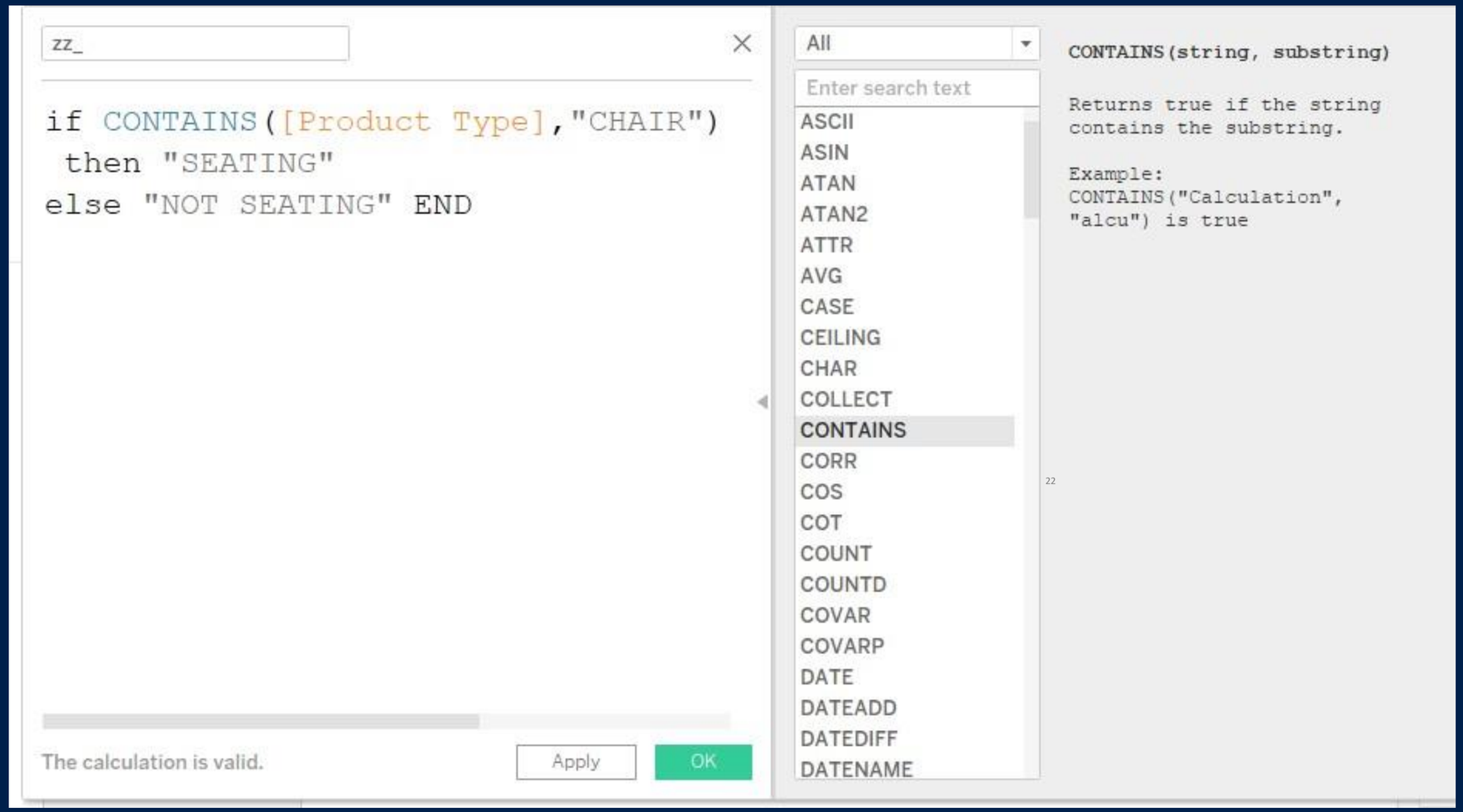


Tableau – General Overview Basic Analytics

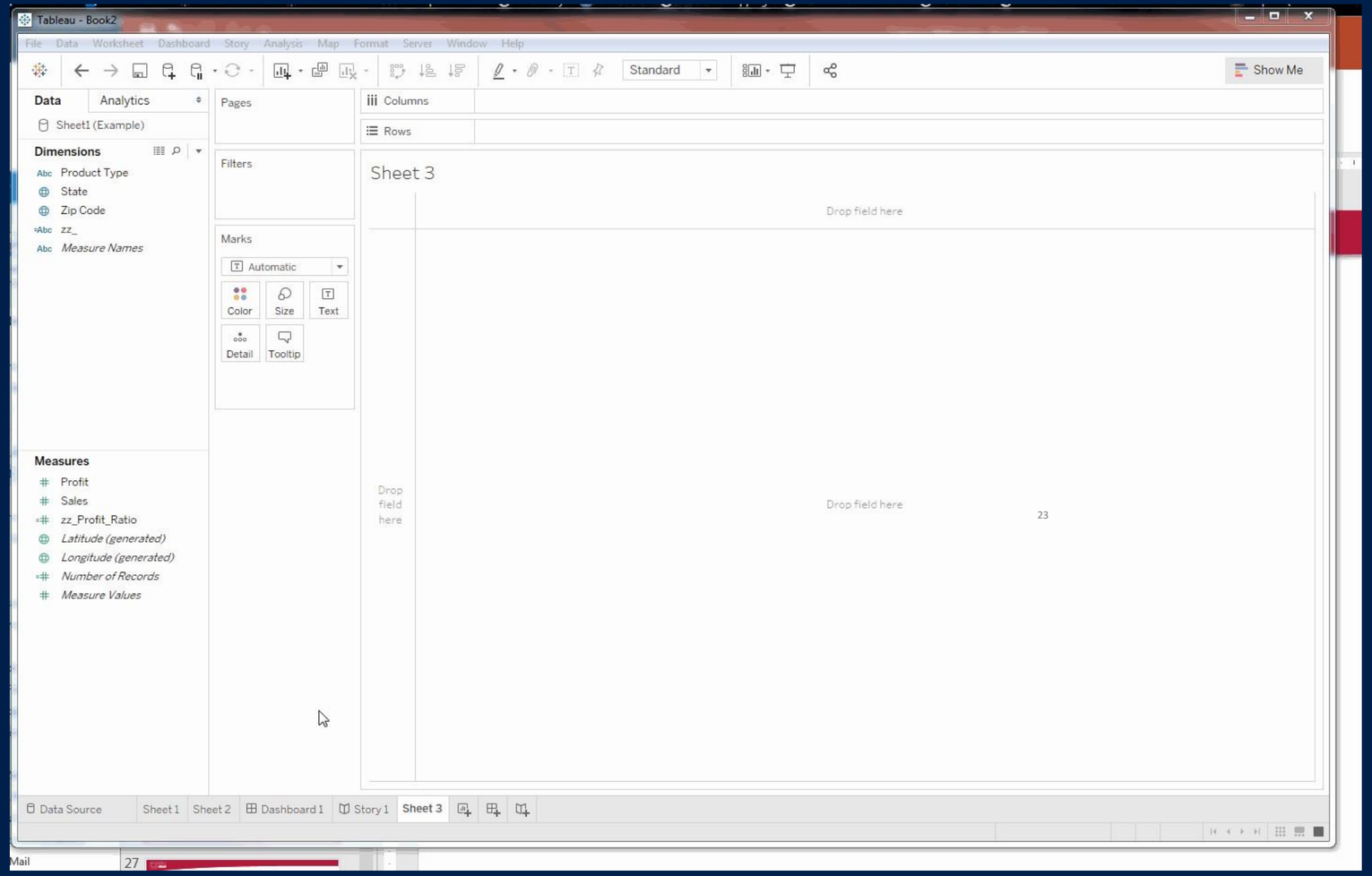
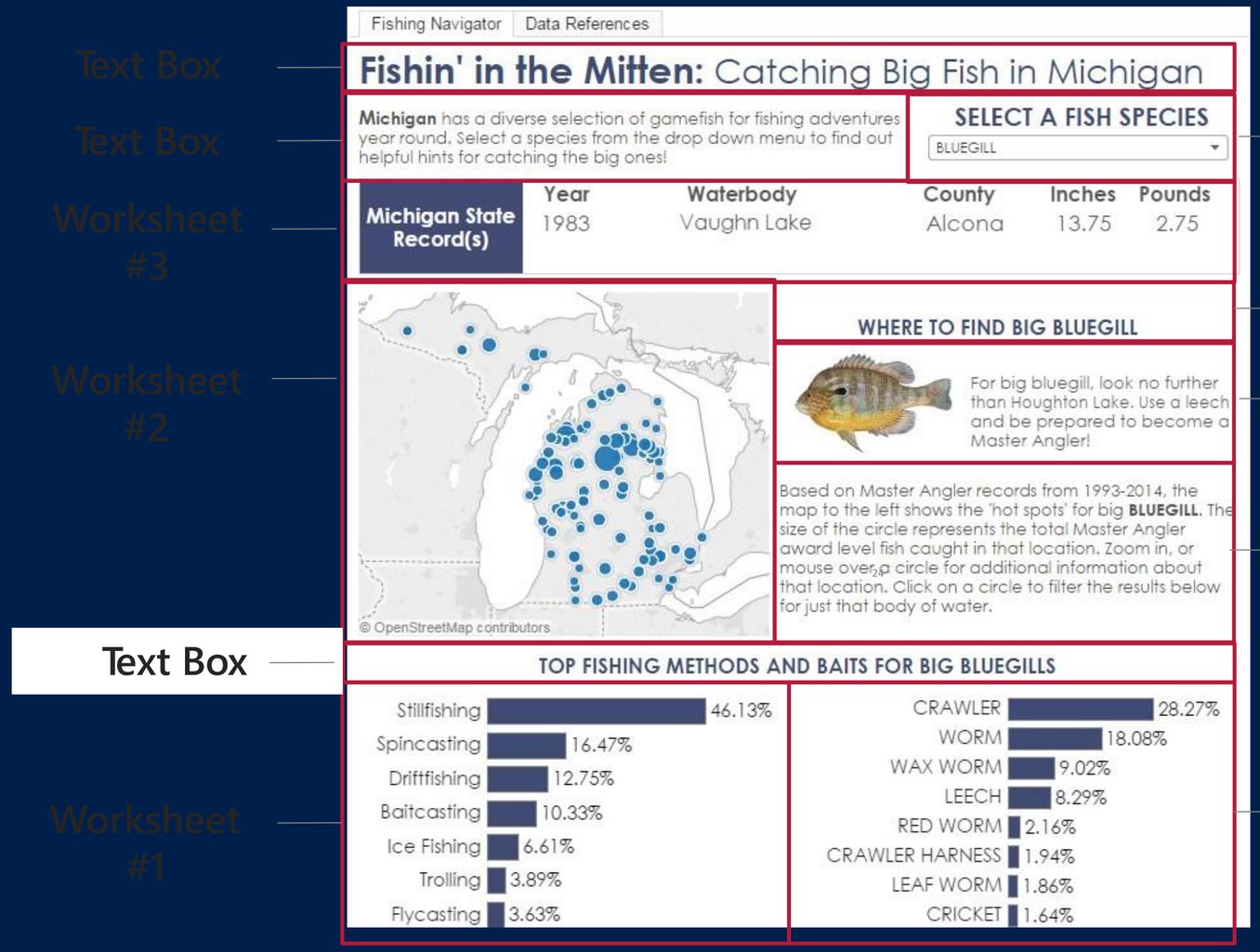


Tableau – General Overview: Bringing it all together

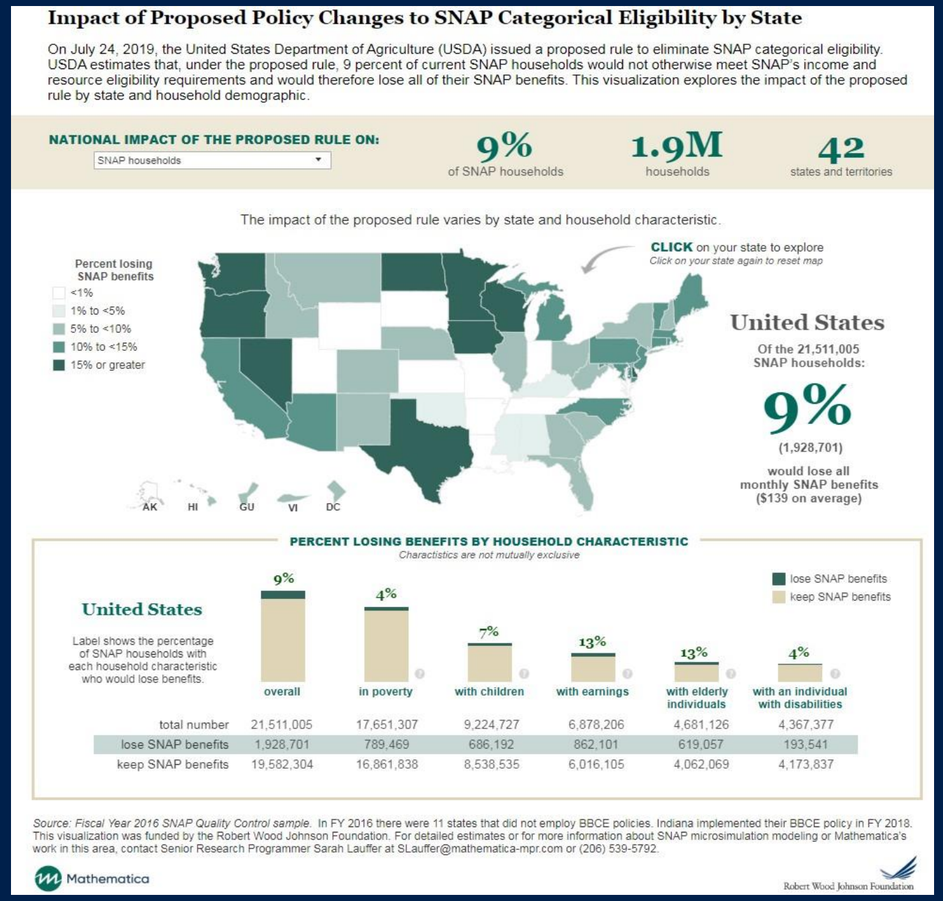
- Many different worksheets, text boxes, parameters, and filters come together to create a dashboard
- Multiple dashboards can be 'chained' together so that users are guided through multiple analytical paths



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.

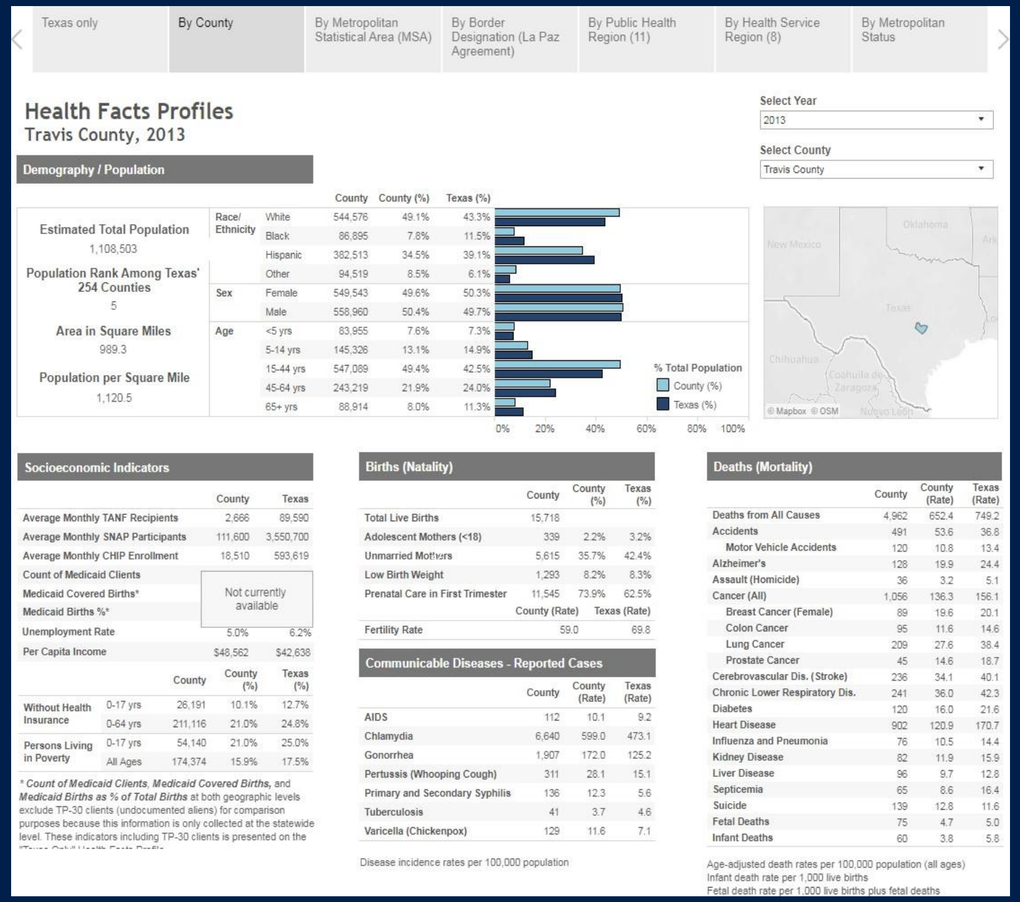


https://public.tableau.com/en-us/gallery/impact-proposed-policy-changes-snap?tab=featured&type=featured

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

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.

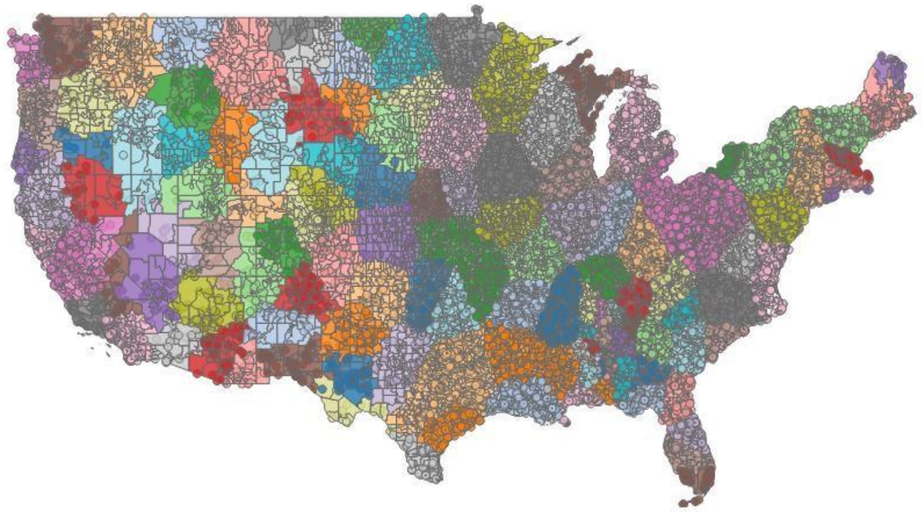


Use for Storytelling - Examples

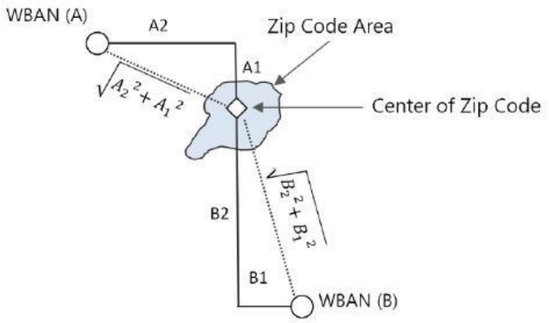
Tableau – Storytelling Example (Story Points)

Visualizing Weather Data in Tableau Software

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

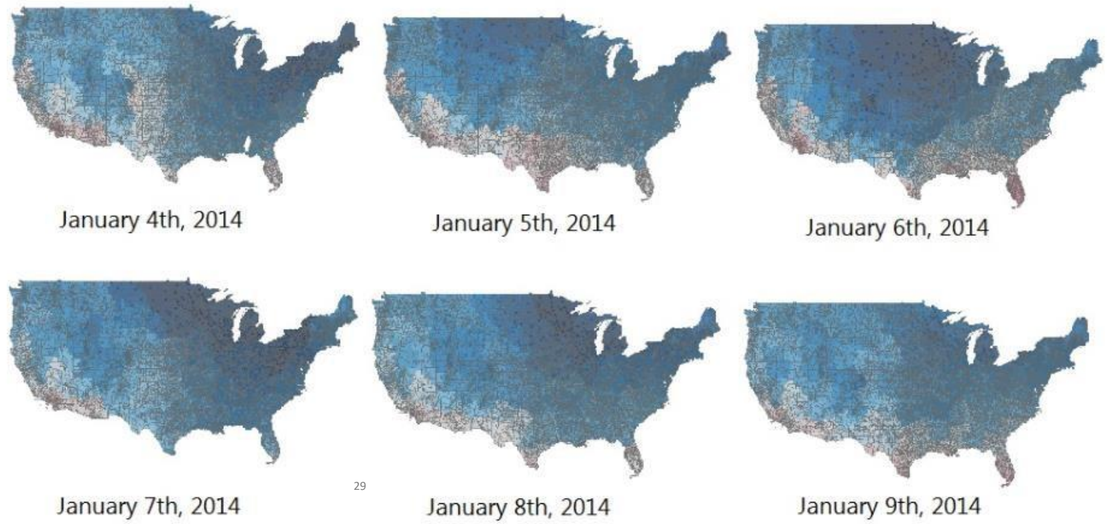


To map each ZIP code to the nearest WBAN, a simple pairwise comparison was used. For each ZIP code, a distance was calculated from the center of the ZIP code to the location of each WBAN utilizing the Latitude and Longitude data from the dataset. Each ZIP code was assigned to one WBAN sensor based on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



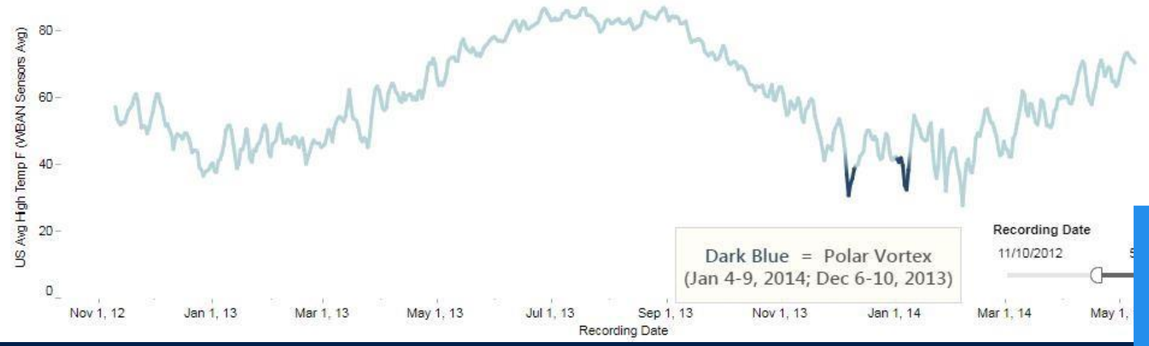
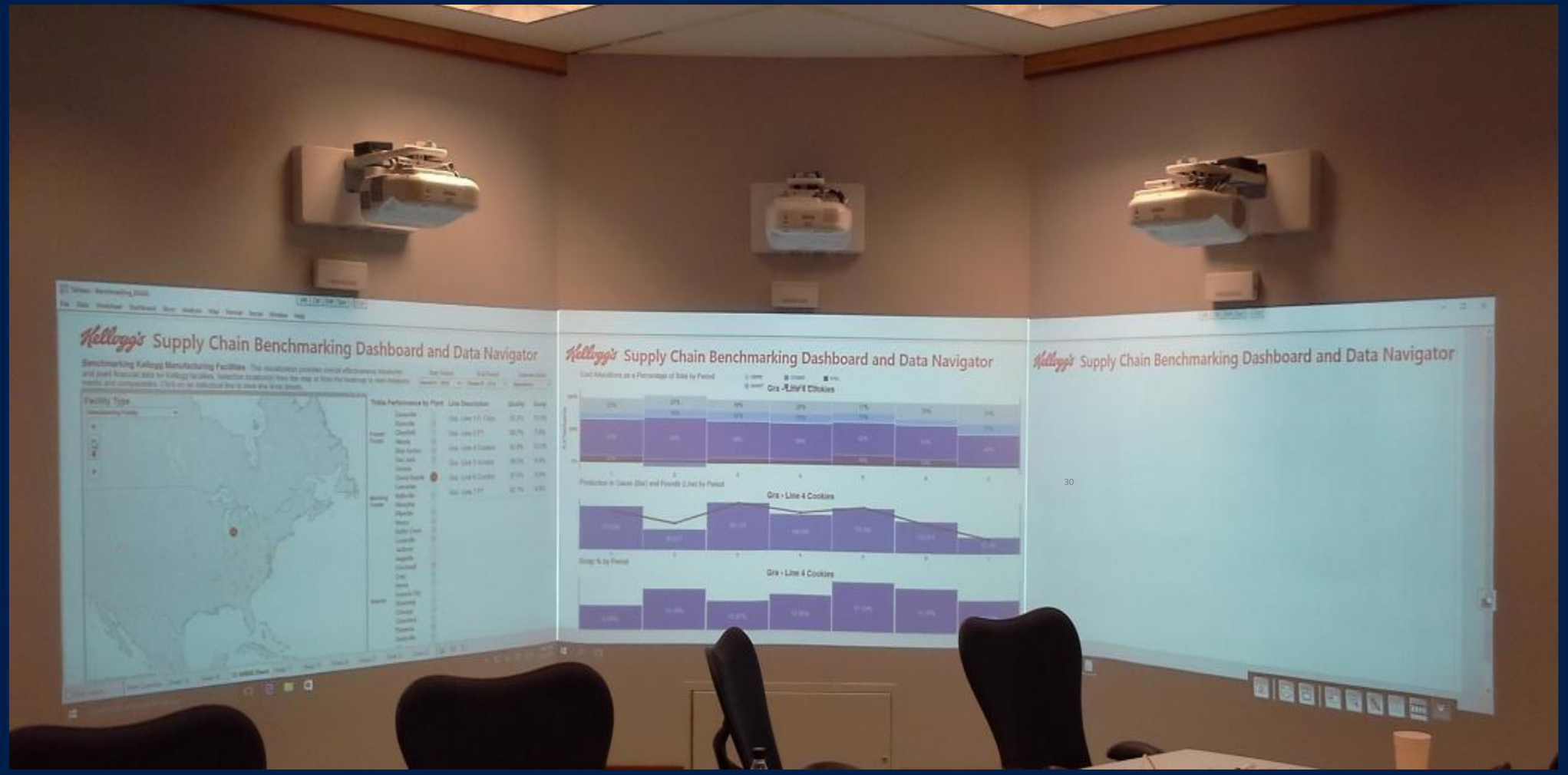


Tableau – Storytelling Example (K-MAX)



Advanced Features - Examples

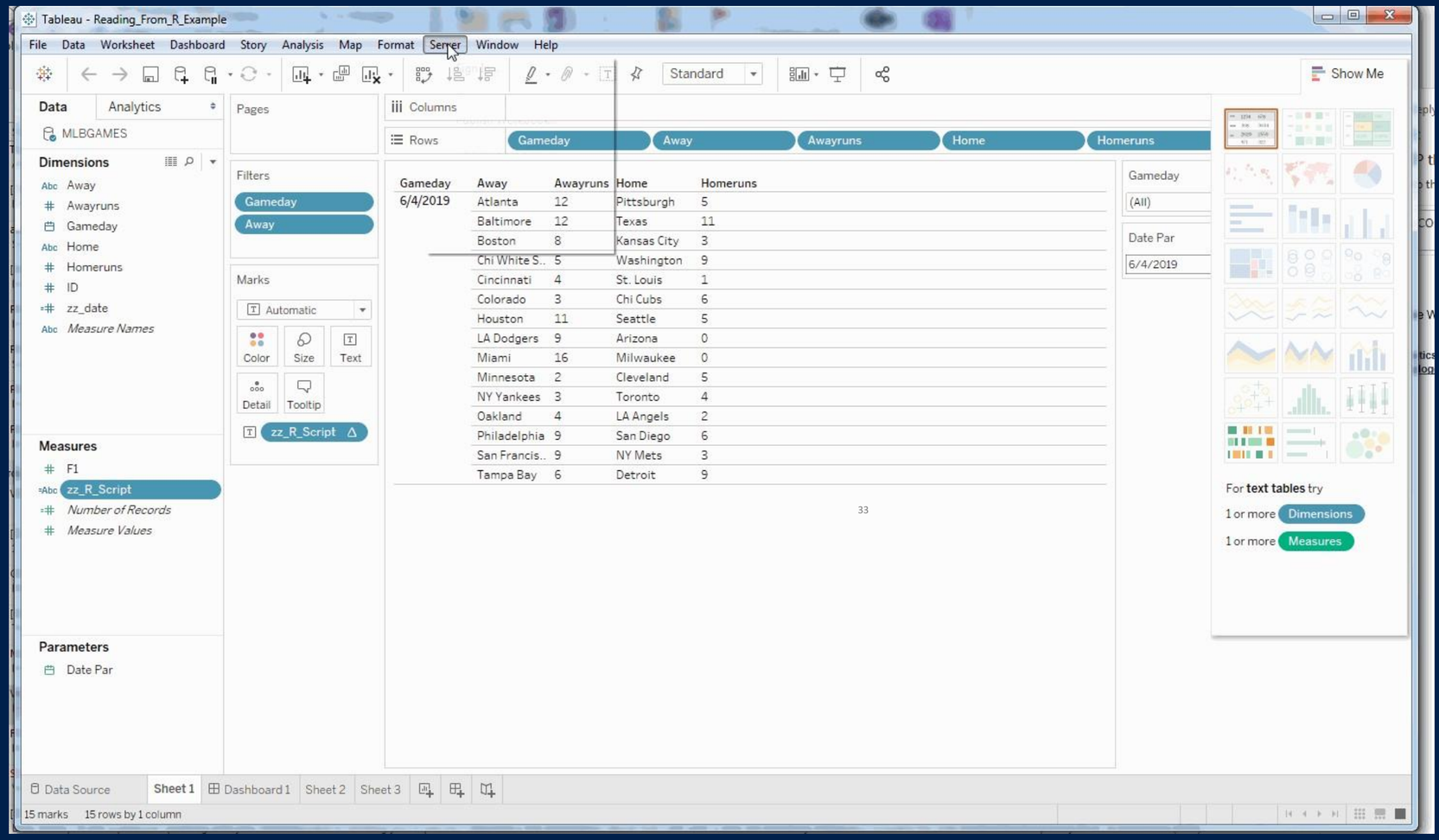
Advanced Features – Connecting Tableau to “R”

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

```
Console ~/

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\DOCUME~1\R\R-36~1.0\library\Rserve\libs\x64\Rserve.exe"
```

Advanced Features – Connecting Tableau to “R”

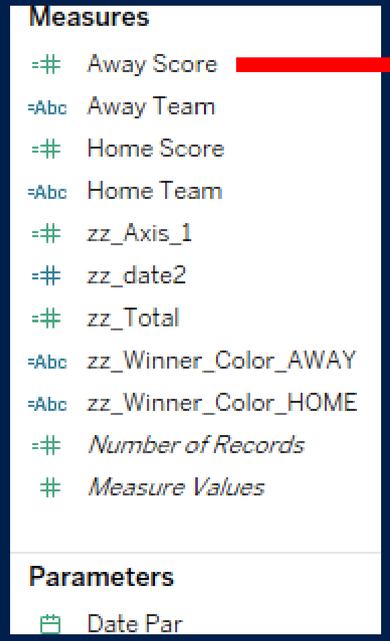
- Step #2
  - Connect Tableau to your Rserve instance



Advanced Features – Connecting Tableau to “R”

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

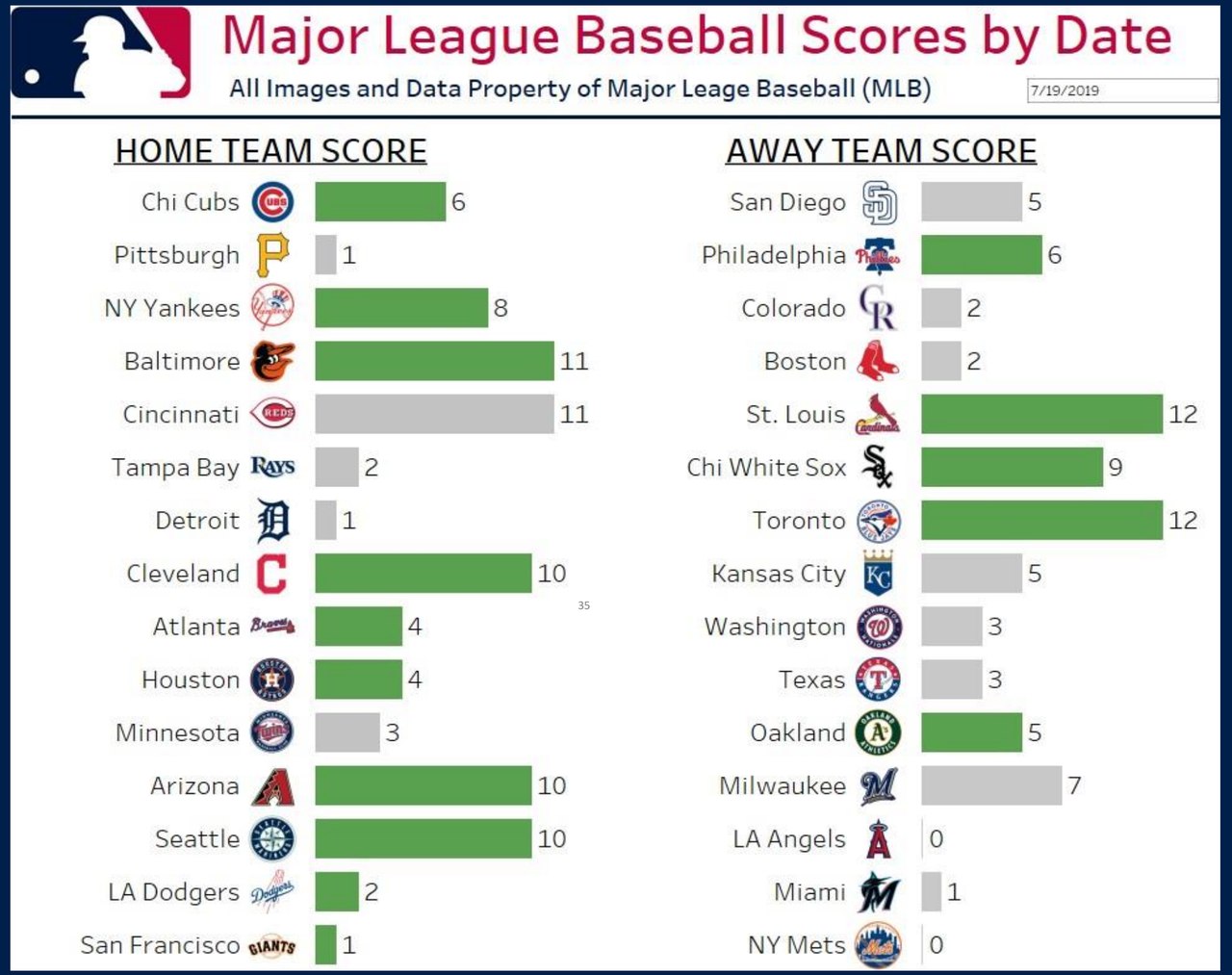
Note: This is also generally the same way to connect Tableau to Python in Anaconda – with a few small configuration differences.




```
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);
awayrunsdf <- as.data.frame(awayruns);
awayrunsdf$ID <- seq.int(nrow(awayrunsdf));
toString(awayrunsdf[.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.



Institut Informatika & Bisnis DARMAJAYA
Yayasan Alfian Husin

SAINS DATA
Institut Informatika & Bisnis Darmajaya

Kampus Merdeka INDONESIA JAYA

MERDEKA BELAJAR

THANK YOU!!

DATA SCIENCE DARMAJAYA “YOUR BEST FUTURE IN DATA”