Visualization and Dashboarding

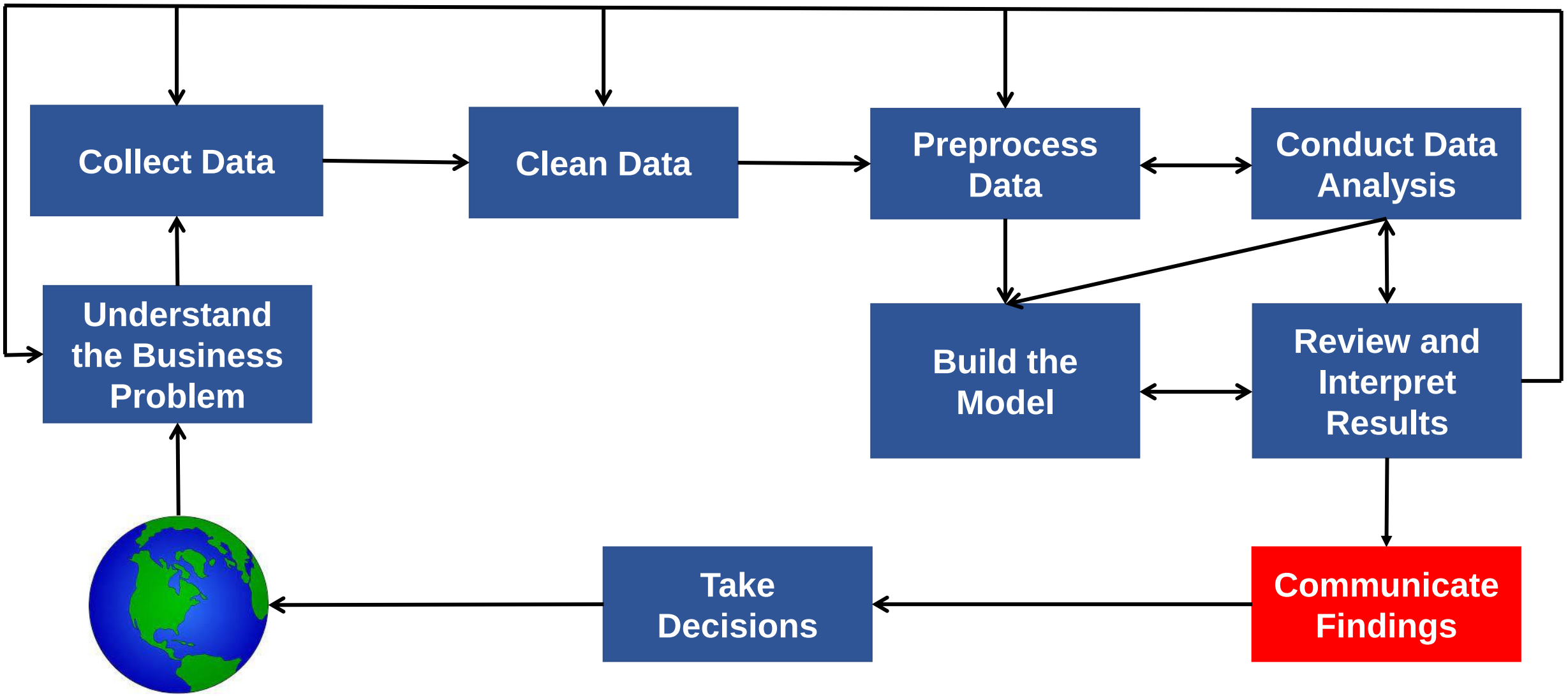
Dr. Sandareka Wickramanayake

sandarekaw@cse.mrt.ac.lk

Recommended Reading

- The Big Book of Dashboarding – Steve Wexler, Jeffrey Shaffer, Andy Cotgreave
  - Chapter 1
- Visualization Analysis & Design - Tamara Munzner
  - Chapter 1,2 and 5
- Tableau Viz Design Workshop - https://www.tableau.com/community/toolkit/viz-design-workshop
- Toward a Deeper Understanding of the Role of Interaction in Information Visualization - J. S. Yi, Y. a. Kang, J. Stasko and J. A. Jacko
- Fundamentals of Data Visualization - Claus O. Wilke https://clauswilke.com/dataviz/index.html

Data Science Process



In Today's Lecture

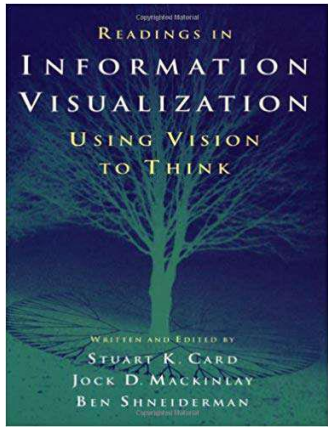
- Visualization
- Dashboarding
- Interactivity

In Today's Lecture

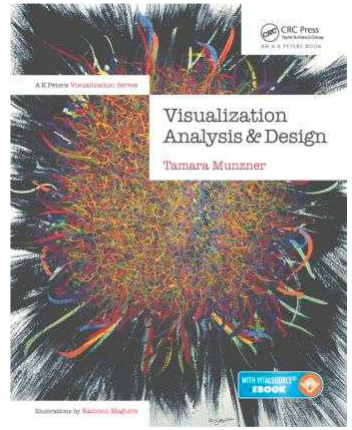
- **Visualization**
- Dashboarding
- Interactivity

Visualization – What is Visualization?

- Converting raw data to a form that is viewable and understandable to humans. - *Oxford English Dictionary*



- The use of computer-supported, interactive, visual representations of abstract data to amplify cognition. - *Stuart Card*



- Visual representations of datasets designed to help people carry out tasks more effectively - *Tamara Munzner*

Visualization – What is Visualization?

- The purpose of visualization is insight, not pictures
- A good visualization
  - Makes data accessible
  - Combines strengths of humans and computers
  - Enables insights
  - Communicates

Visualization – Why Visualize?

- Human has limited memory and cognition
  - E.g., Change blindness: the door study

Visualization – Why Visualize?

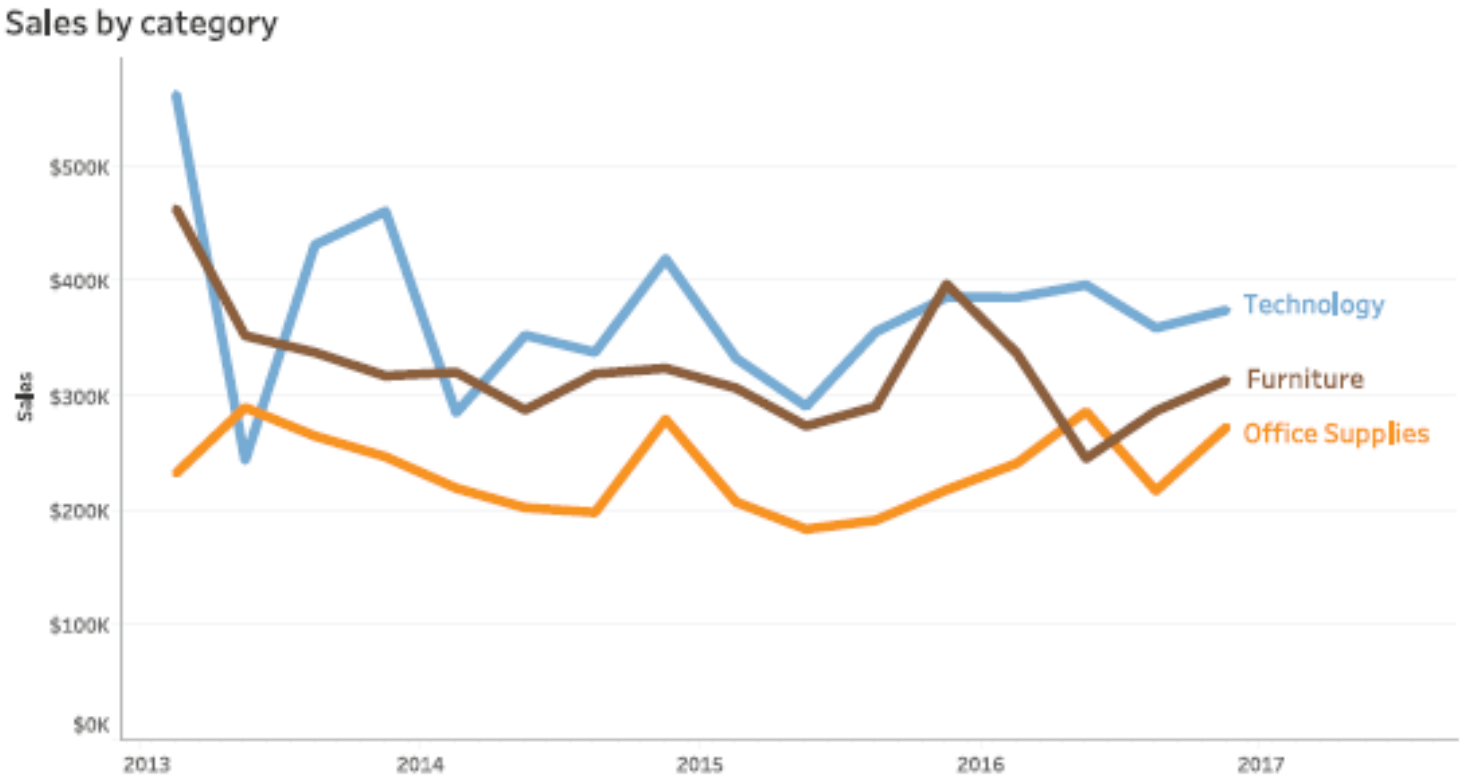
- Visualization helps human memory.
  - E.g., The following table shows sales numbers for three categories, by quarter, over a four-year period. What trends can you see?

Category	2013 Q1	2013 Q2	2013 Q3	2013 Q4	2014 Q1	2014 Q2	2014 Q3	2014 Q4
Furniture	$463,988	$352,779	$338,169	$317,735	$320,875	$287,934	$319,537	$324,319
Office Supplies	$232,558	$290,055	$265,083	$246,946	$219,514	$202,412	$198,268	$279,679
Technology	$563,866	$244,045	$432,299	$461,616	$285,527	$353,237	$338,360	$420,018
Category	**2015 Q1**	**2015 Q2**	**2015 Q3**	**2015 Q4**	**2016 Q1**	**2016 Q2**	**2016 Q3**	**2016 Q4**
Furniture	$307,028	$273,836	$290,886	$397,912	$337,299	$245,445	$286,972	$313,878
Office Supplies	$207,363	$183,631	$191,405	$217,950	$241,281	$286,548	$217,198	$272,870
Technology	$333,002	$291,116	$356,243	$386,445	$386,387	$397,201	$359,656	$375,229

Source – The Big Book of Dashboards

Visualization – Why Visualize?

- Visualization helps human memory.
  - E.g., The following table shows sales numbers for three categories, by quarter, over a four-year period. What trends can you see?



Source – The Big Book of Dashboards

Visualization – Why Visualize?

- Visualization indicates trends and patterns.
  - E.g., Consider the table below. What does each group tell you?

Group A		Group B		Group C		Group D	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89

Anscombe's quartet, 1973

Source – The Big Book of Dashboards

Summary Statistics:

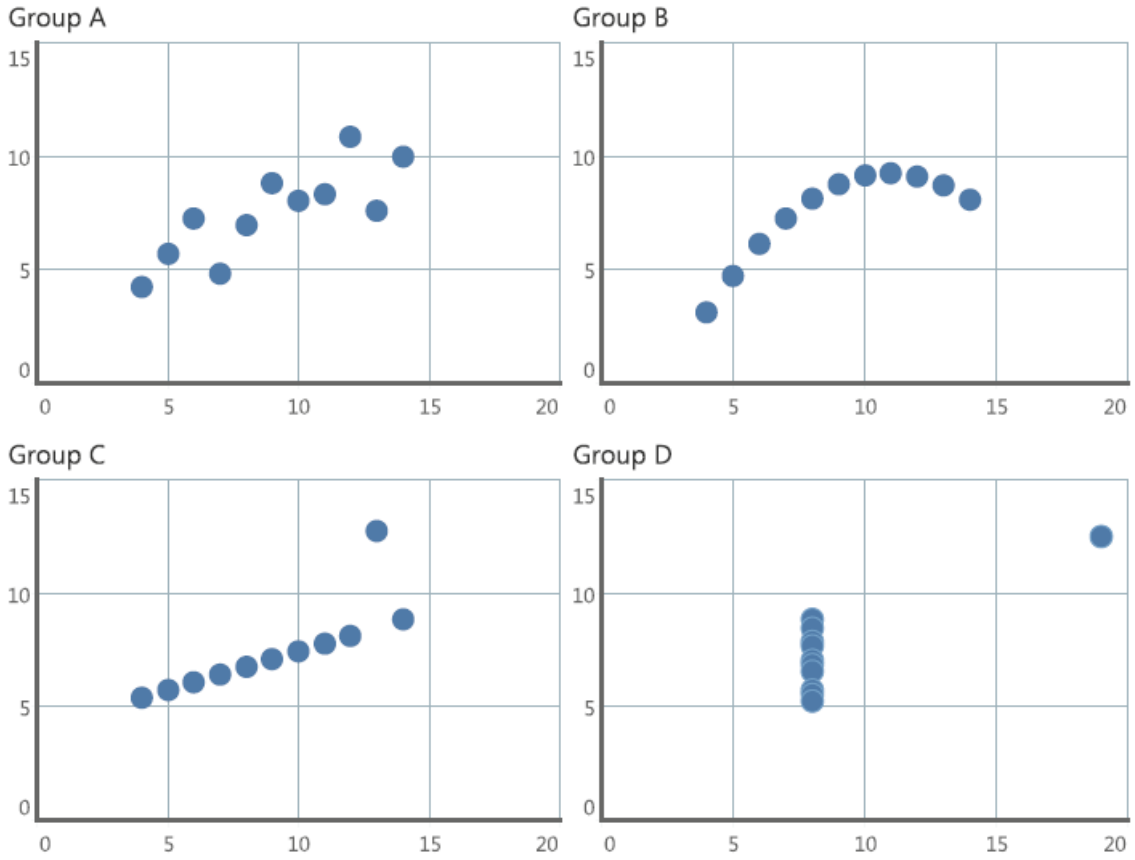
$\mu_x = 9 \quad \sigma_x = 3.31$

$\mu_y = 7.5 \quad \sigma_y = 2.03$

Visualization – Why Visualize?

- Visualization indicates trends and patterns.
  - E.g., Consider the table below. What does each group tell you?

Group A		Group B		Group C		Group D	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89



Anscombe's quartet, 1973

Source – The Big Book of Dashboards

Visualization – Why Visualize?

- Visualization provides **insights**.
  - Facts about a given problem that we were not aware of.
  - Discovering the unknown.
  - Answers to concrete questions about a given problem.

Visualization – Types of Visualizations

- Exploratory - #Analysis
  - Appropriate when you have a whole bunch of data and you're not sure what's in it.
  - When you need to get a sense of what's inside your data set, translating it into a visual medium can help you quickly identify its features, including interesting curves, lines, trends, or anomalous outliers.
- Explanatory - #Communication
  - Appropriate when you already know what the data has to say, and you are trying to tell that story to somebody else.
  - It could be the head of your department, a grant committee, or the general public.

Visualization – How Do We Visualize Data ?

- E.g., How many 9s are there in Figure 1?

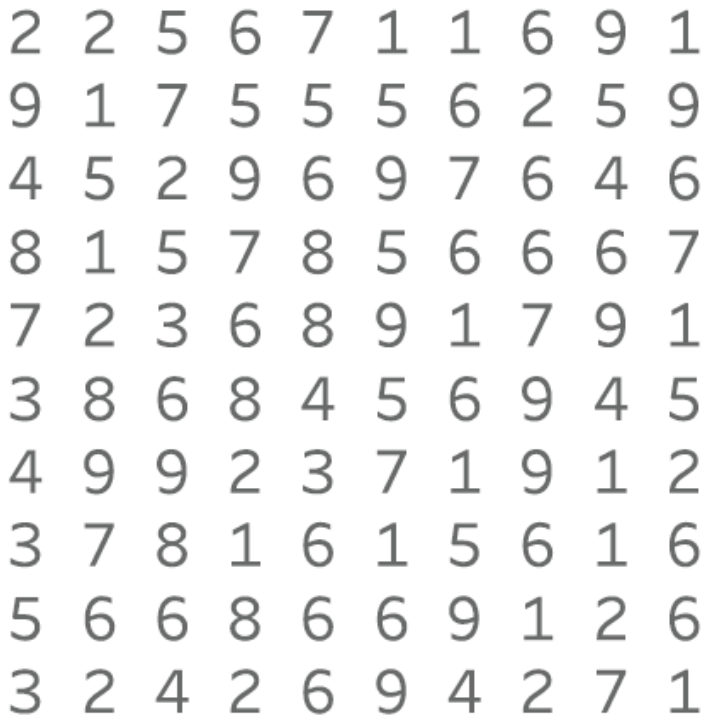


Figure 1

Source – The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- E.g., How many 9s are there in Figure 1?
- Color (hue) difference

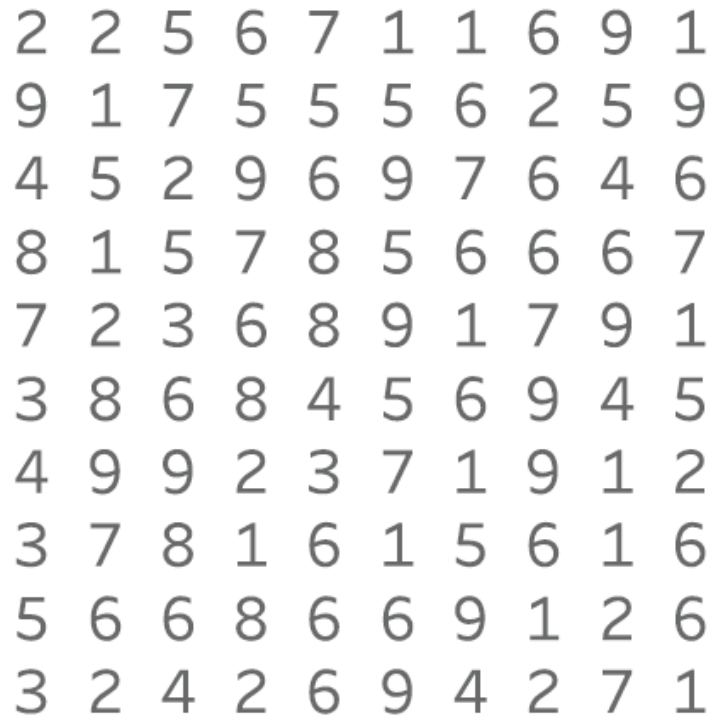


Figure 1

Source – The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- E.g., How many 9s are there in Figure 1?
- Size

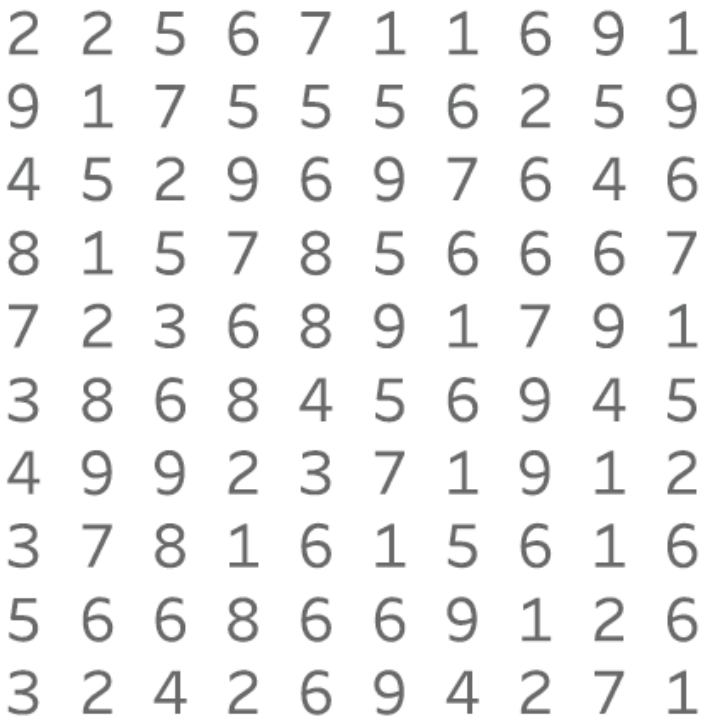
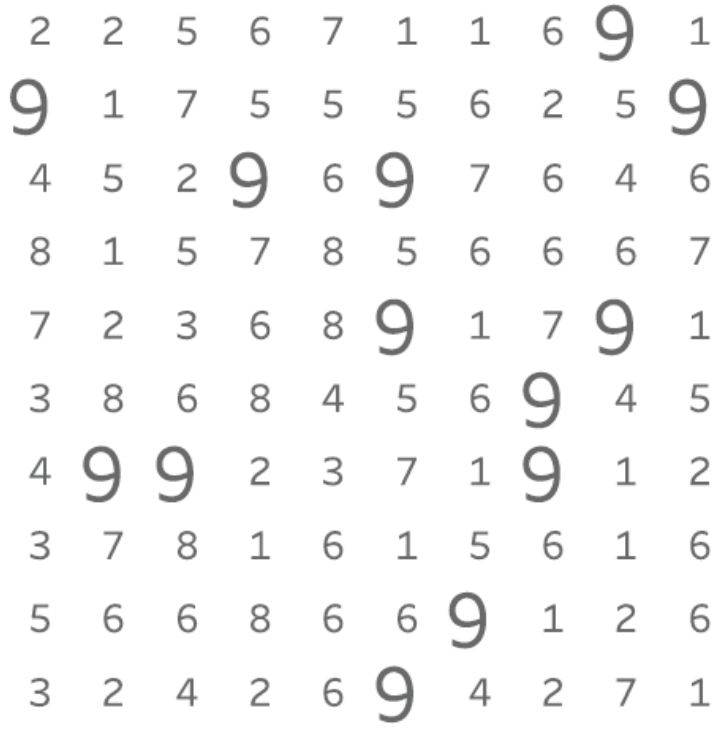


Figure 1



Source – The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- E.g., Count the frequency of each digit in Figure 1?
- Color?

2	2	5	6	7	1	1	6	9	1
9	1	7	5	5	5	6	2	5	9
4	5	2	9	6	9	7	6	4	6
8	1	5	7	8	5	6	6	6	7
7	2	3	6	8	9	1	7	9	1
3	8	6	8	4	5	6	9	4	5
4	9	9	2	3	7	1	9	1	2
3	7	8	1	6	1	5	6	1	6
5	6	6	8	6	6	9	1	2	6
3	2	4	2	6	9	4	2	7	1

Figure 1

Not going to work!

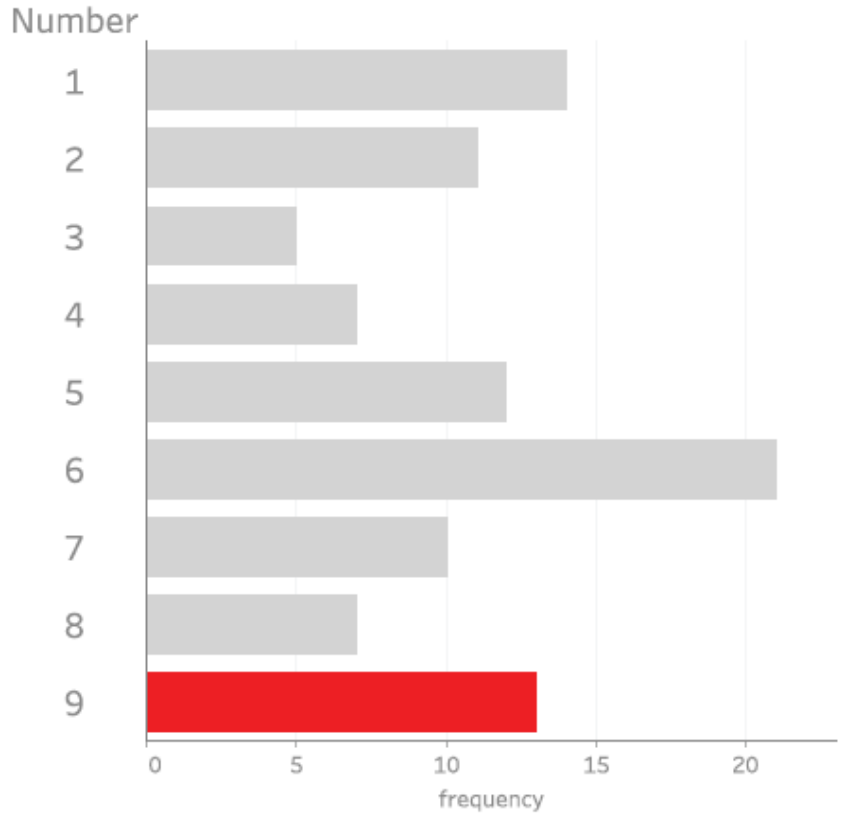
Source – The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- E.g., Count the frequency of each digit in Figure 1?
- Length

2 2 5 6 7 1 1 6 9 1
9 1 7 5 5 5 6 2 5 9
4 5 2 9 6 9 7 6 4 6
8 1 5 7 8 5 6 6 6 7
7 2 3 6 8 9 1 7 9 1
3 8 6 8 4 5 6 9 4 5
4 9 9 2 3 7 1 9 1 2
3 7 8 1 6 1 5 6 1 6
5 6 6 8 6 6 9 1 2 6
3 2 4 2 6 9 4 2 7 1

Figure 1



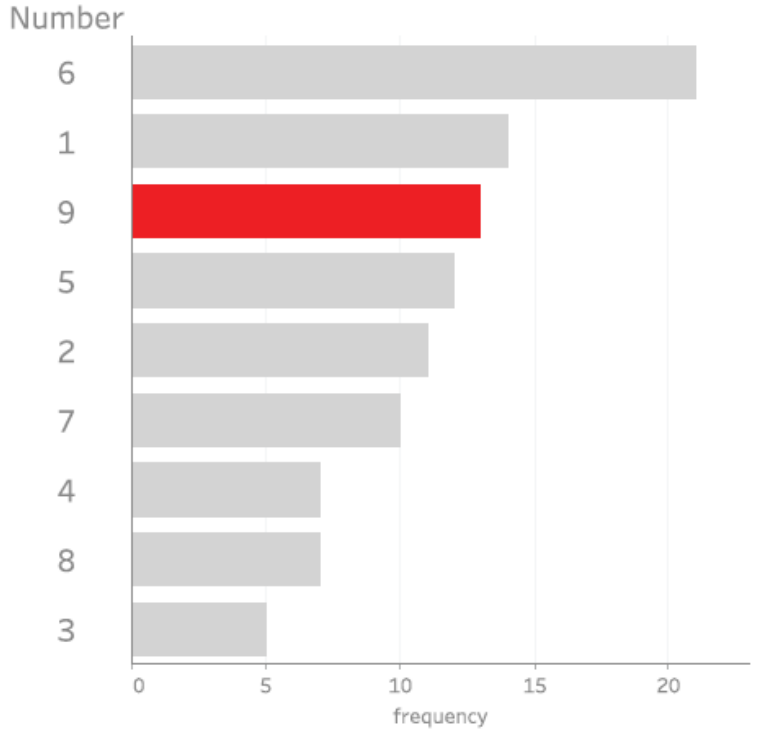
Source – The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- E.g., Count the frequency of each digit in Figure 1?
- Length and Position

2	2	5	6	7	1	1	6	9	1
9	1	7	5	5	5	6	2	5	9
4	5	2	9	6	9	7	6	4	6
8	1	5	7	8	5	6	6	6	7
7	2	3	6	8	9	1	7	9	1
3	8	6	8	4	5	6	9	4	5
4	9	9	2	3	7	1	9	1	2
3	7	8	1	6	1	5	6	1	6
5	6	6	8	6	6	9	1	2	6
3	2	4	2	6	9	4	2	7	1

Figure 1



Difficult-to-read table → Easy-to-read bar chart

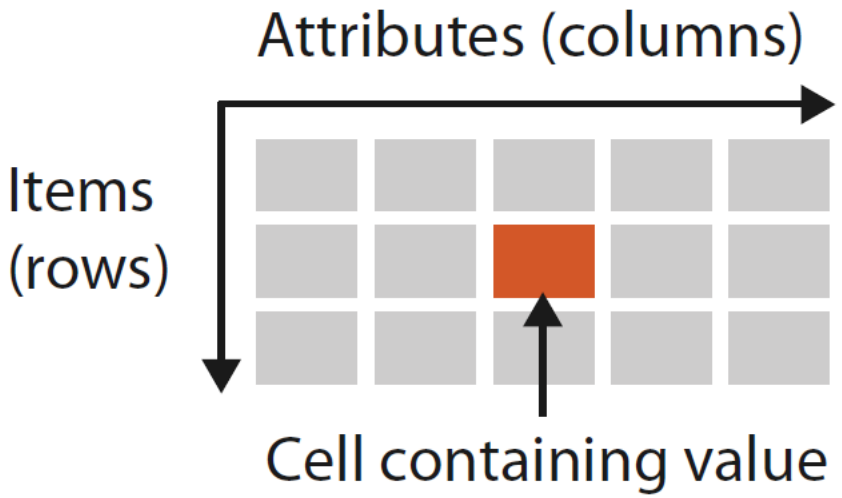
Source – The Big Book of Dashboards

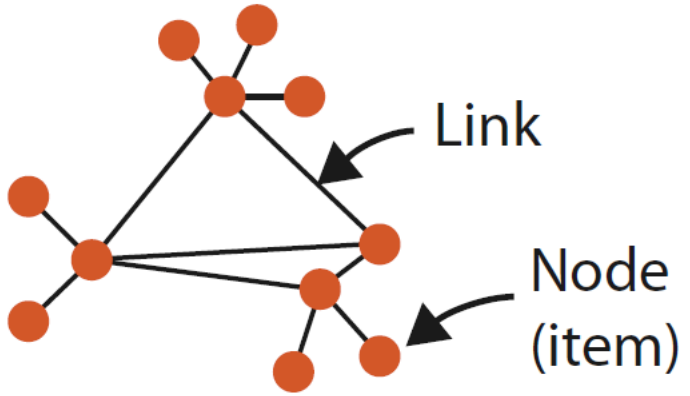
Visualization – Know Your Data

- You should know the semantics and types!
- Semantics – Real-world meaning.
  - Java: Language, coffee, or a city?
  - 10: a day of the month, or an age, or a measurement of height, or a unique code or a position in sports?
- Types – Structural or mathematical interpretation
  - What kind of thing is it: an item, an attribute, a link
  - How are these data types combined into a larger structure: a table, a tree, a network
  - What kinds of mathematical operations are meaningful for it?

Visualization – Know Your Data

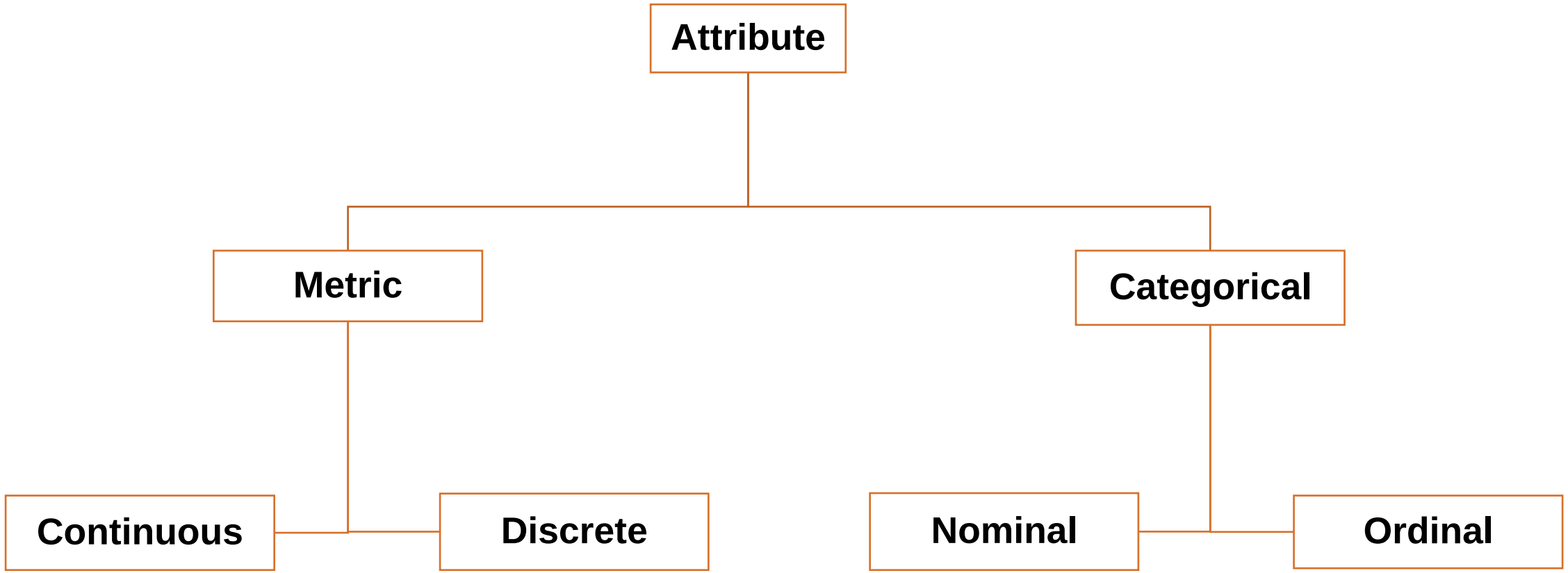
- Data Types
  - Items – An individual entity that is discrete, e.g., a row in a table
  - Attributes – Some specific properties that can be measured, observed, or logged, e.g., salary, price
  - Links – A relationship between items.





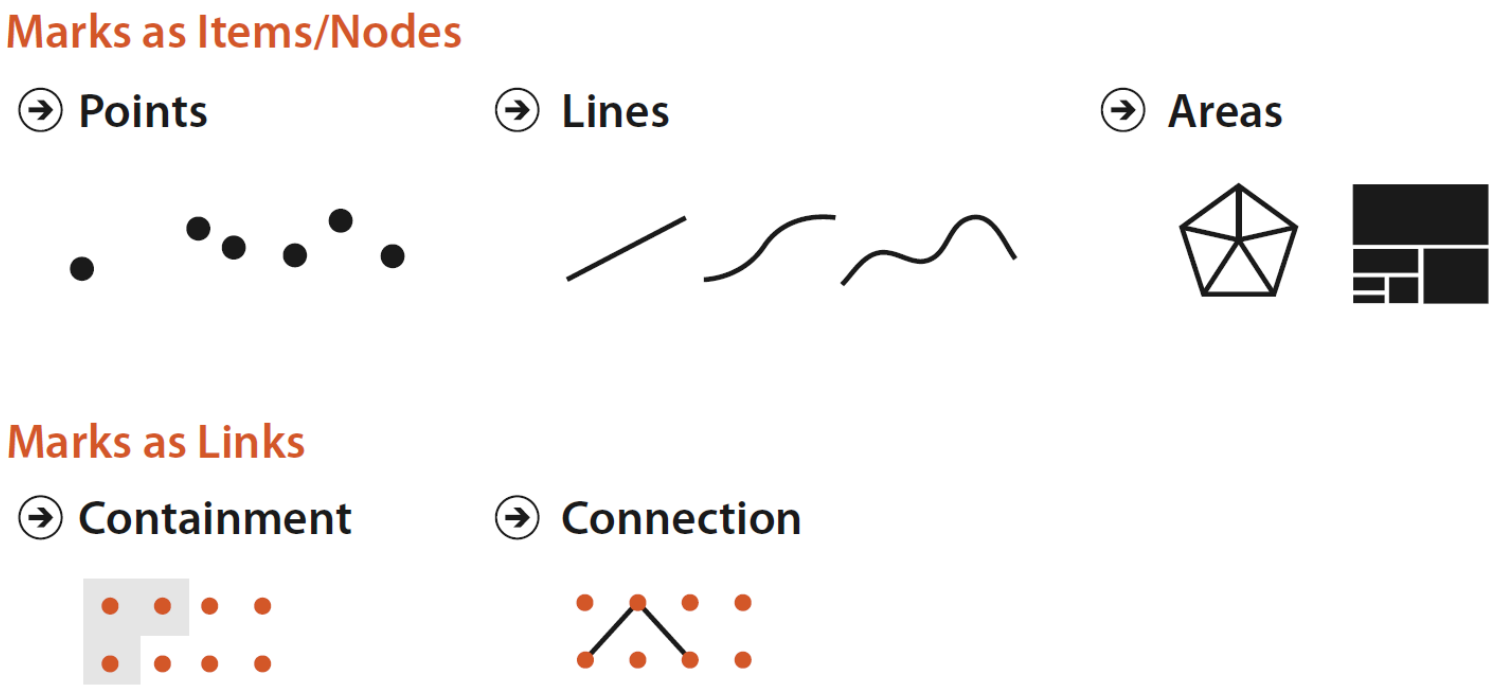
Visualization - Know Your Data - Attribute Types (Recap)

- Four fundamental attribute types.



Visualization – How Do We Visualize Data?

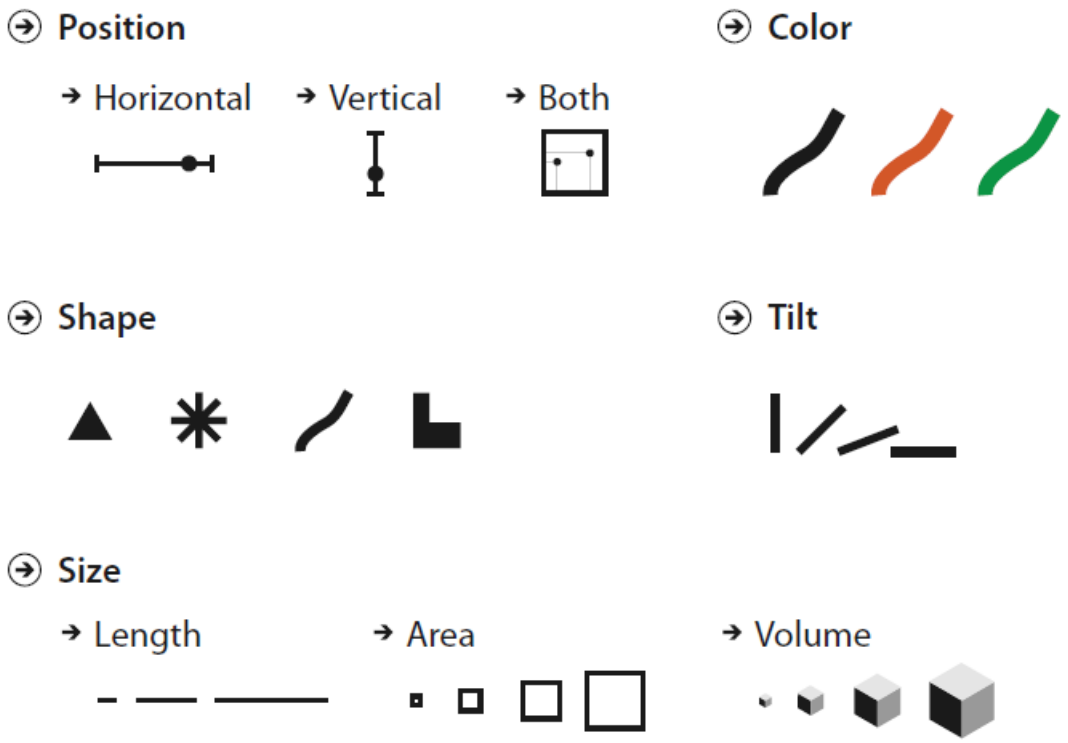
- Marks and Channels [1]
  - Marks - Basic geometric elements
    - Represents ITEMS and LINKS



[1] Visualization Analysis and Design, Tamara Munzer

Visualization – How Do We Visualize Data ?

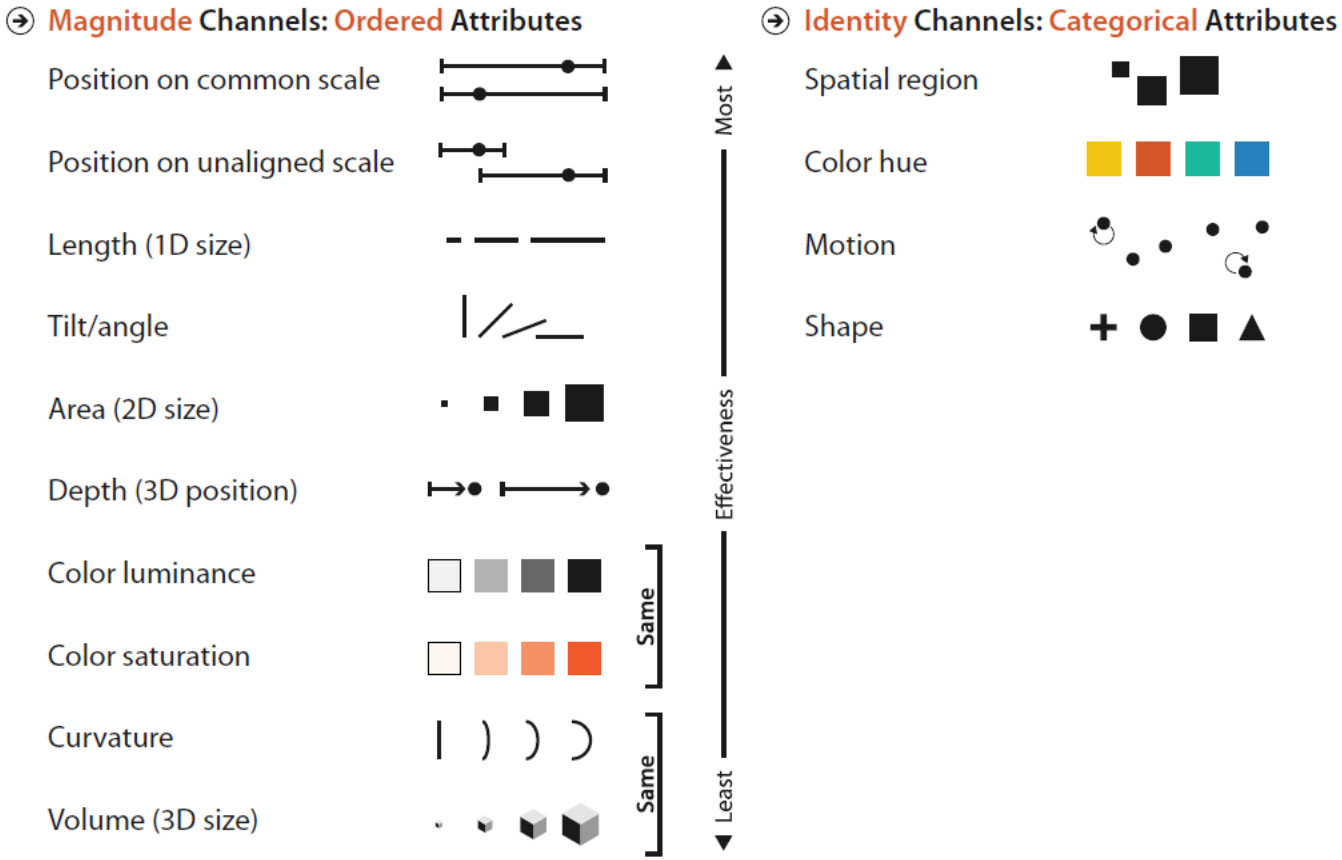
- Marks and Channels [1]
  - Channels – Control the appearance of MARKS.
    - AKA visual channels, visual variables, visual attributes, visual dimensions



[1] Visualization Analysis and Design, Tamara Munzer
[2] The Big Book of Dashboards

Visualization – How Do We Visualize Data?

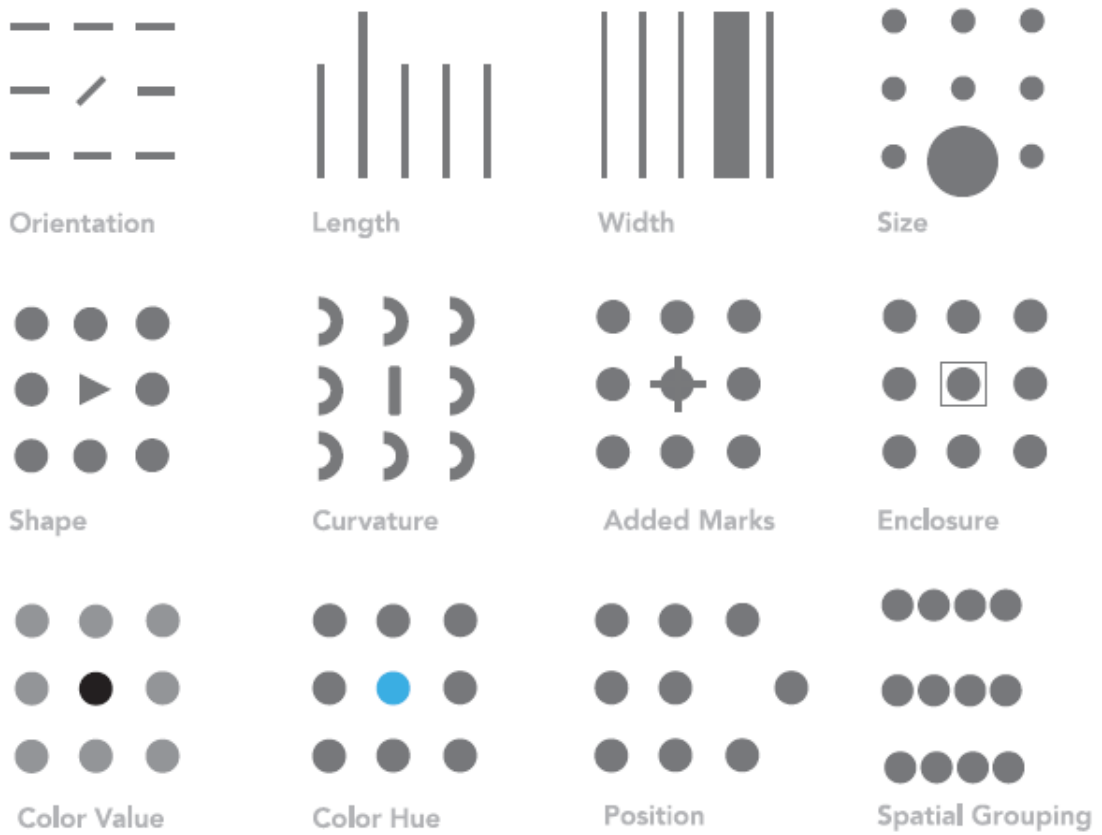
- Marks and Channels [1]
  - Channels – Control the appearance of MARKS.
    - AKA visual channels, visual variables, visual attributes, visual dimensions
    - Magnitude – HOW much something is
    - Identity – WHAT something is or WHERE it is



[1] Visualization Analysis and Design, Tamara Munzer
[2] The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- Marks and Channels [1]
  - Channels – Control the appearance of MARKS.
    - Pre-attentive Attributes [2] - Things that our brain processes in milliseconds, before we pay attention to everything else.



Example of using pre-attentive attributes - https://help.tableau.com/current/blueprint/en-us/bp_why_visual_analytics.htm#:~:text=Pre%2Dattentive%20attributes%20are%20information,processing%20parts%20of%20our%20brain.

[1] Visualization Analysis and Design, Tamara Munzer
[2] The Big Book of Dashboards

Visualization – How Do We Visualize Data ?

- Visual encoding – Choice of marks and channels for data representation.[1]

	Nominal	Ordinal	Metric
Position	Y	Y	Y
Length	N	Y	Y
Angle	N	Y	Y
Orientation	Y	Y	Y
Size	Y	Y	Y
Volume	N	Y	Y
Density	N	Y	Y
Saturation	N	Y	Y
Hue	Y	N	Y
Shape	Y	N	N
Texture	Y	N	N

[1] The Big Book of Dashboards

Visualization – Tips for Reading Visualizations (Graphs)

- Identify what information the chart is meant to convey.
- Identify information contained on each axis.
- Identify the range covered by each axis.
- Look for patterns or trends.
- Look for averages and/or exceptions.
- Look for bold or highlighted data.
- Look for citations for the data to see where it originated and ensure it is credible.

In Class Activity 1

Visualization - Data Types Mater

- E.g., Star Trek: The Next Generation.

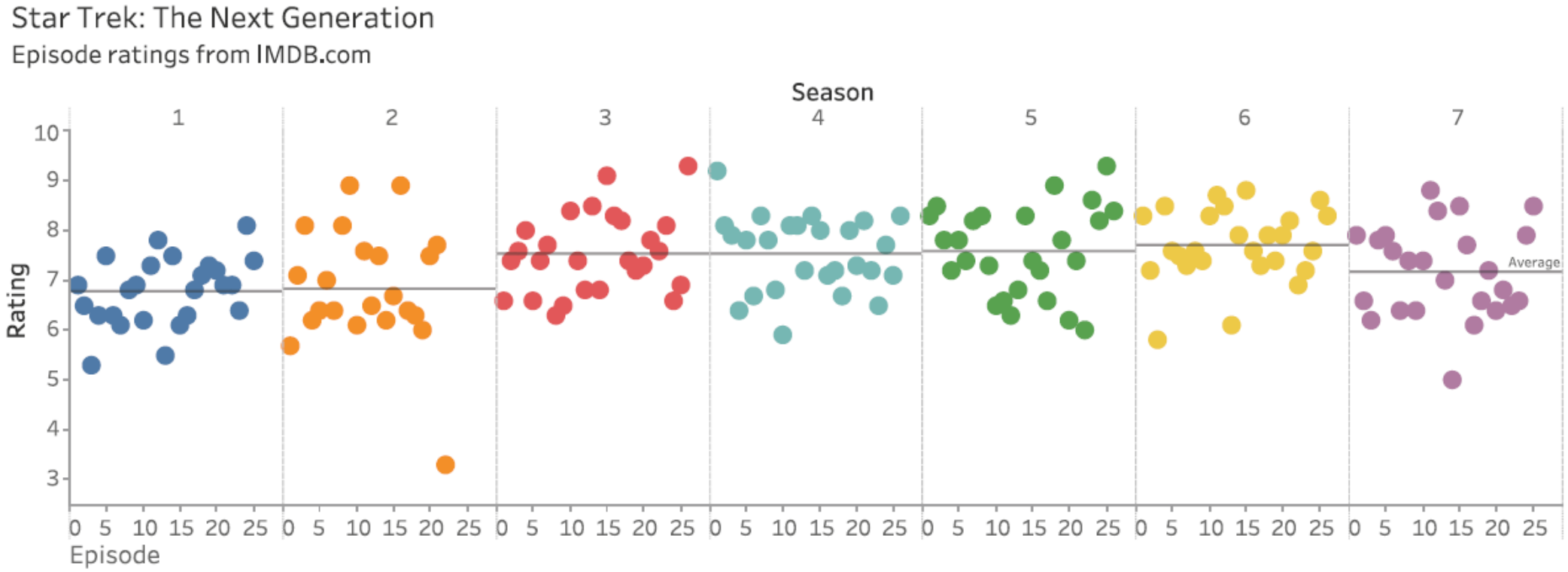


FIGURE 1.12 Every episode of *Star Trek: The Next Generation* rated.

Source: IMDB.com

Visualization - Data Types Mater

- E.g., Star Trek: The Next Generation.

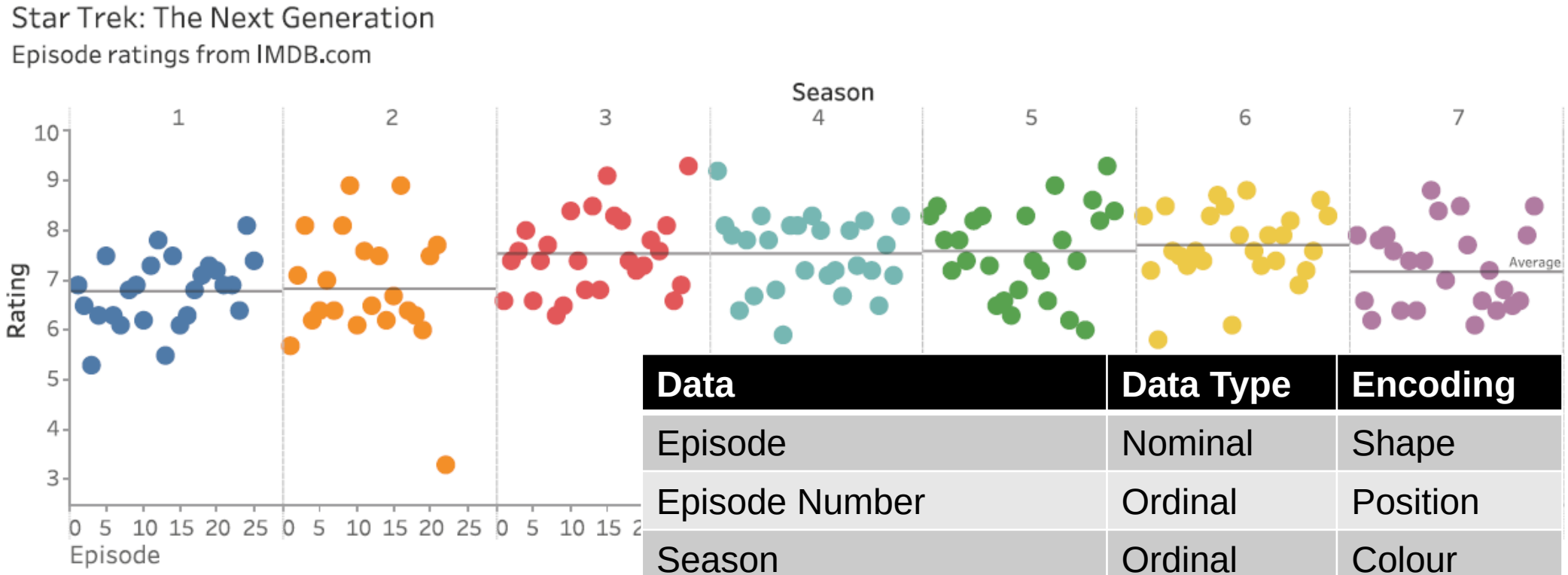


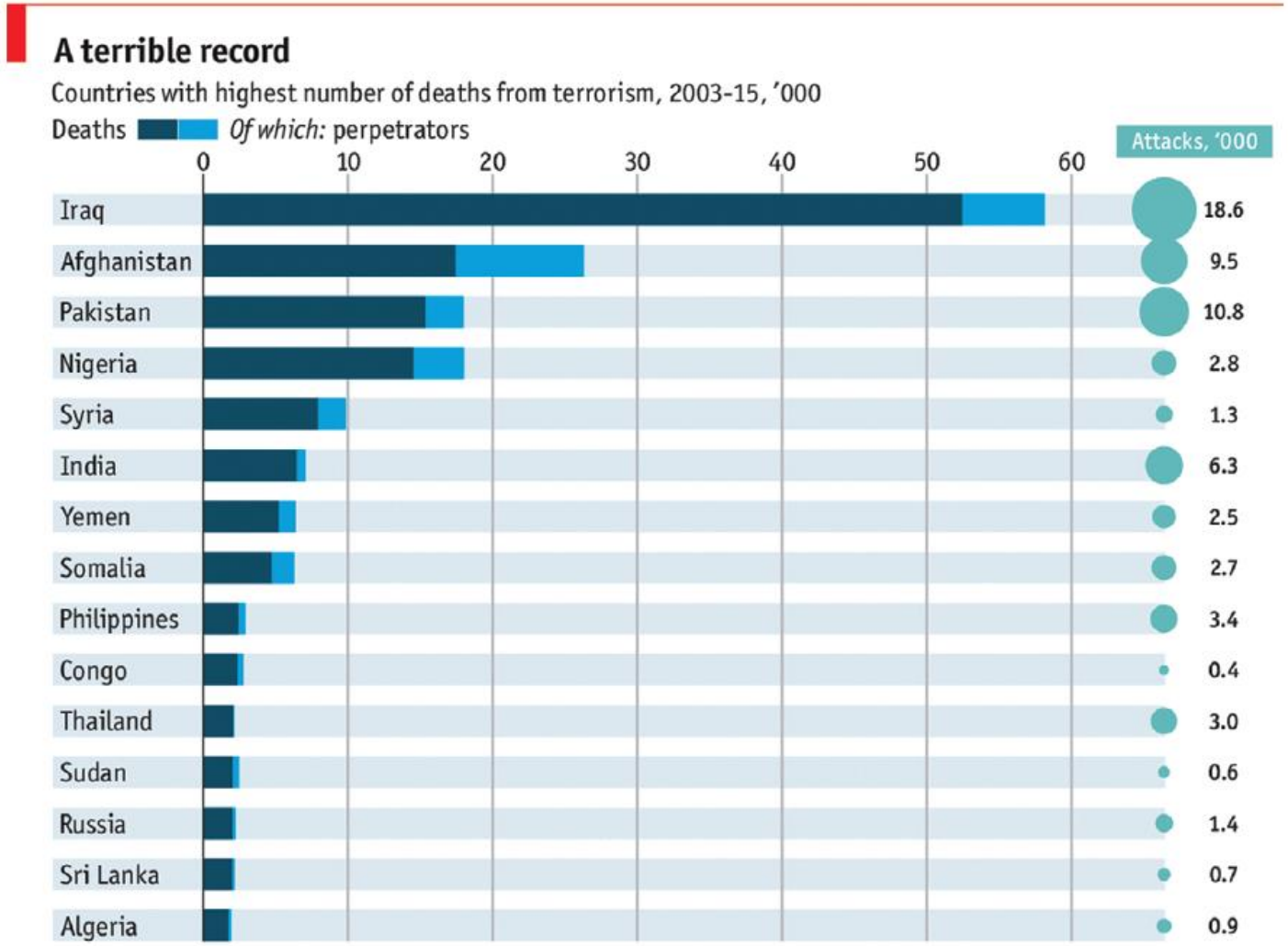
FIGURE 1.12 Every episode of *Star Trek: The N*

Source: IMDB.com

Data	Data Type	Encoding
Episode	Nominal	Shape
Episode Number	Ordinal	Position
Season	Ordinal	Colour
		Position
Average IMDB Rating	Continuous	Position
Average Season Rating	Continuous	Position

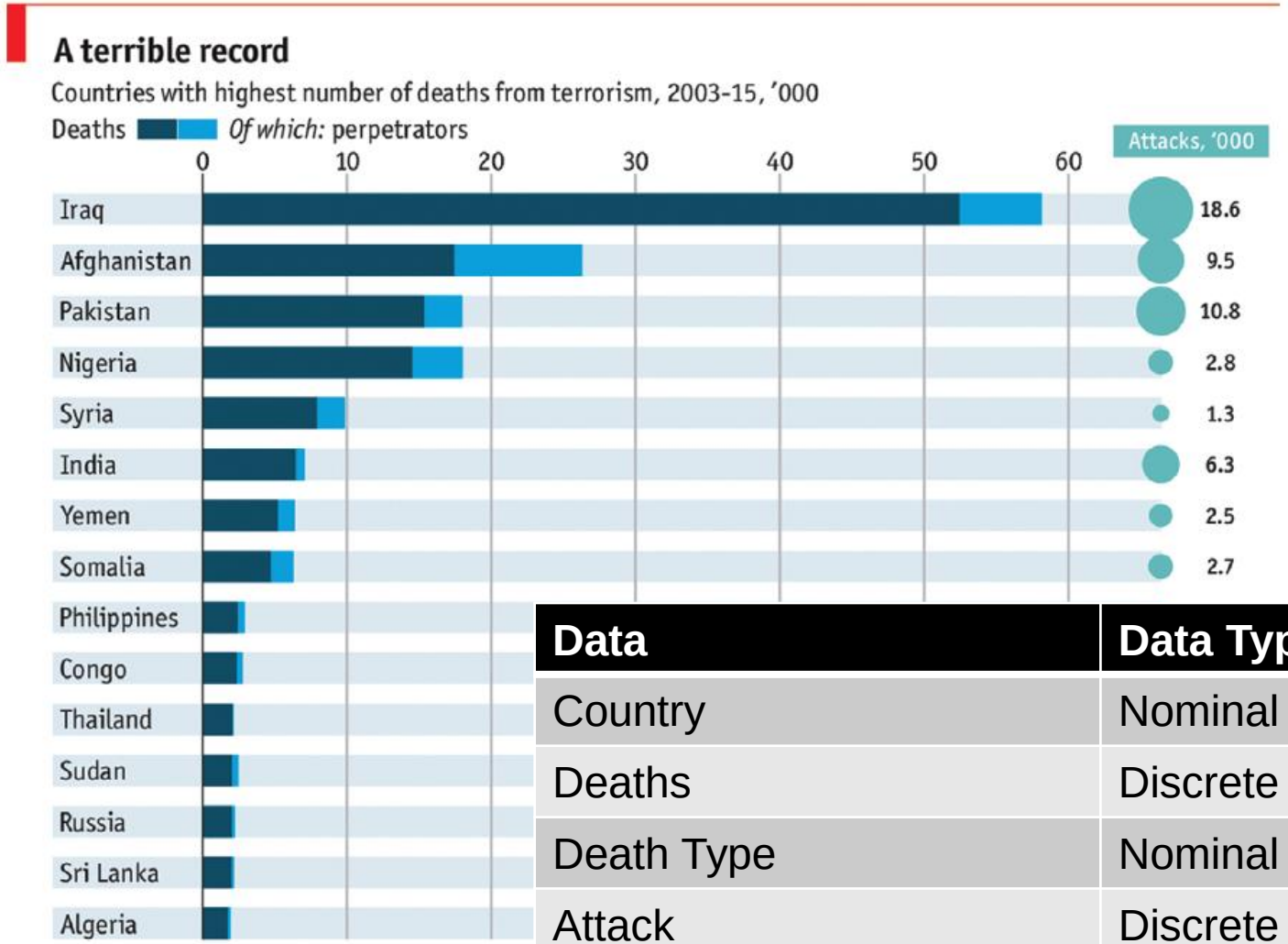
Visualization - Data Types Mater

- E.g., “A terrible record” from The Economist, July 2016.
  - Source: START, University of Maryland. The Economist, http://tabsoft.co/2agK3if



Visualization - Data Types Mater

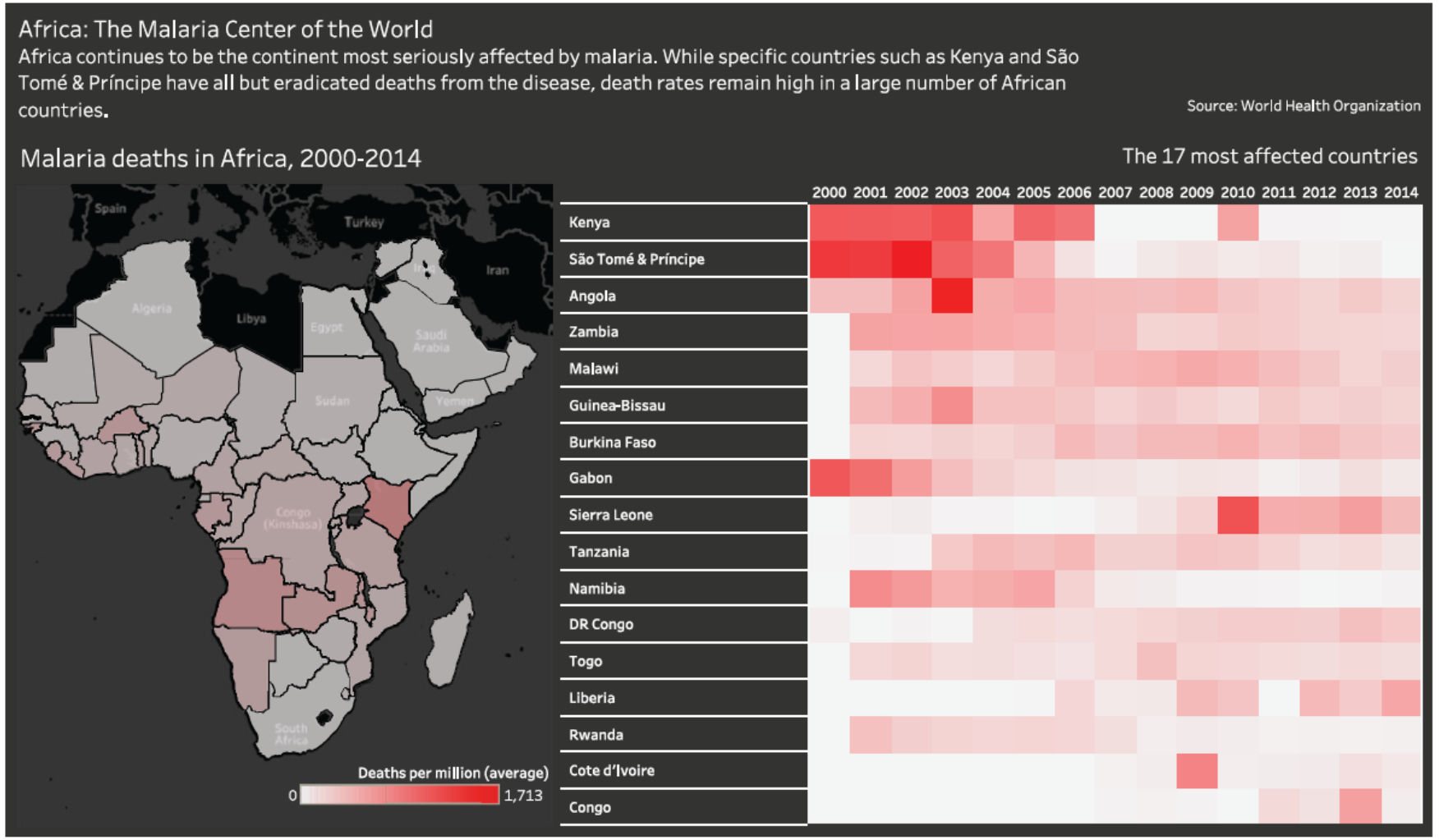
- E.g., “A terrible record” from The Economist, July 2016.
  - Source: START, University of Maryland. The Economist, http://tabsoft.co/2agK3if



Data	Data Type	Encoding
Country	Nominal	Position
Deaths	Discrete	Length
Death Type	Nominal	Colour
Attack	Discrete	Size

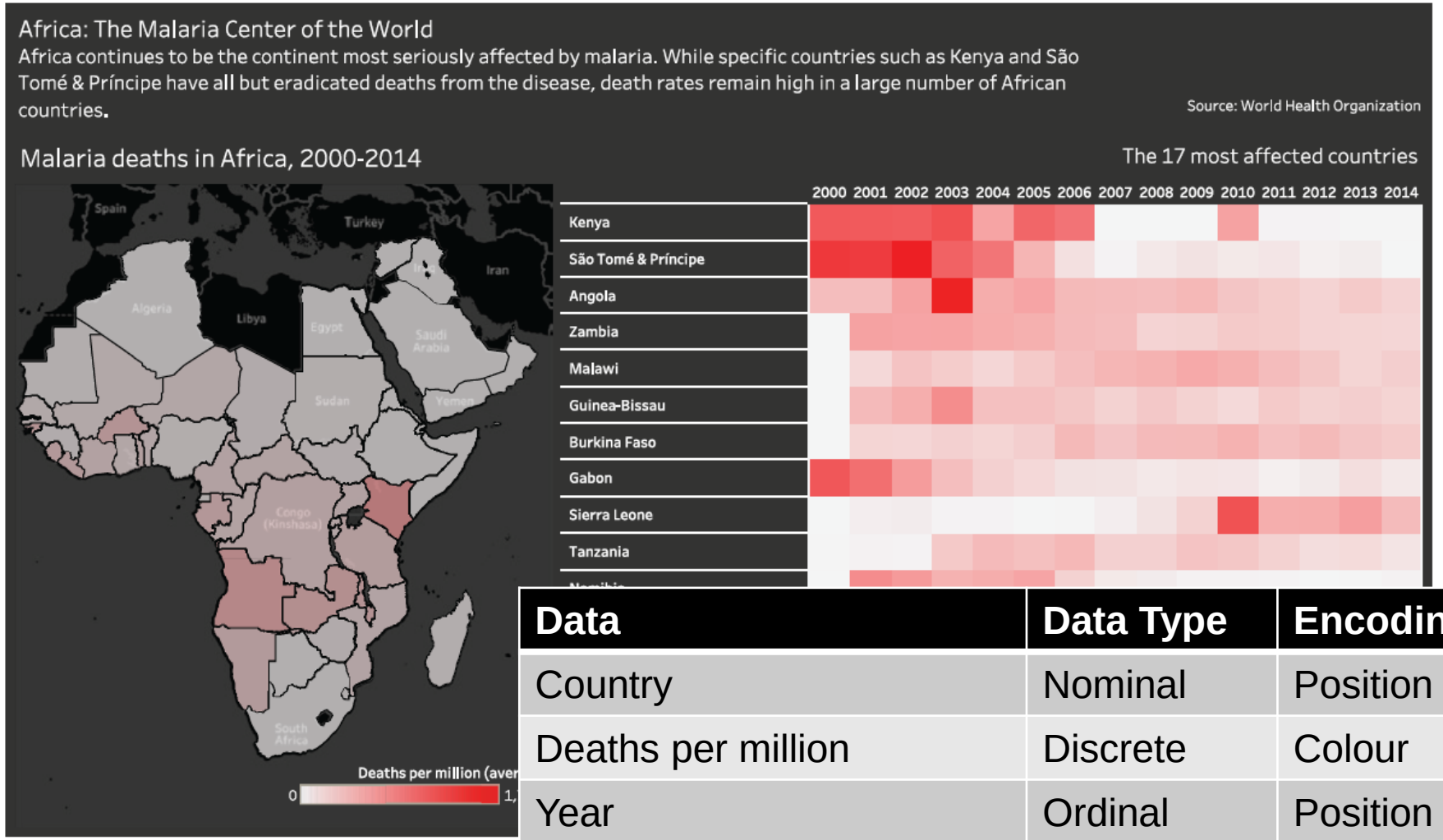
Visualization - Data Types Mater

- E.g., Deaths from malaria, 2000–2014.
  - Source: World Health Organization. Chart part of the Makeover Monday project



Visualization - Data Types Mater

- E.g., Deaths from malaria, 2000–2014.
  - Source: World Health Organization. Chart part of the Makeover Monday project

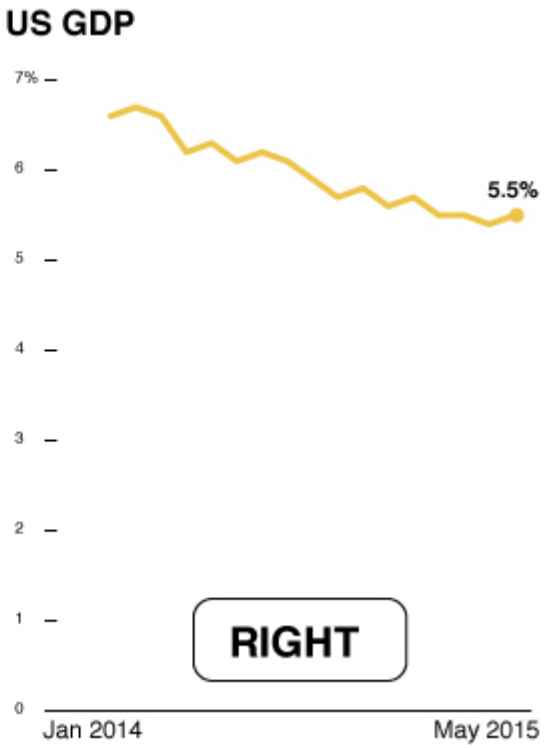


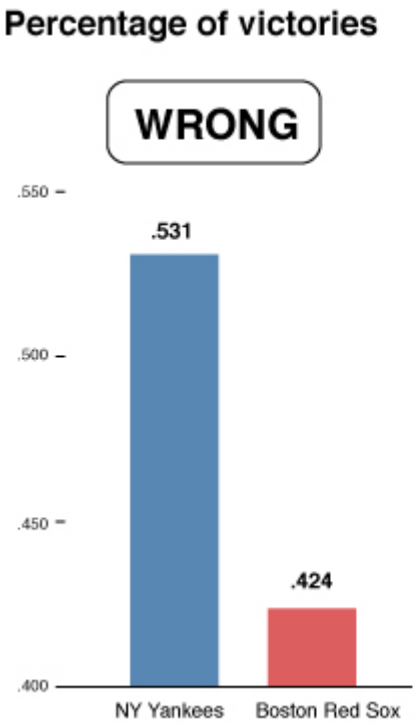
Data	Data Type	Encoding
Country	Nominal	Position
Deaths per million	Discrete	Colour
Year	Ordinal	Position

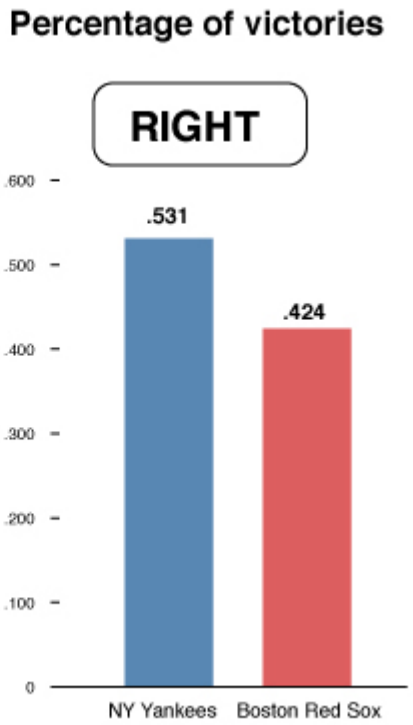
In Class Activity 2

Visualization – Common Mistakes

- Truncated (shortened) Y Axis ("broken scale“)

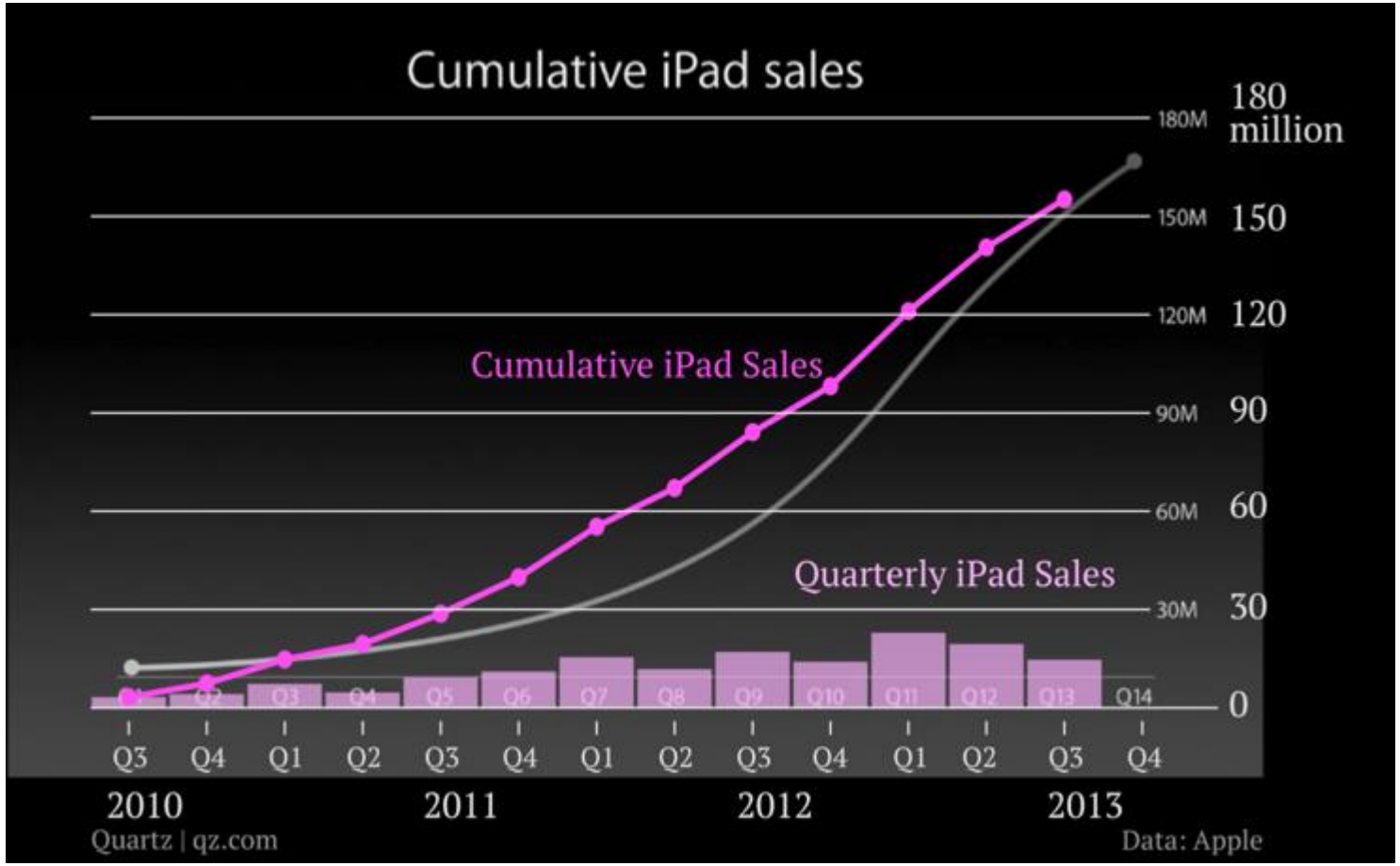






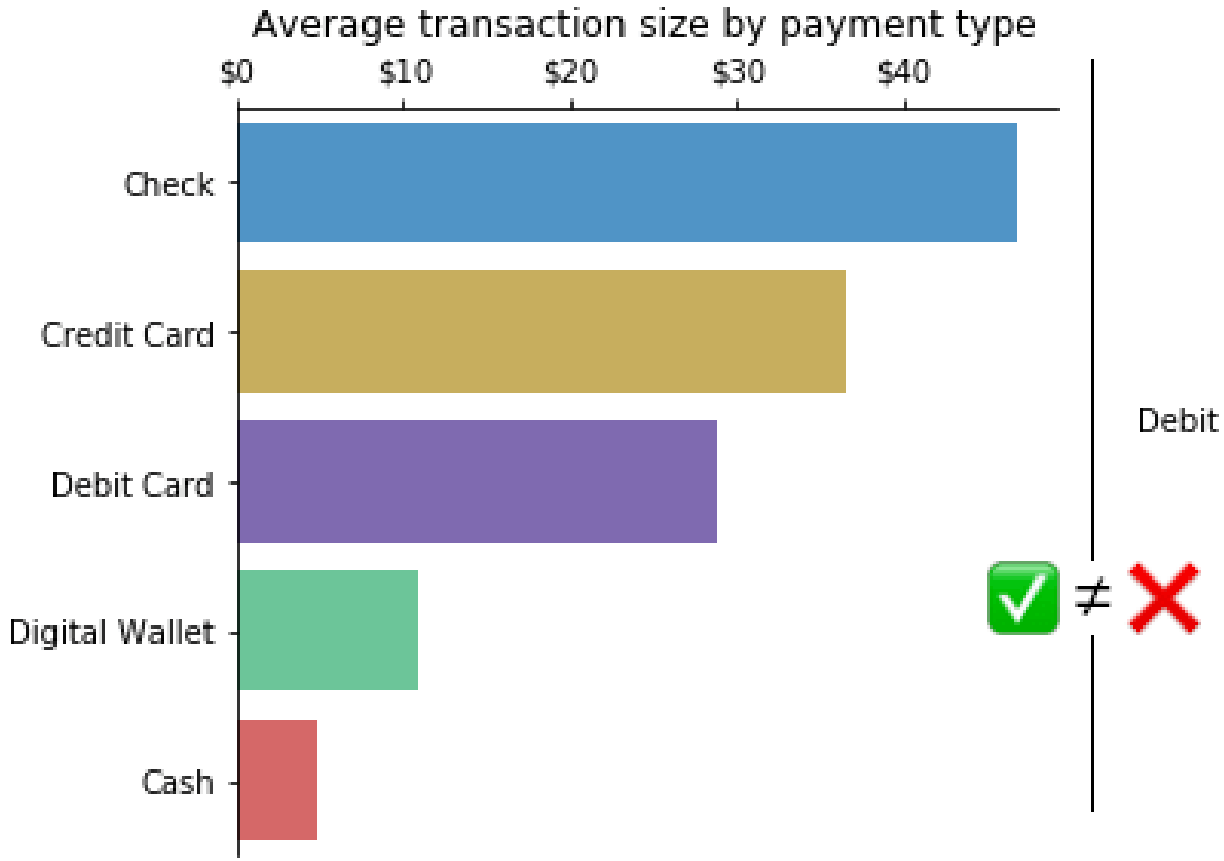
Visualization – Common Mistakes

- Misleading Cumulative Graphs (showing data that is increasing in quantity)

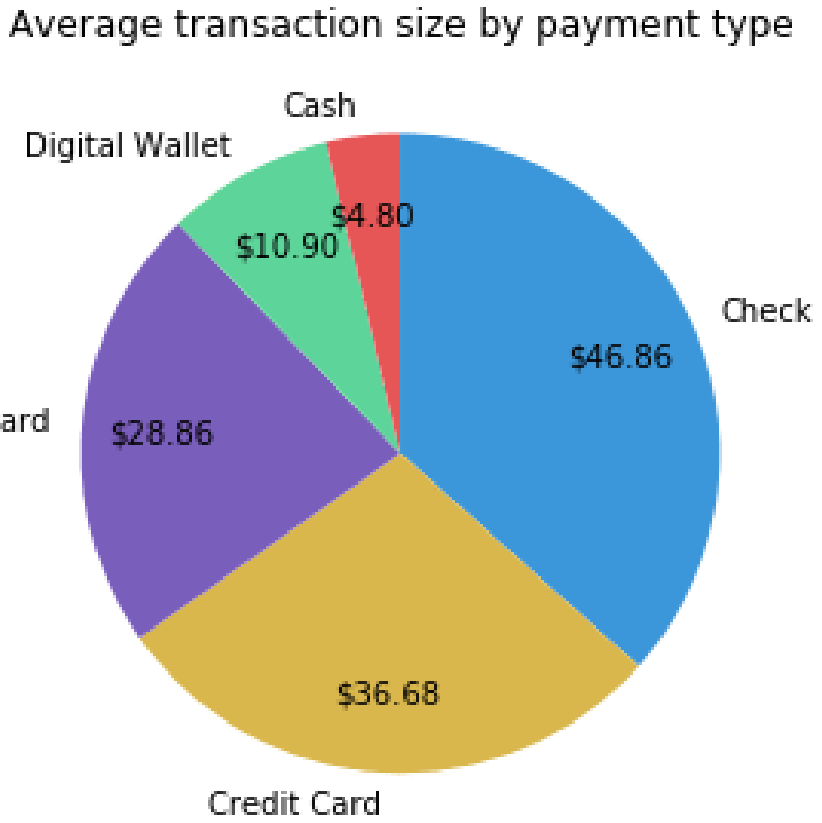


Visualization – Common Mistakes

- Ignoring Conventions
- Using the wrong type of visualizations

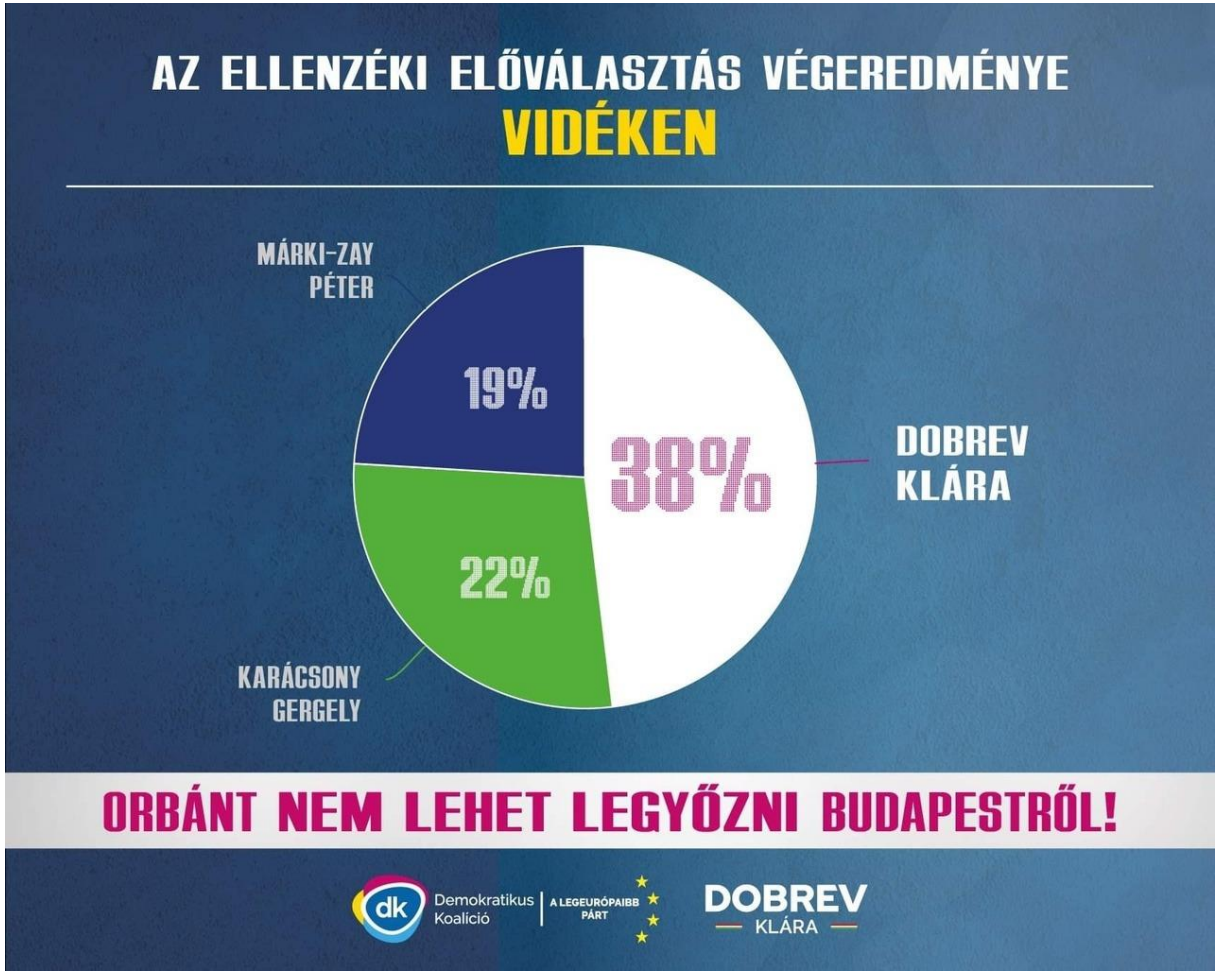


✅ ≠ ❌



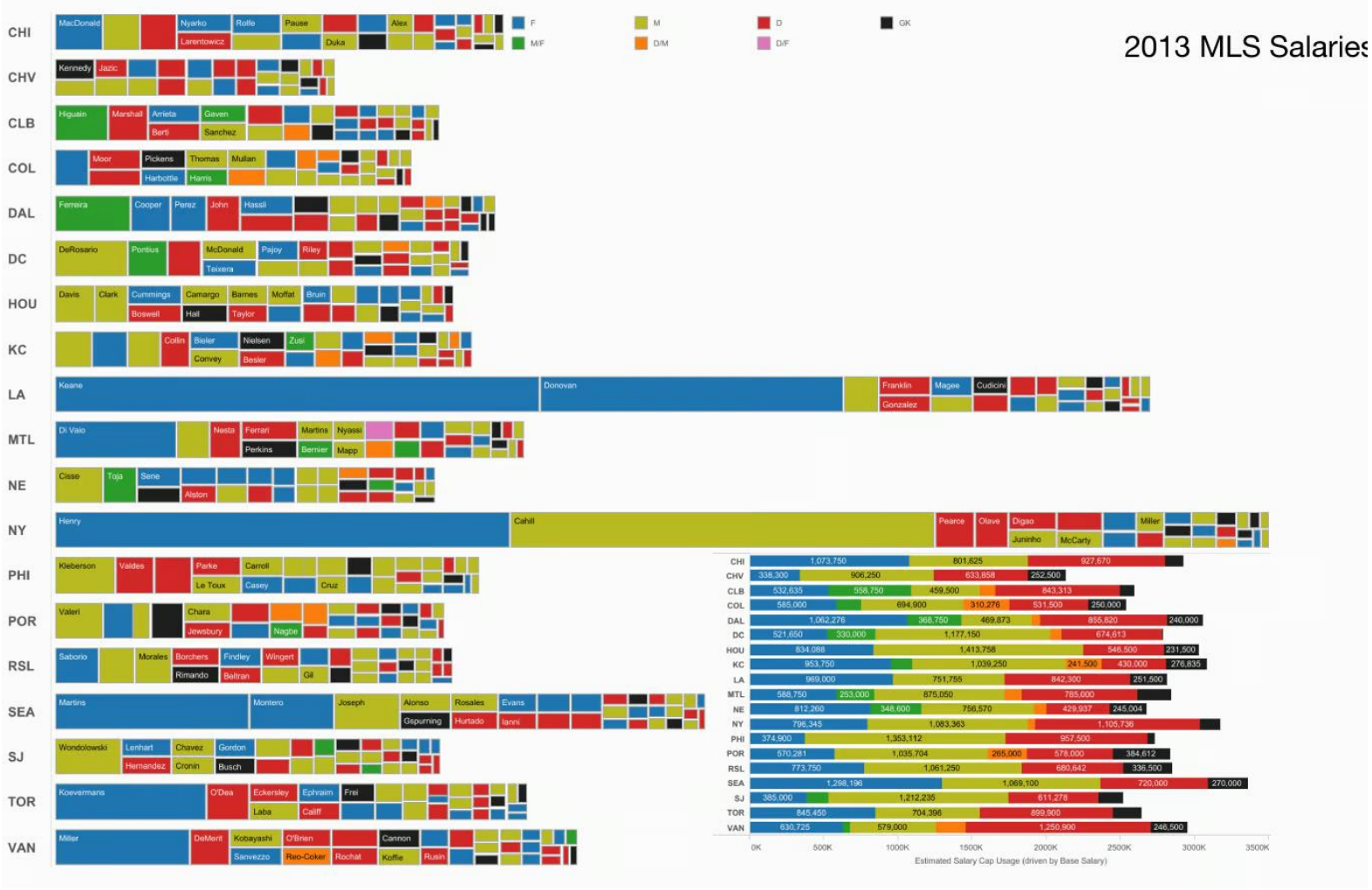
Visualization – Common Mistakes

- Failed Calculations



Visualization – Common Mistakes

- Displaying too much of data

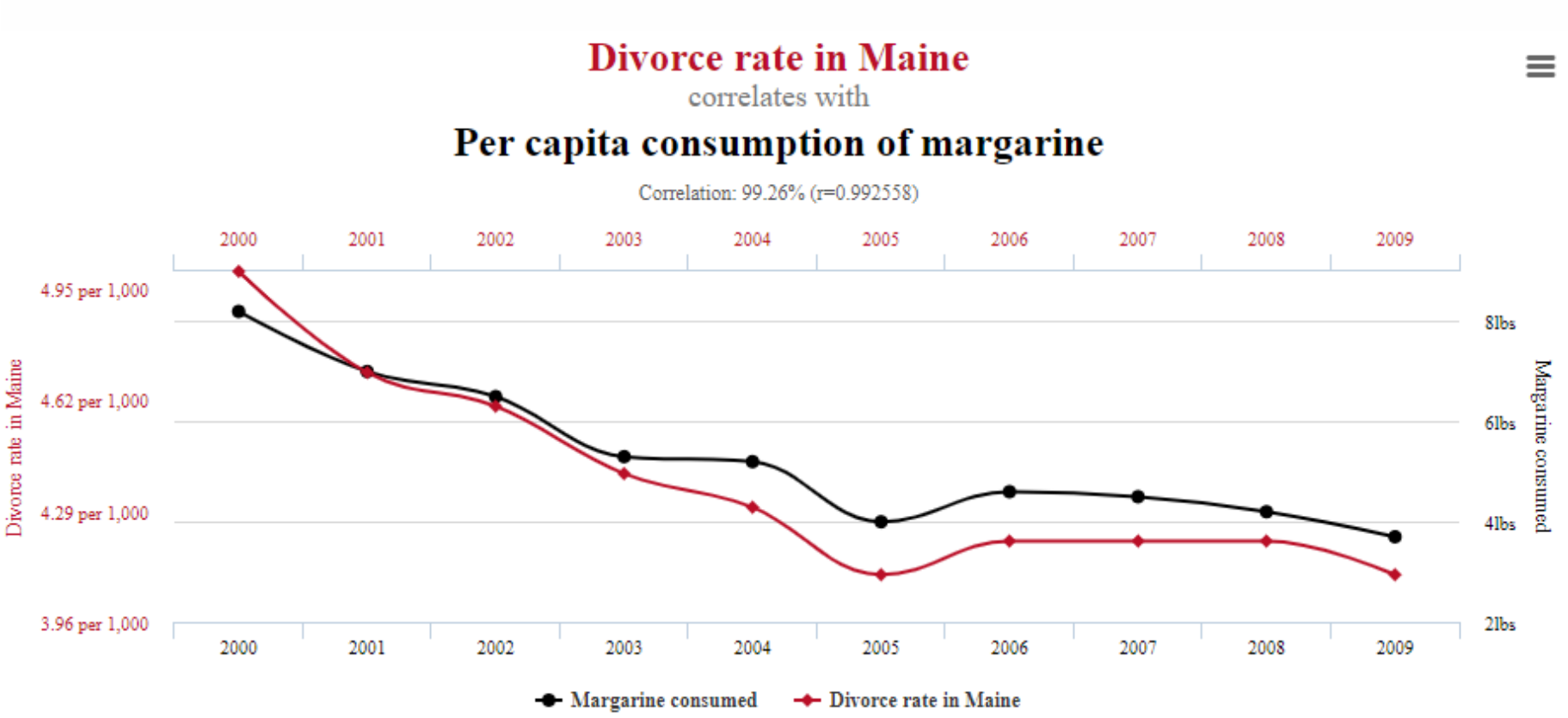


Visualization – Common Mistakes

- Purposeful & Selective bias
  - **Purposeful bias**: a deliberate attempt to influence data, most likely to take the form of data omissions or adjustments
  - **Selective bias**: slightly more discreet and passes by those who do not, or are not able, to read between the lines, such as the nature of the sample of people surveyed.

Visualization – Common Mistakes

- Correlation implying causation

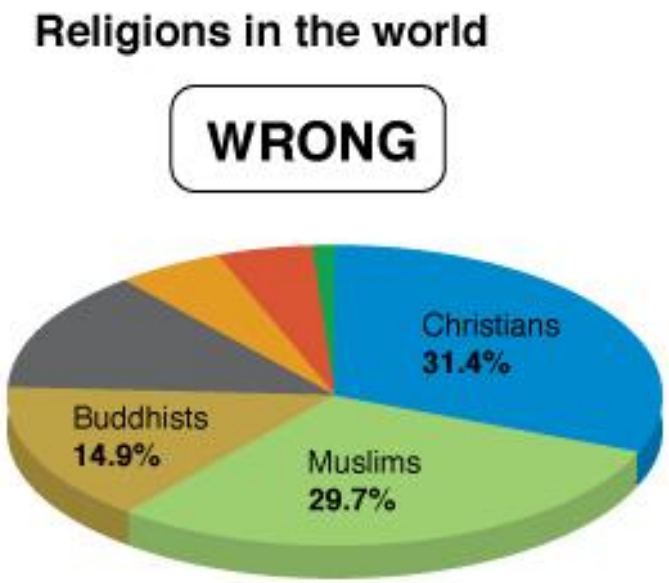


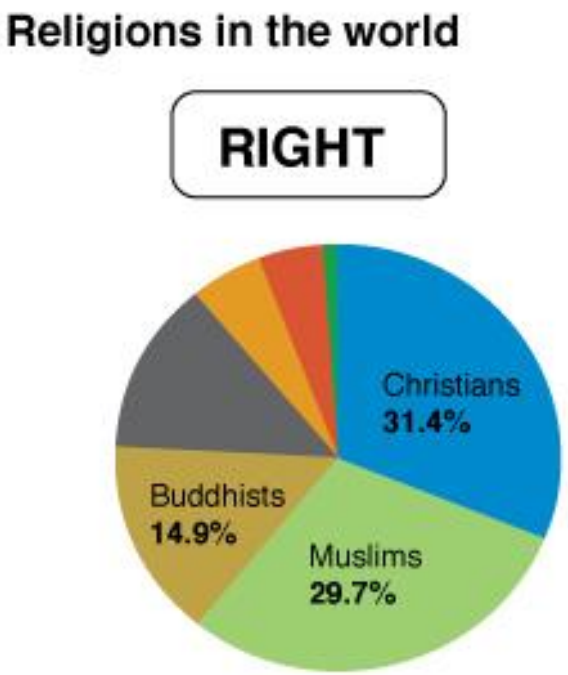
Visualization – Common Mistakes

- Ignoring population size makes accurate comparisons impossible

Visualization – Common Mistakes

- Decoration can be distracting and misleading





In Today's Lecture

- Visualization
- **Dashboarding**
- Interactivity

Dashboard – What is a dashboard?

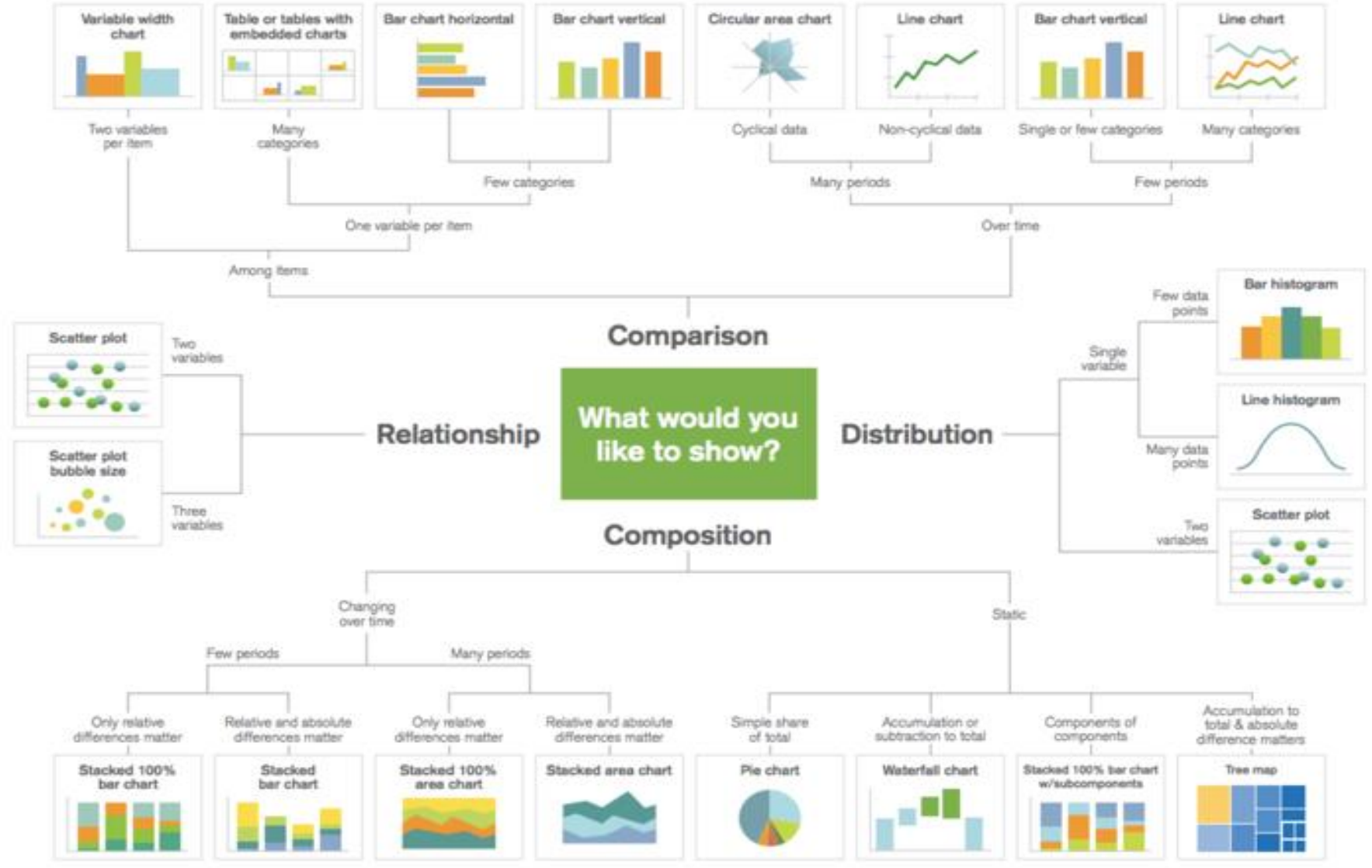
- A dashboard is a visual display of data used to monitor conditions and/or facilitate understanding. – *The Big Book of Dashboards*
- Some examples
  - An interactive display that allows people to explore worker compensation claims by region, industry, and body part.
  - A PDF showing key measures that gets e-mailed to an executive every Monday morning.
  - A large wall-mounted screen that shows support center statistics in real-time.
  - A mobile application that allows sales managers to review performance across different regions and compare year-to-date sales for the current year with the previous year.

Dashboard – What Makes a Good Dashboard?

- Answers a set of questions
- Follows a flow and invites interactivity
- Condensed; primarily in the form of summaries and exceptions
- Specific to and customized for the dashboard's audience and objectives
- Provides appropriate text for clarity and direction, if needed
- Makes strategic use of color

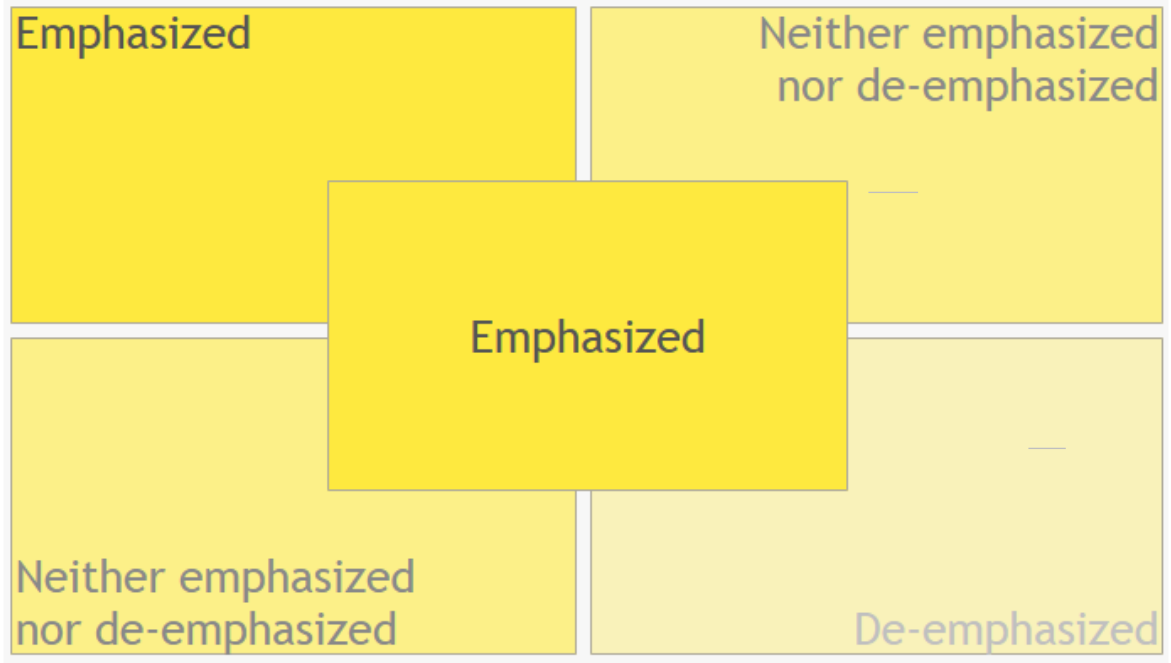
Dashboard – What to Consider?

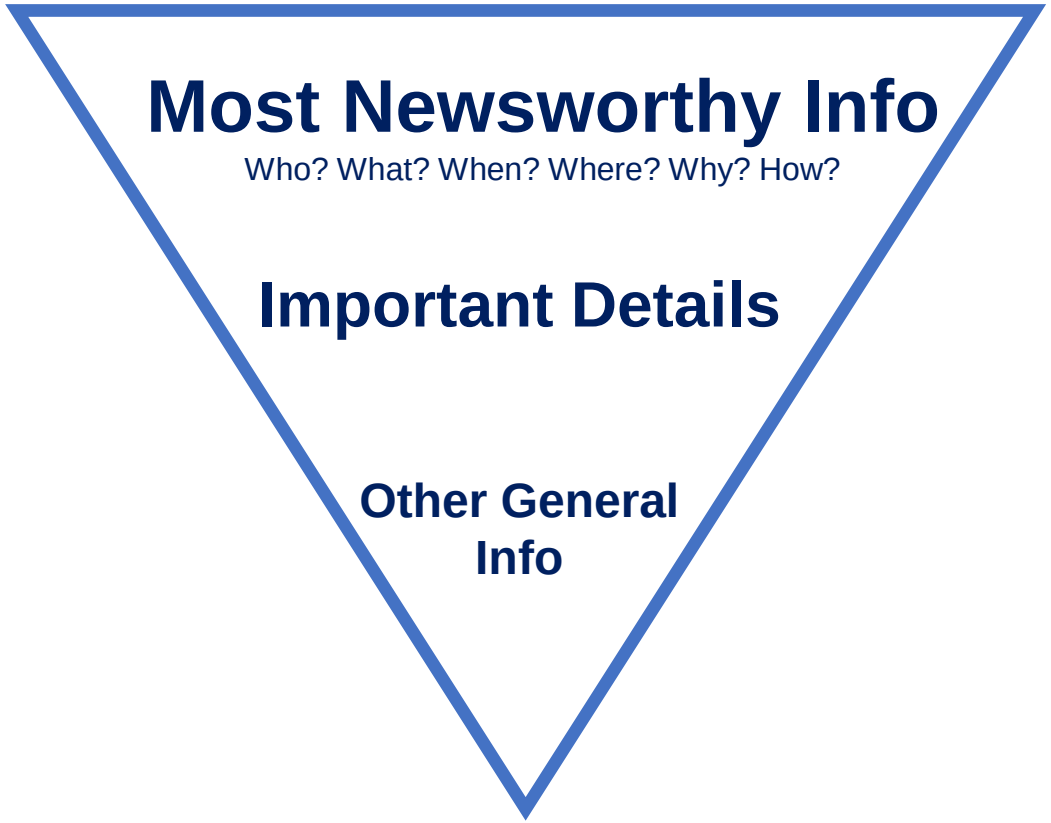
- Chart Type



Dashboard – What to Consider?

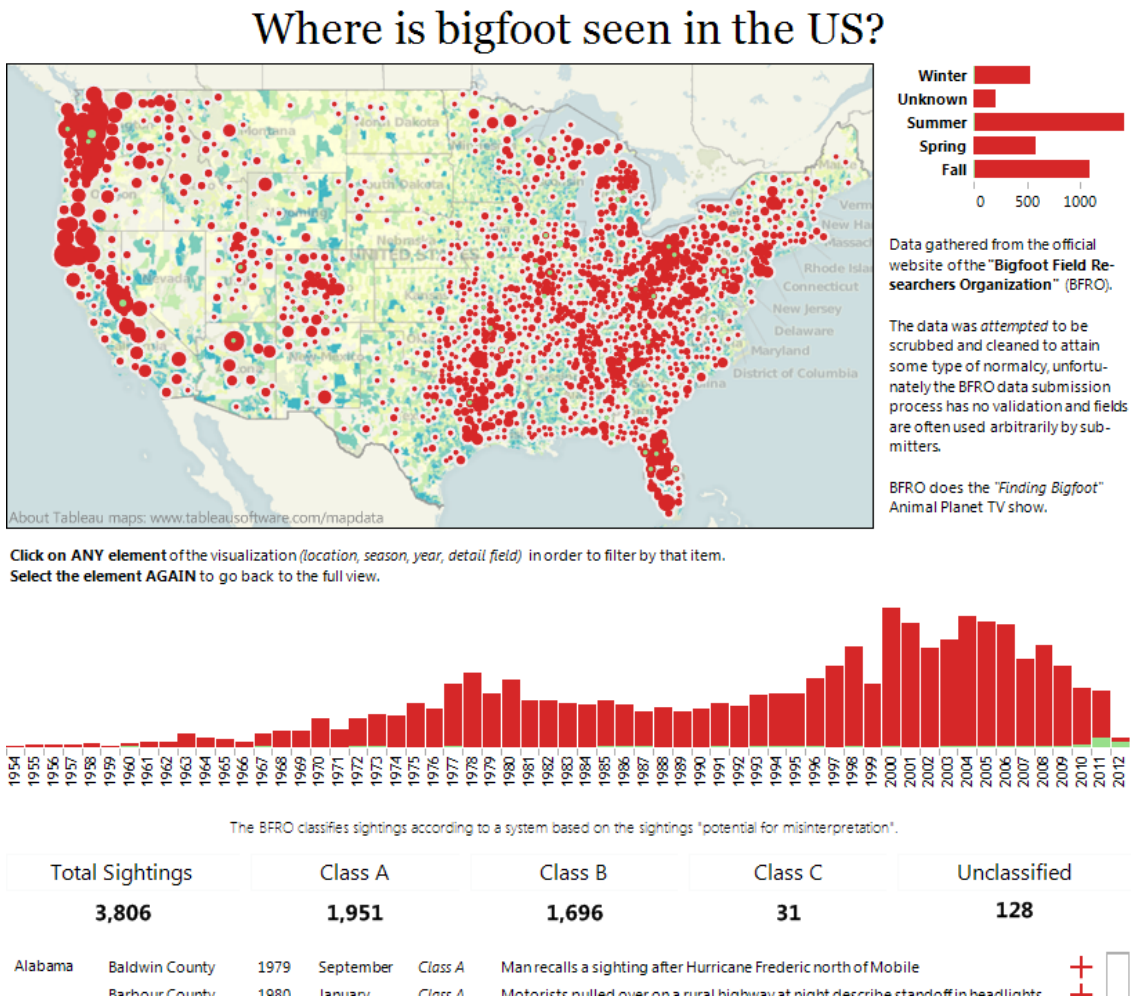
- Layout

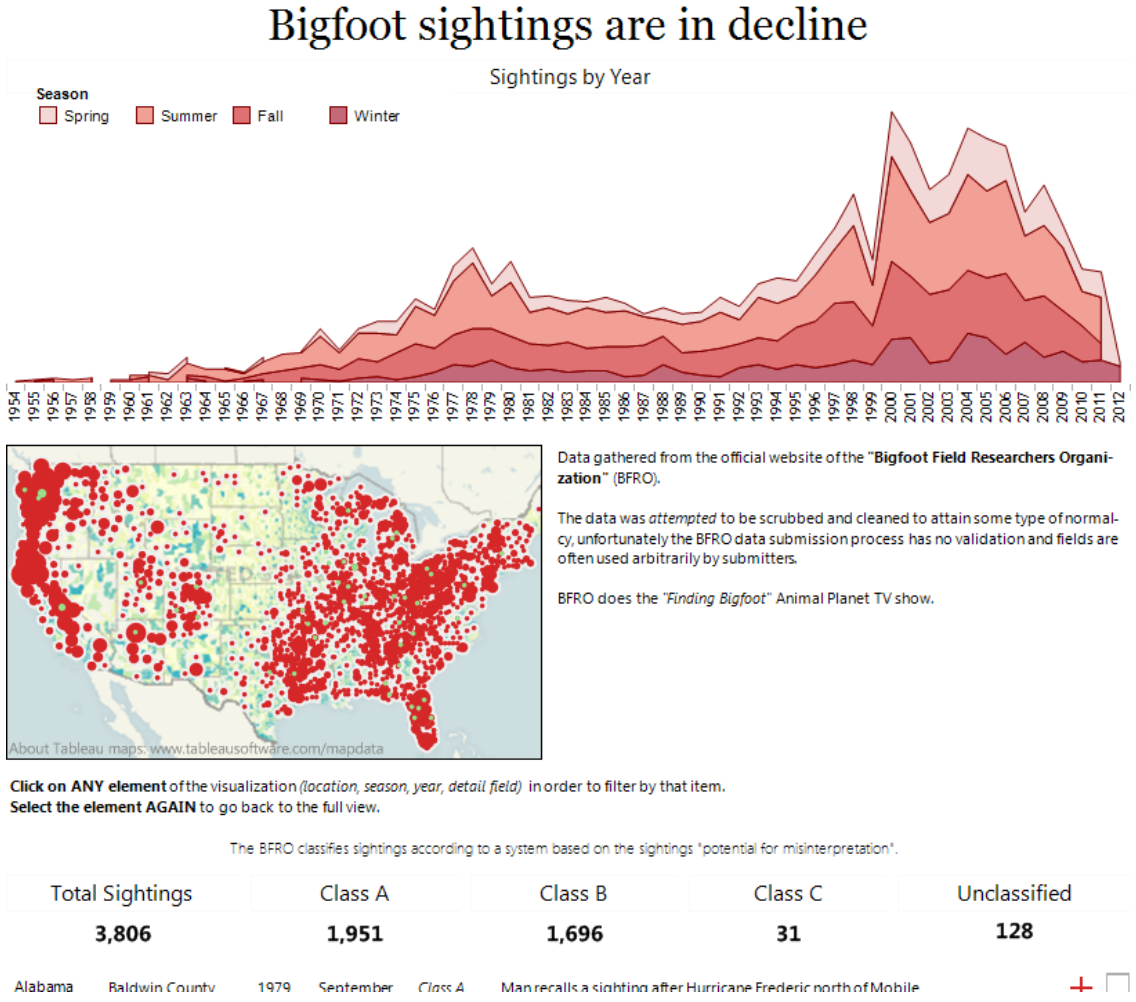




Dashboard – What to Consider?

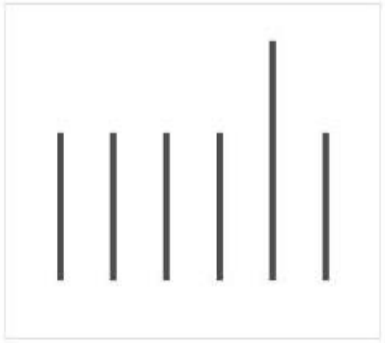
- Layout





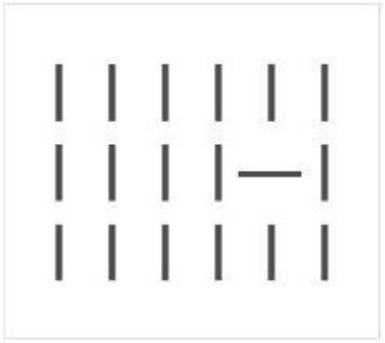
Dashboard – What to Consider?

- Pre-Attentive Attributes

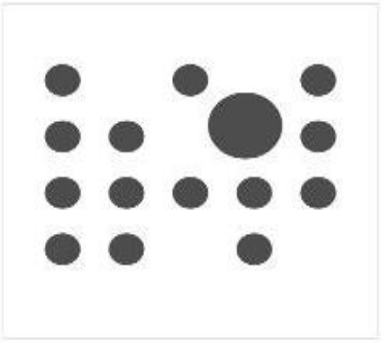


Length

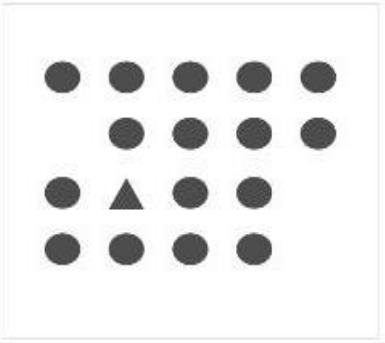
Width



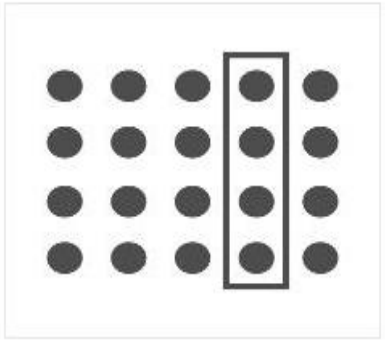
Orientation



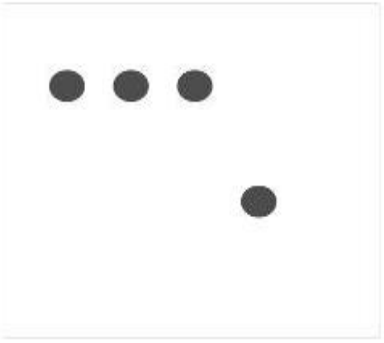
Size



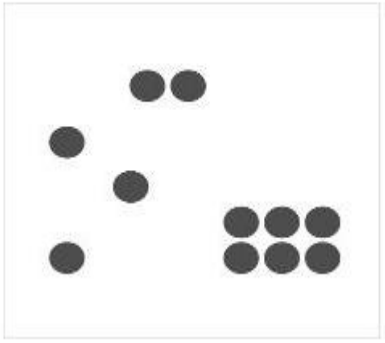
Shape



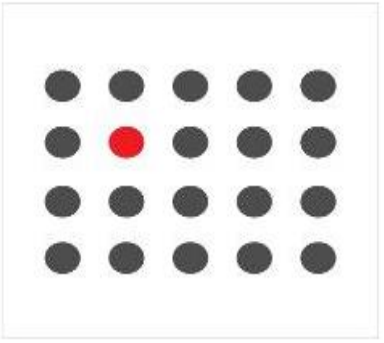
Enclosure



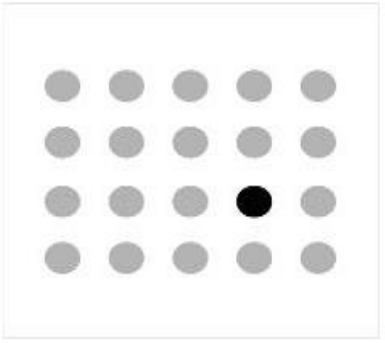
2D Position



Grouping



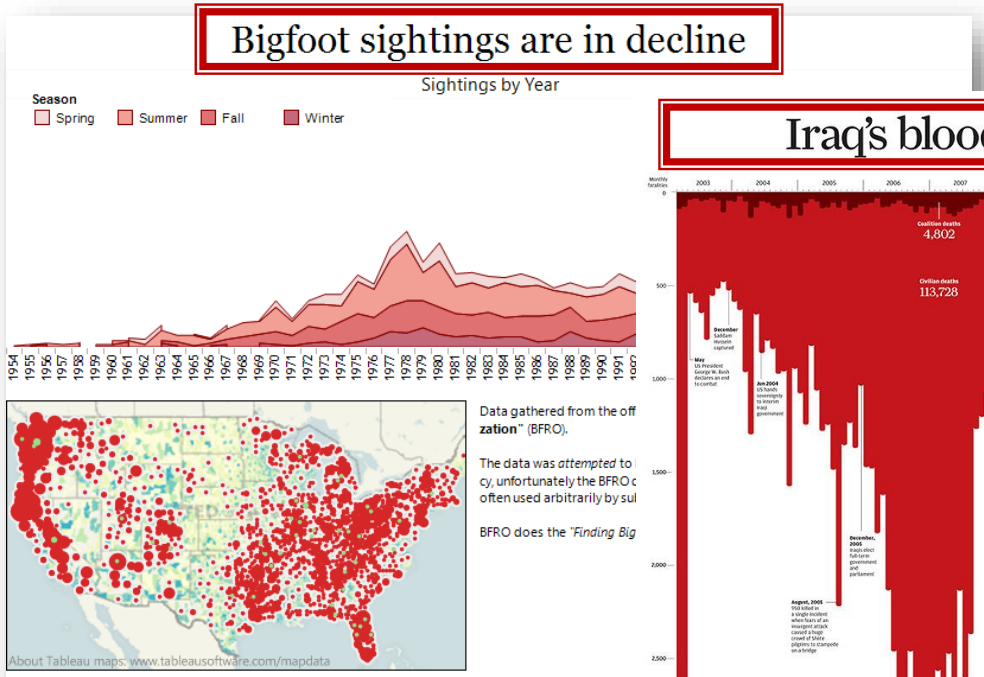
Color (Hue)

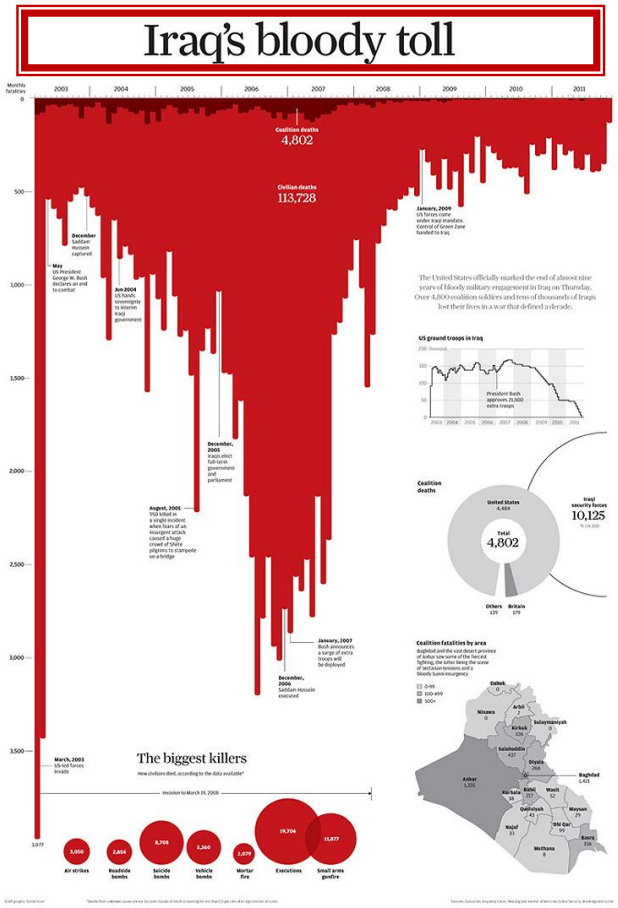


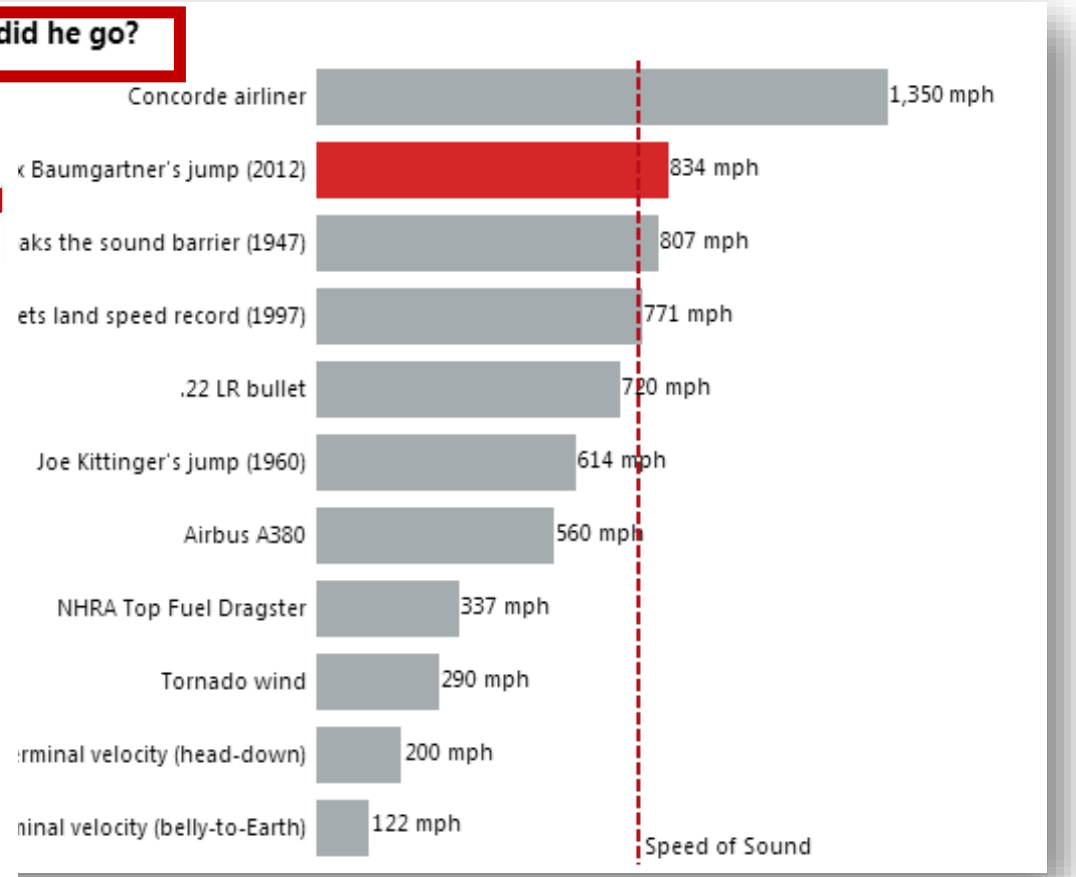
Color (Intensity)

Dashboard – What to Consider?

- Title
  - A concise, descriptive title can make a huge difference in garnering attention and making a chart more memorable.

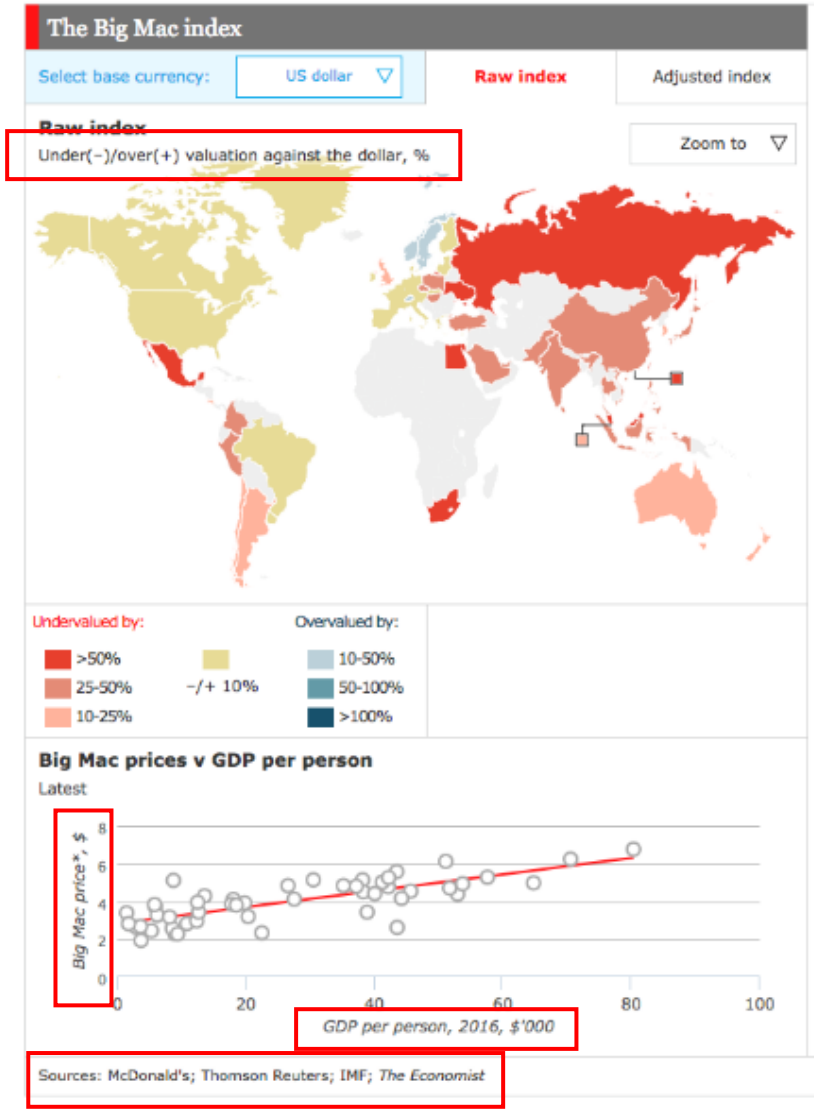


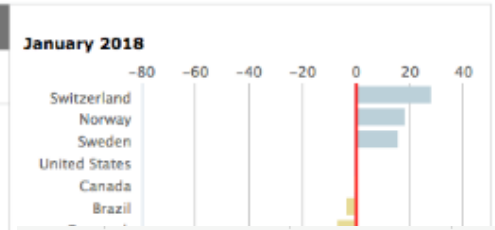


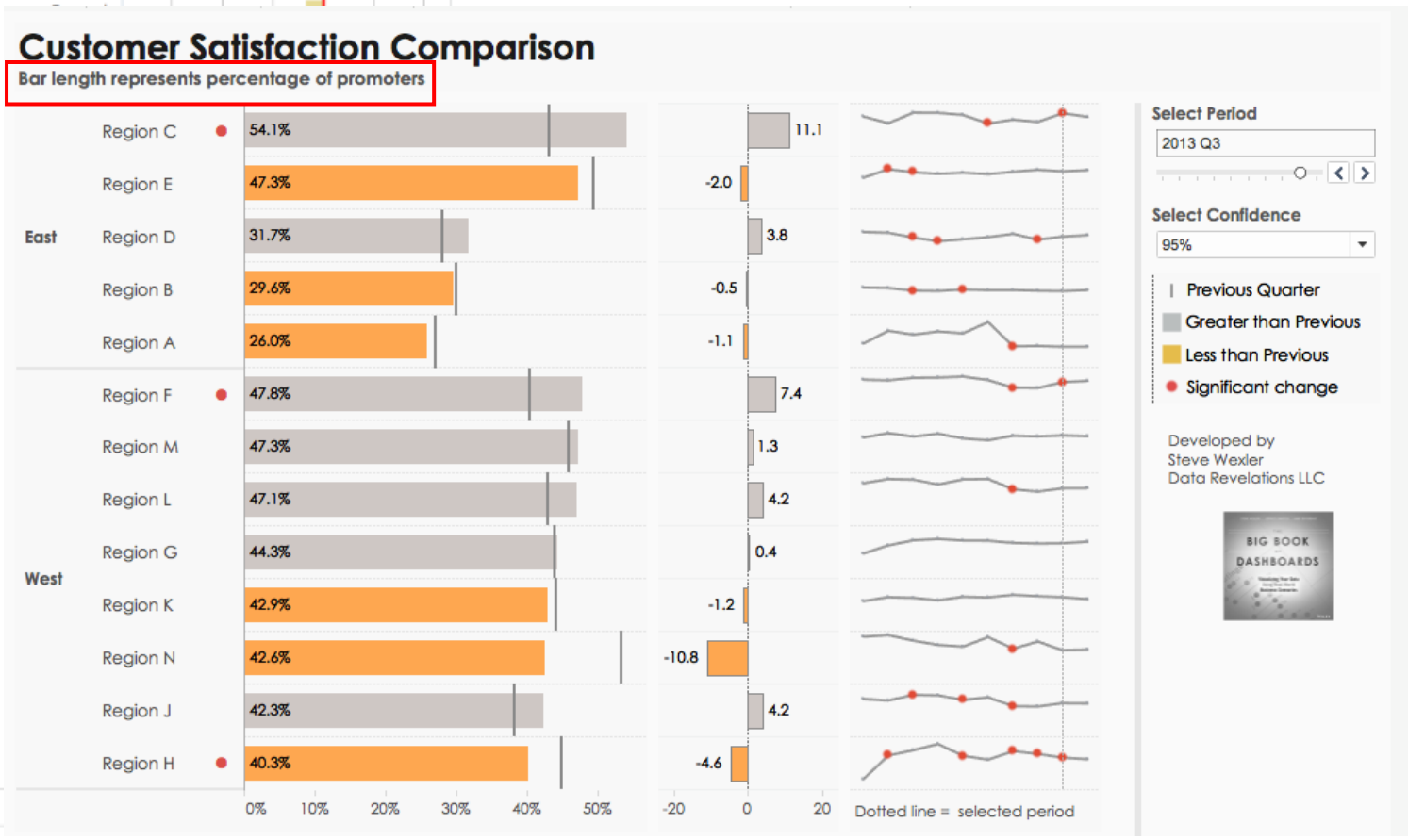


Dashboard – What to Consider?

- Adding additional context



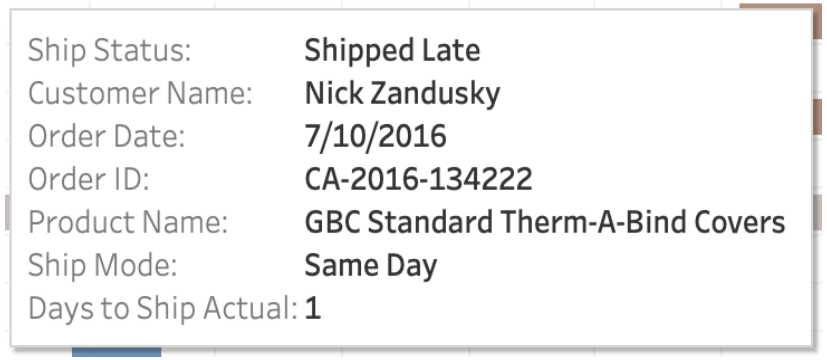




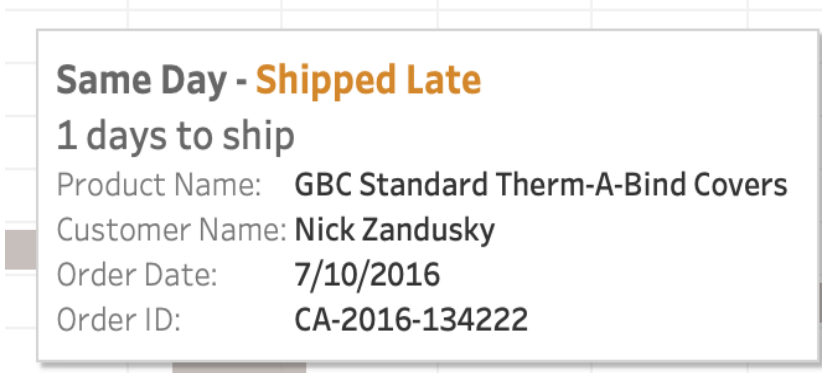
Dashboard – What to Consider?

- Tooltip
  - Tooltips can make the difference between a user loving your visualization and not understanding it.

  - Tips to improve tooltip designs
    - Use a proper font
    - Identify the most important part of the tool tip and make it your title
    - Change measure names/values to make them specific and understandable
    - Include proper units
    - Remove command prompts

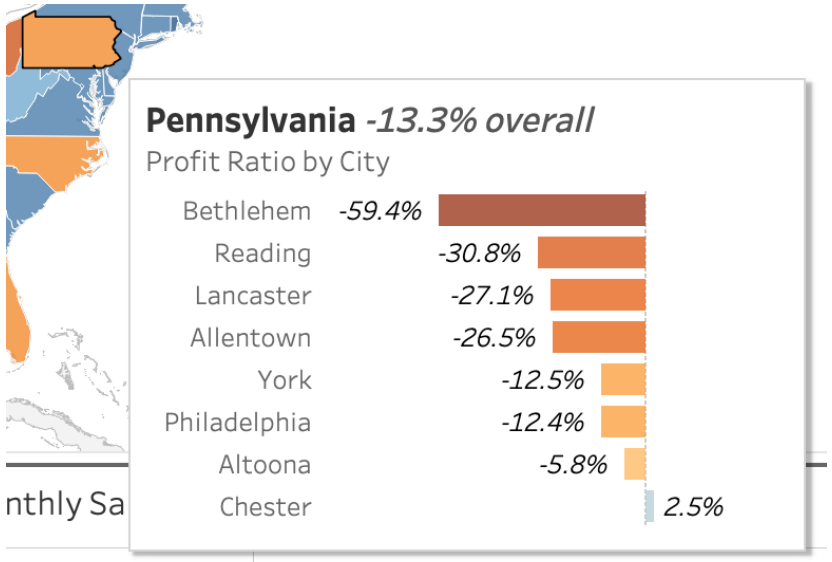


Default vs Modified



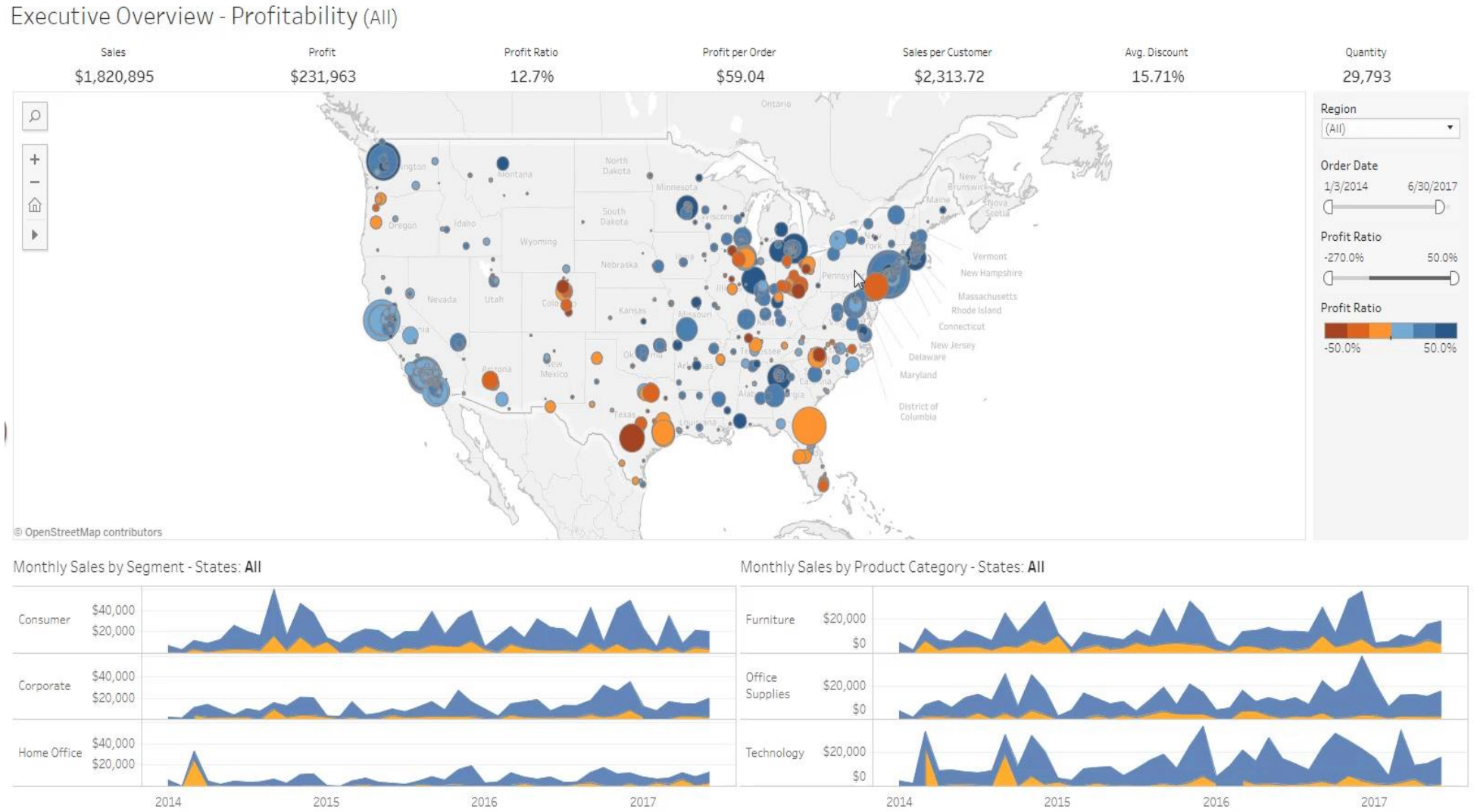
Dashboard – What to Consider?

- Tooltip
  - Tooltips can make the difference between a user loving your visualization and not understanding it.
  - Using visualizations in tooltips
    - Highlighting trends
    - Describing the distribution of data
    - Layering different levels of time or geography



Dashboard – What to Consider?

- Tooltip



Dashboard – What to Consider?

- Does your dashboard pass the 5-second test?
  - Most important view is on top or top left
  - Legends are near their views
  - Avoid multiple color schemes
  - Uses 5 or fewer views (charts)

In Today's Lecture

- Visualization
- Dashboarding
- **Interactivity**

Interactivity

- Interaction in Human-Computer Interaction
  - The communication between the user and the system.
  - Way of using a physical input/output device to perform a generic task in human-computer dialogue.

- Interaction and interactive behavior in visualization
  - Interactive features provide users with the ability to directly or indirectly manipulate and interpret representations.
  - The amount of data flowing from visualization systems to users is far greater than from users to systems.

J. S. Yi, Y. a. Kang, J. Stasko and J. A. Jacko, "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," in IEEE Transactions on Visualization and Computer Graphics

Interactivity – 7 Categories of Interaction Techniques

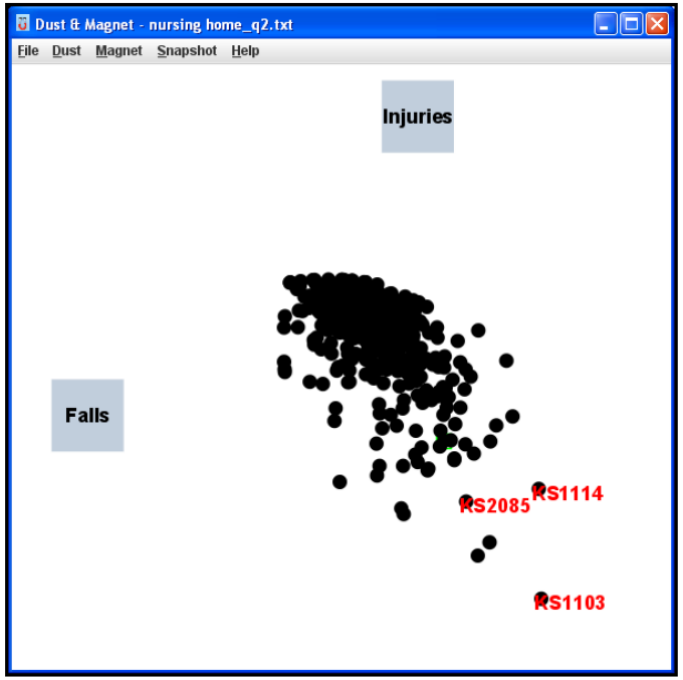
- Select: mark something as interesting
- Explore: show me something else
- Reconfigure: show me a different arrangement
- Encode: show me a different representation
- Abstract/Elaborate: show me more or less detail
- Filter: show me something conditionally
- Connect: show me related items

J. S. Yi, Y. a. Kang, J. Stasko and J. A. Jacko, "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," in IEEE Transactions on Visualization and Computer Graphics

Interactivity – 7 Categories of Interaction Techniques

1. Select: mark something as interesting

- Provide users with the ability to mark a data item(s) of interest to keep track of it.
- This is useful when too many data items are presented on a view, or when representations are changed, and it gets difficult for users to follow items of interest.

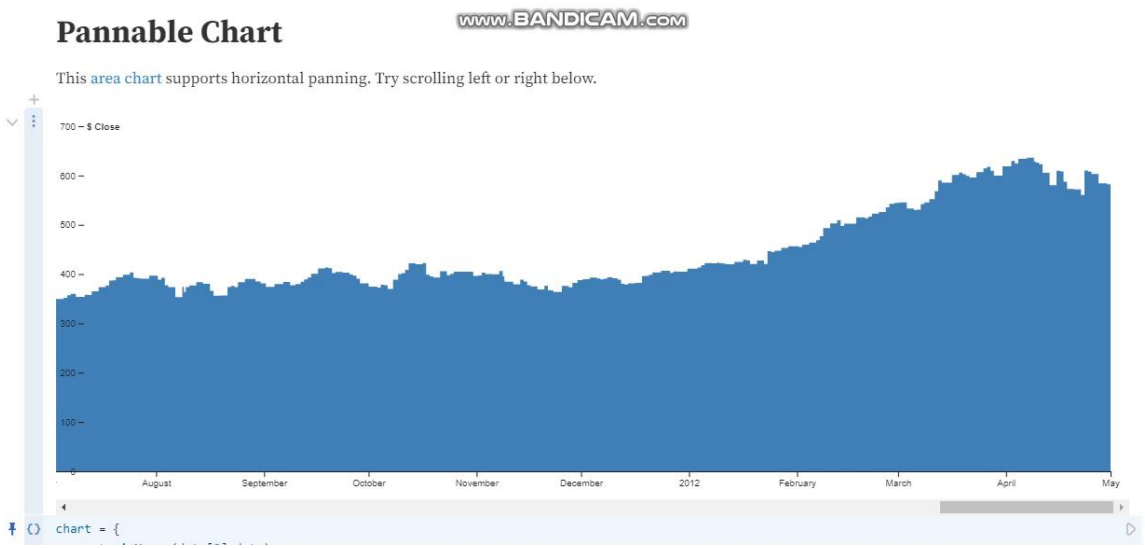


Interactivity – 7 Categories of Interaction Techniques

2. Explore: show me something else

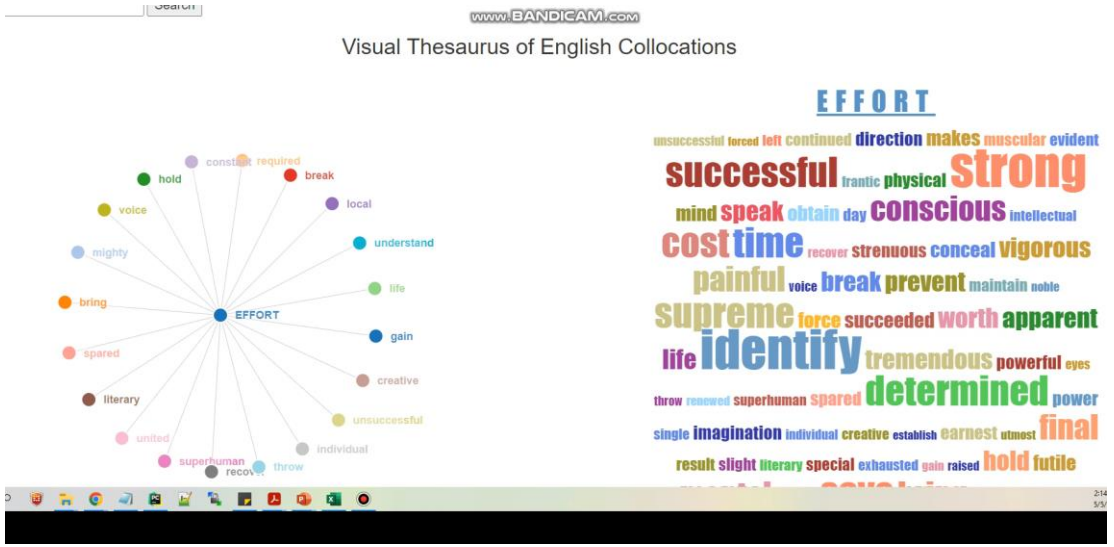
- Enable users to examine a different subset of data.

Panning



https://observablehq.com/@d3/pannable-chart

Direct Walk

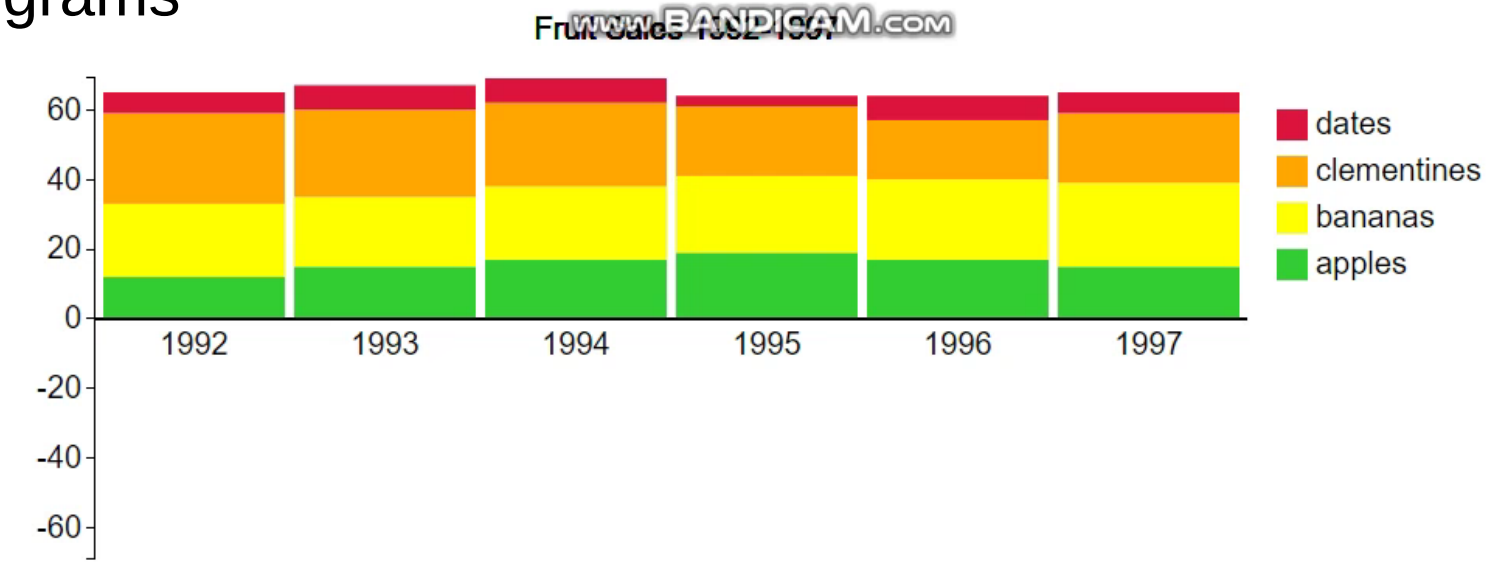


https://www.visual-thesaurus.com/

Interactivity – 7 Categories of Interaction Techniques

3. Reconfigure: show me a different arrangement

- Provide users with different perspectives on the data set by changing the spatial arrangement of representations.
- Reconfiguring methods
  - Sorting
  - Rearranging
  - Clustering
- E.g., Dancing Histograms



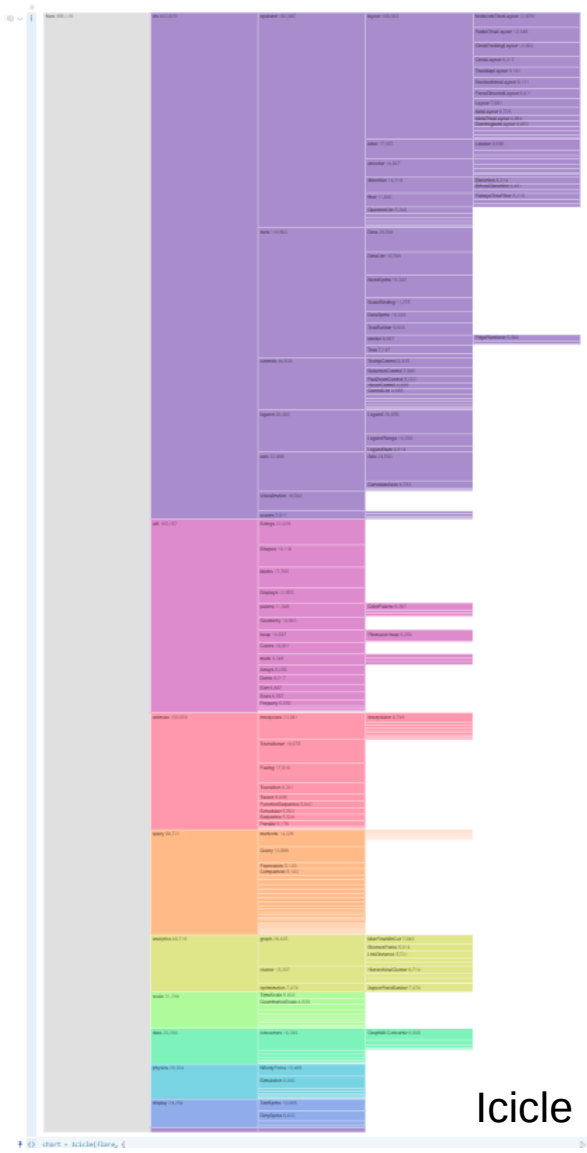
Interactivity – 7 Categories of Interaction Techniques

4. Encode: show me a different representation

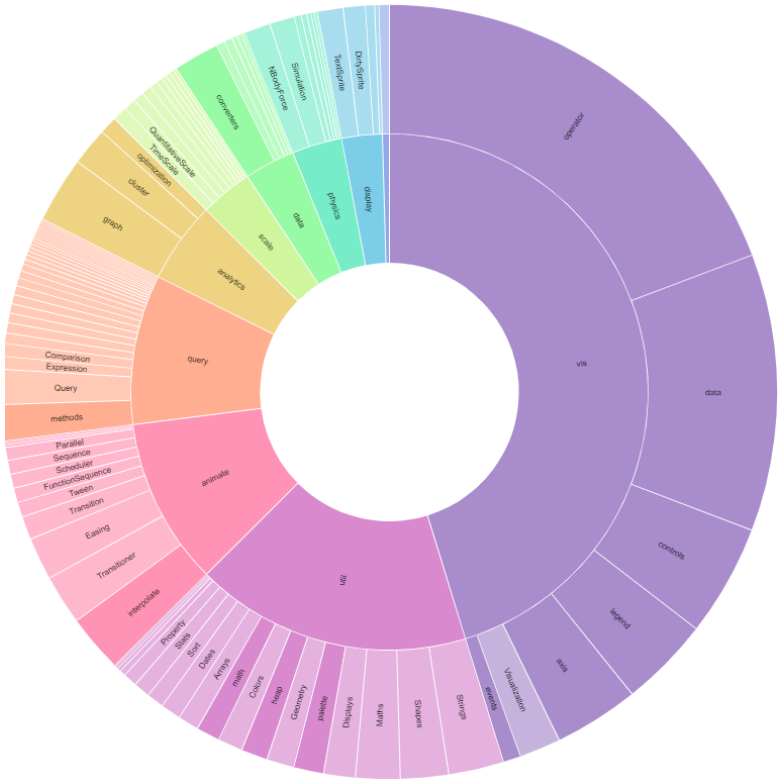
- Enable users to alter the fundamental visual representation of the data including the visual appearance (e.g., color, size, and shape) of each data element.
- Allows uncovering new aspects of a relationships
- E.g., Changing pie chat to histogram

Interactivity – 7 Categories of Interaction Techniques

4. Encode: show me a different representation



Icicle



Sunburst

Zoom-Sunburst

https://observablehq.com/@d3/zoomable-icicle

Interactivity – 7 Categories of Interaction Techniques

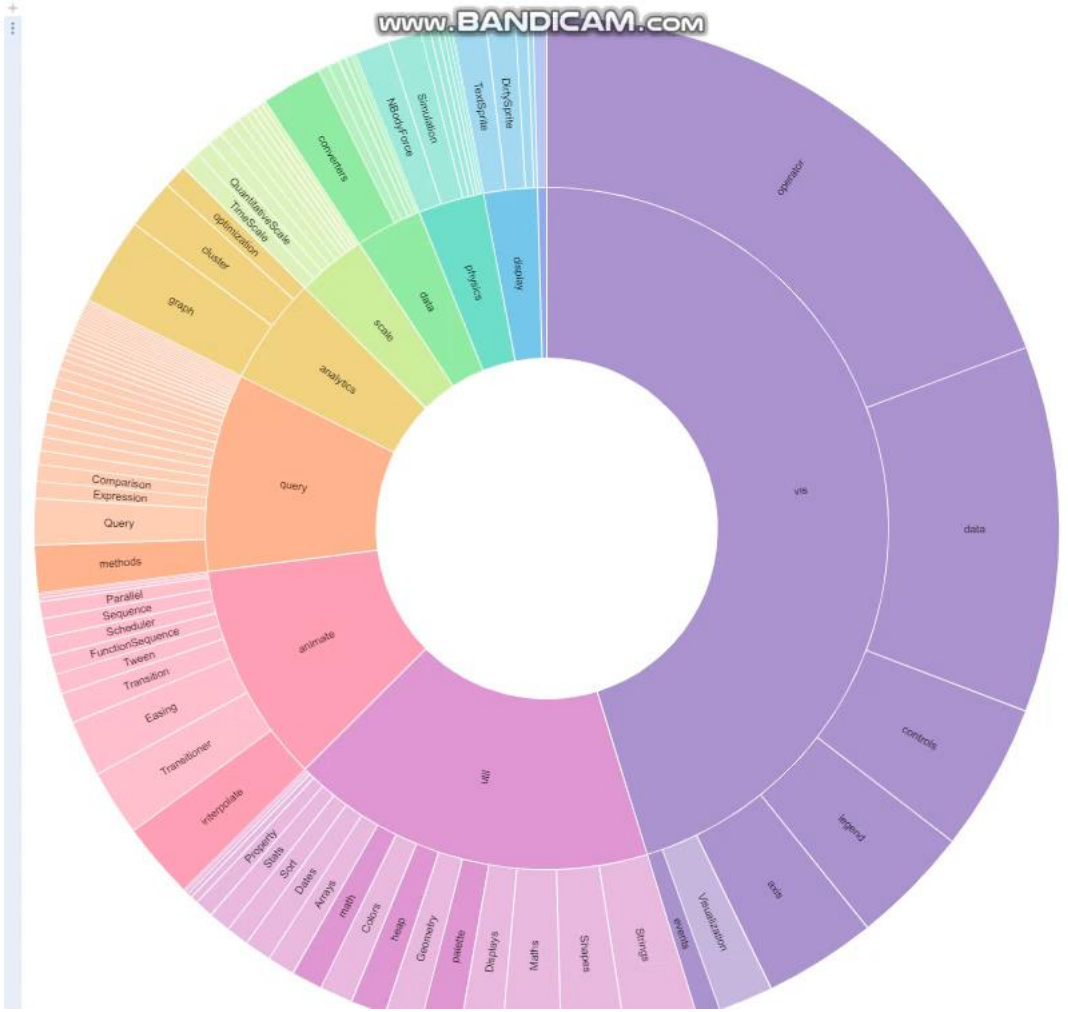
5. Abstract/Elaborate: show me more or less detail

- These techniques provide users with the ability to adjust the level of abstraction of a data representation.
- These types of interactions allow users to alter the representation from an overview down to details of individual data cases and often many levels in-between.

Interactivity – 7 Categories of Interaction Techniques

5. Abstract/Elaborate: show me more or less detail

- Details-on-demand
  - E.g., Zoom-Sunburst

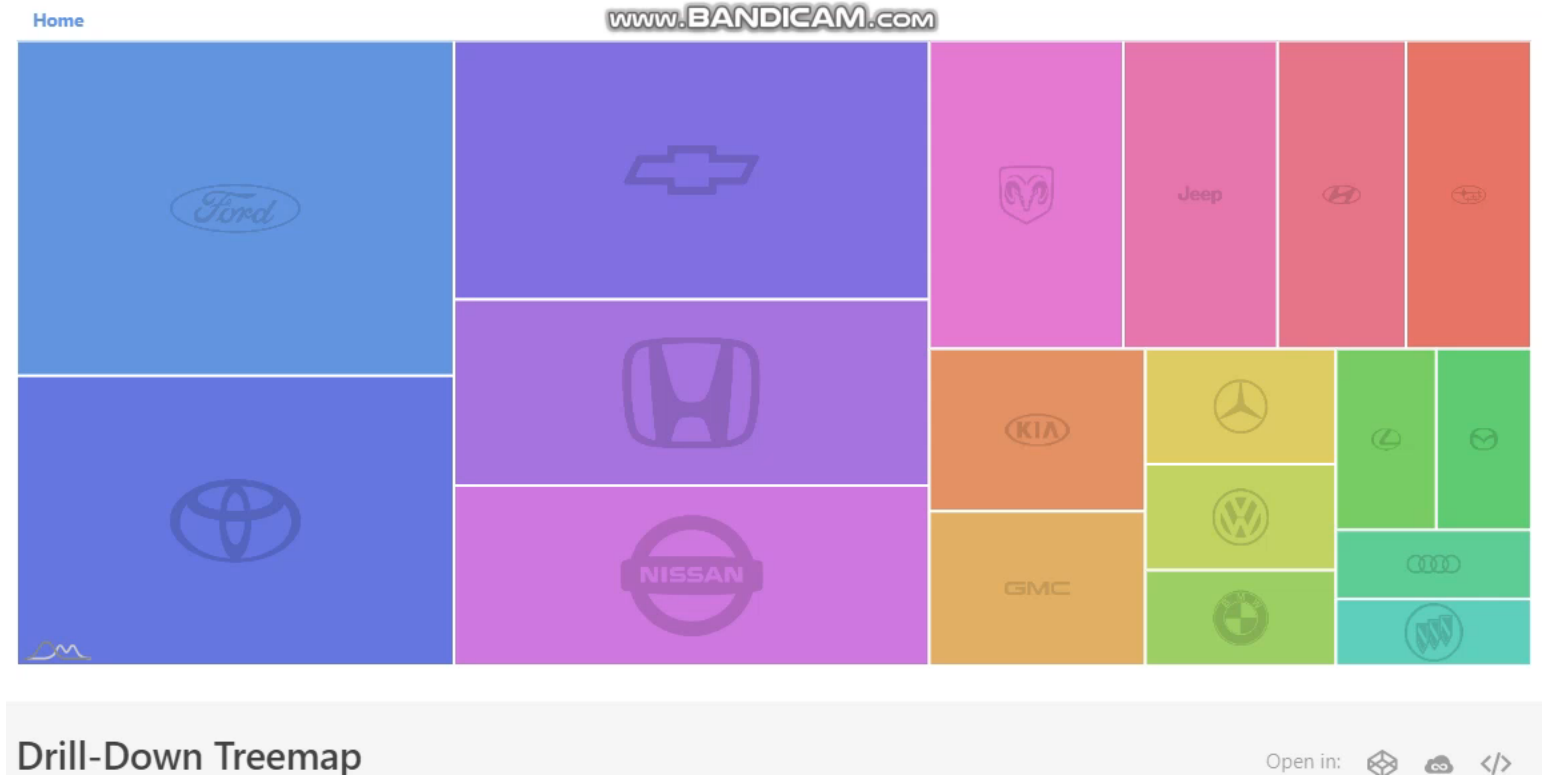


https://observablehq.com/@d3/zoomable-icicle

Interactivity – 7 Categories of Interaction Techniques

5. Abstract/Elaborate: show me more or less detail

- Drill-down
  - E.g., Drill-down-treemap



https://www.amcharts.com/demos/drill-down-treemap/

Interactivity – 7 Categories of Interaction Techniques

5. Abstract/Elaborate: show me more or less detail

- Zooming
  - Geometric zooming: the objects get larger as if you're getting closer to them.
  - Semantic zooming: they stay the same size but are just further apart when you zoom in

- Tooltips and visualization in tooltips

Interactivity – 7 Categories of Interaction Techniques

6. Filter: show me something conditionally

- Enable users to change the set of data items being presented based on some specific conditions.
- In this type of interaction, users specify a range or condition

Interactivity – 7 Categories of Interaction Techniques

7. Connect: show me related items

- Enable users to change the set of data items being presented based on some specific conditions. In this type of interaction, users specify a range or condition
- Interaction techniques that are used to
  - Highlight associations and relationships between data items that are already represented
  - Show hidden data items that are relevant to a specified item.

Tools for Data Visualization

- Microsoft Excel
- Power BI
- Google Charts
- Tableau
- Zoho Analytics
- Datawrapper
- Infogram

Questions?