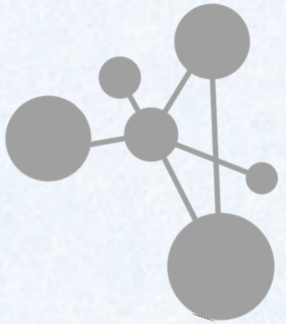


An Introduction to Data Visualization

Amy Trost, Data Services Librarian

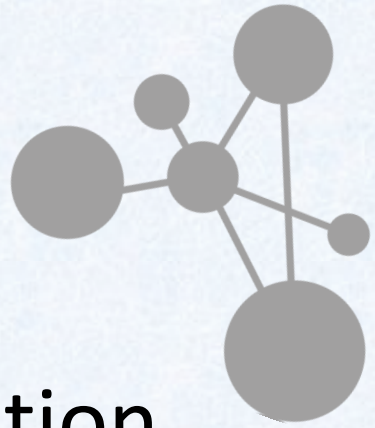
atrost1@umd.edu

Workshop materials available at

https://github.com/amytrost/dataviz_intro_2020

Priddy Library

During this session we will:

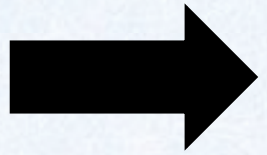


- Review some definitions around data visualization
- Explore the elements of an effective visualization
- Look at a few beautiful examples
- Demonstrate visualization creation in R and Tableau

Why create visualizations?

1. Raw data (in Excel)

	A	B	C	D
1	**Order Date**	**Sales**	**Product Category**	
2	1/1/2009	872.48	Office Supplies	
3	1/1/2009	180.36	Office Supplies	
4	1/2/2009	124.81	Office Supplies	
5	1/2/2009	1239.06	Furniture	
6	1/2/2009	4083.19	Furniture	
7	1/2/2009	4902.38	Office Supplies	
8	1/2/2009	137.63	Office Supplies	
9	1/2/2009	614.8	Office Supplies	
10	1/3/2009	85.56	Office Supplies	
11	1/3/2009	754.6555	Technology	
12	1/3/2009	172.51	Office Supplies	
13	1/3/2009	698	Furniture	
14	1/3/2009	522.49	Technology	
15	1/3/2009	28359.4	Technology	
16	1/3/2009	123.76	Office Supplies	
17	1/3/2009	122.23	Office Supplies	
18	1/3/2009	262.76	Furniture	
19	1/3/2009	896.49	Technology	
20	1/4/2009	63.34	Office Supplies	
21	1/4/2009	151.35	Office Supplies	
22	1/4/2009	1039.56	Furniture	
23	1/5/2009	700.73	Technology	
24	1/5/2009	165.75	Technology	
25	1/5/2009	2021.147	Technology	
26	1/5/2009	1244.19	Furniture	
27	1/5/2009	4201.08	Furniture	
28	1/5/2009	8958.46	Furniture	
29	1/5/2009	4913.7	Technology	
30	1/5/2009	78.08	Furniture	



Sum of Sales
14,915,600.82

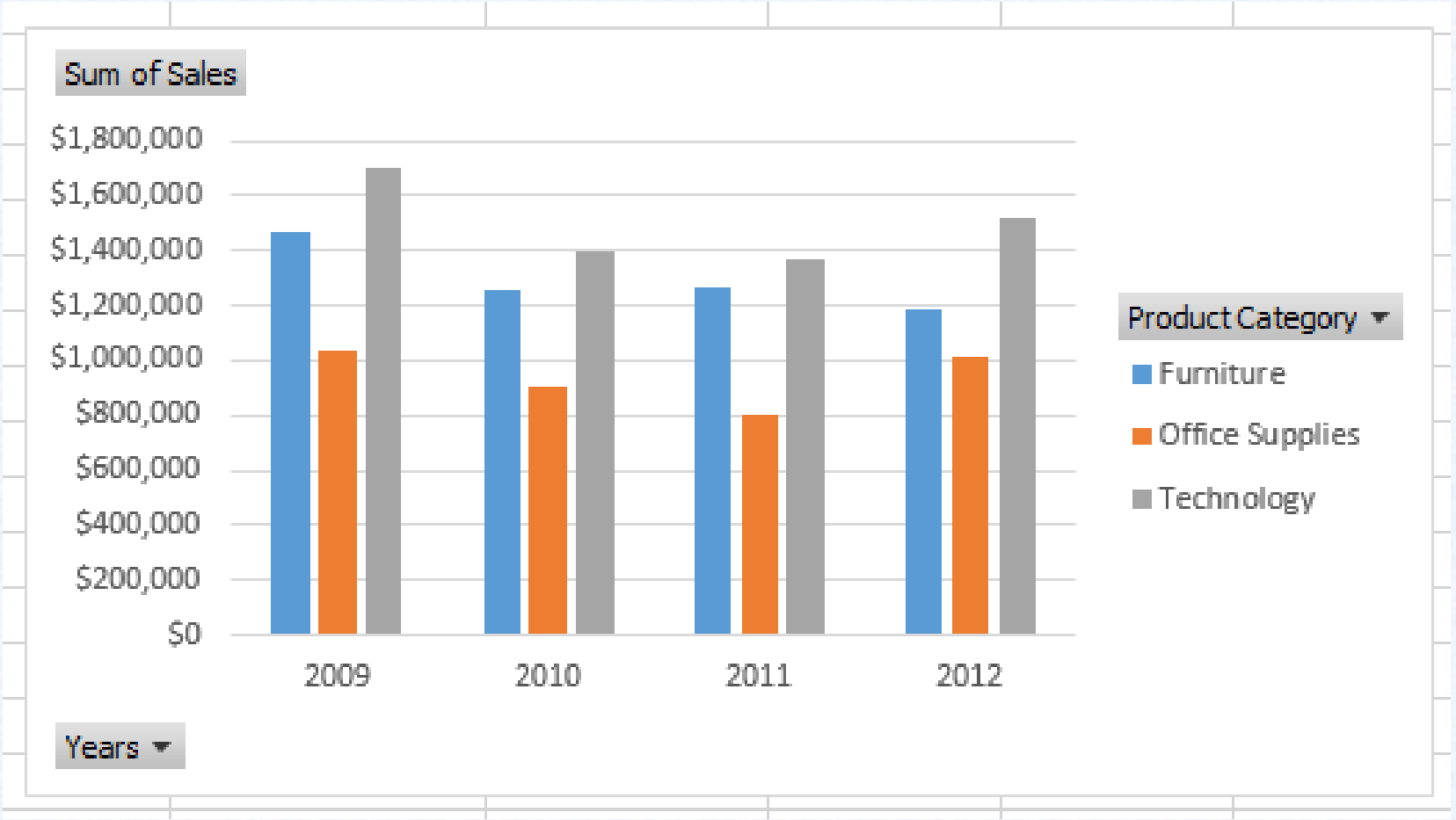
Why create visualizations?

2. Pivot Tables (in Excel)

Row Labels	Sum of Sales
Furniture	5178590.542
Office Supplie	3752762.1
Technology	5984248.182
Grand Total	**14,915,600.82**

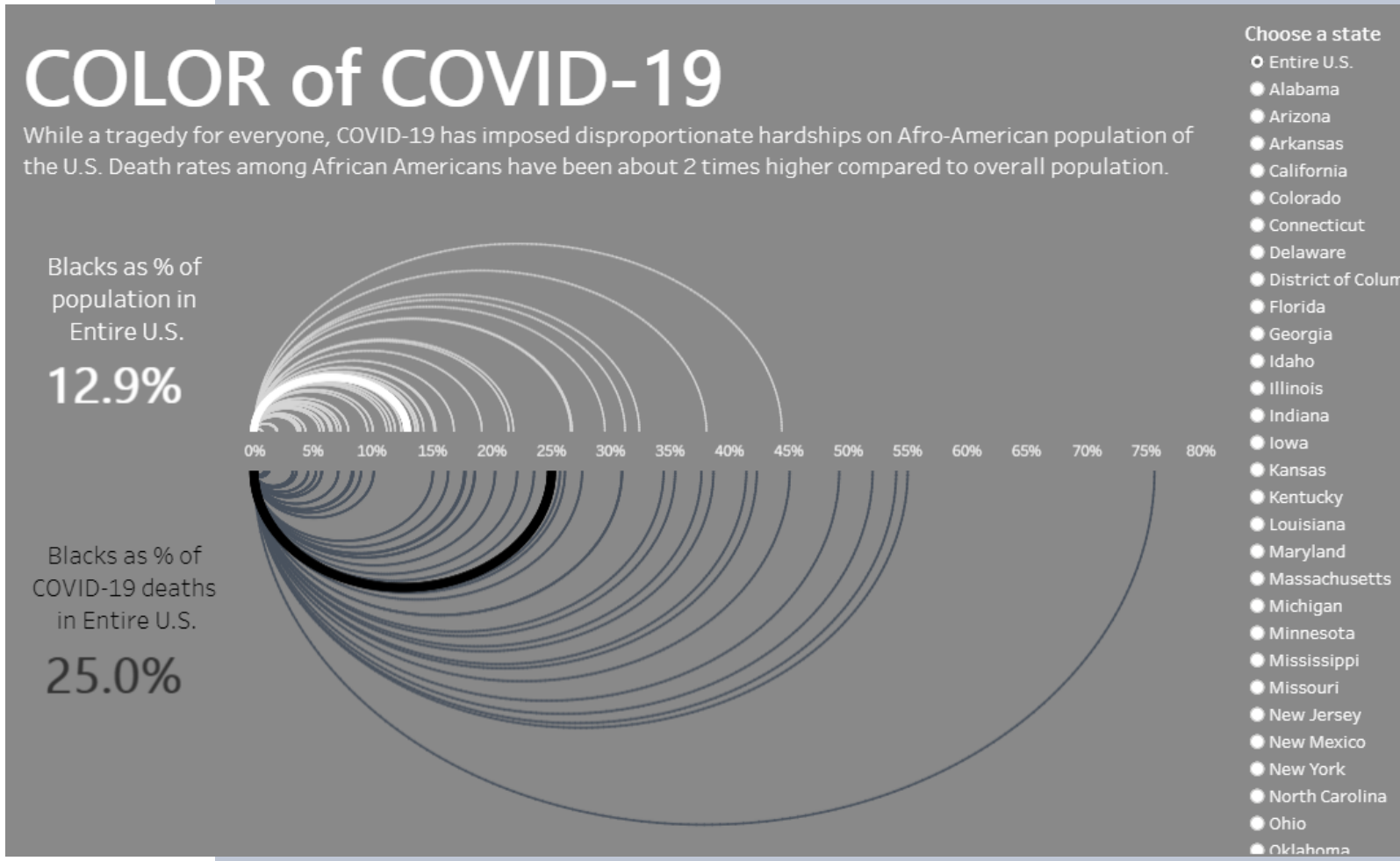
12	Row Labels	Furniture	Office Supplies	Technology	Grand Total
13	2009	$ 1,472,672	$ 1,034,642	$ 1,701,825	$ 4,209,139
14	2010	$ 1,252,665	$ 899,874	$ 1,397,142	$ 3,549,681
15	2011	$ 1,269,661	$ 800,349	$ 1,366,807	$ 3,436,817
16	2012	$ 1,183,593	$ 1,017,897	$ 1,518,474	$ 3,719,964
17	**Grand Total**	**$ 5,178,591**	**$ 3,752,762**	**$ 5,984,248**	**$ 14,915,601**

Why create visualizations?



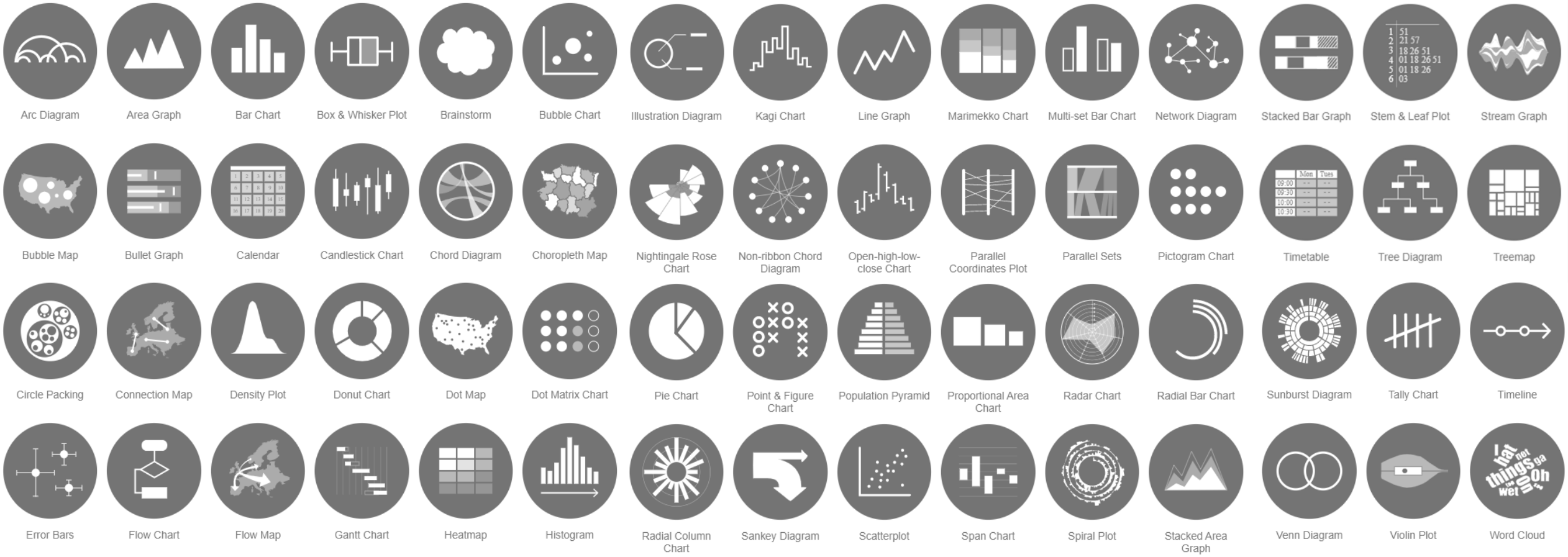
3. Bar Graph (in Excel)

Data Visualization: The graphical representation of information using visual elements like charts, graphs, and maps. An accessible way to see and understand trends, outliers, and patterns in data.



Source: Tableau learning and Tableau public gallery visualization of the day. June 12, 2020. See https://www.tableau.com/learn/articles/data-visualization and https://public.tableau.com/en-us/gallery/color-covid-19?tab=viz-of-the-day&type=viz-of-the-day

Visualization Types



Source: https://datavizcatalogue.com/

Chart Suggestions—A Thought-Starter

What would you like to show?

Comparison

- Among Items
  - Two Variables per Item
    - Variable Width Column Chart
  - Many Categories
    - Table or Table with Embedded Charts
  - One Variable per Item
    - Few Categories
      - Many Items
        - Bar Chart
      - Few Items
        - Column Chart
    - Many Periods
      - Cyclical Data
        - Circular Area Chart
      - Non-Cyclical Data
        - Line Chart
- Over Time
  - Few Periods
    - Single or Few Categories
      - Column Chart
    - Many Categories
      - Line Chart

Relationship

- Two Variables
  - Scatter Chart
- Three Variables
  - Bubble Chart

Distribution

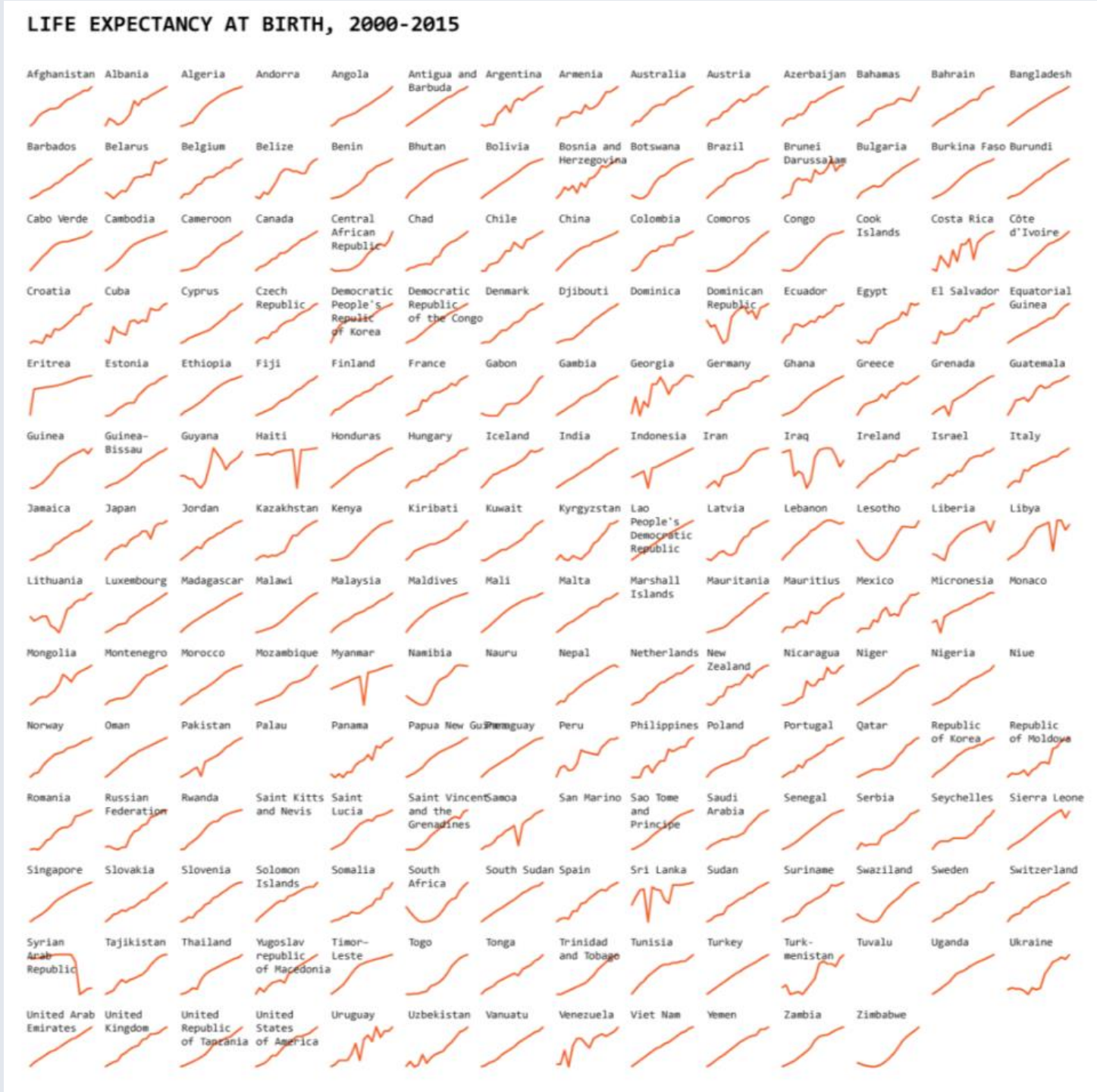
- Single Variable
  - Few Data Points
    - Column Histogram
  - Many Data Points
    - Line Histogram
- Two Variables
  - Scatter Chart
- Three Variables
  - 3D Area Chart

Composition

- Changing Over Time
  - Few Periods
    - Only Relative Differences Matter
      - Stacked 100% Column Chart
    - Relative and Absolute Differences Matter
      - Stacked Column Chart
  - Many Periods
    - Only Relative Differences Matter
      - Stacked 100% Area Chart
    - Relative and Absolute Differences Matter
      - Stacked Area Chart
- Static
  - Simple Share of Total
    - Pie Chart
  - Accumulation or Subtraction to Total
    - Waterfall Chart
  - Components of Components
    - Stacked 100% Column Chart with Subcomponents

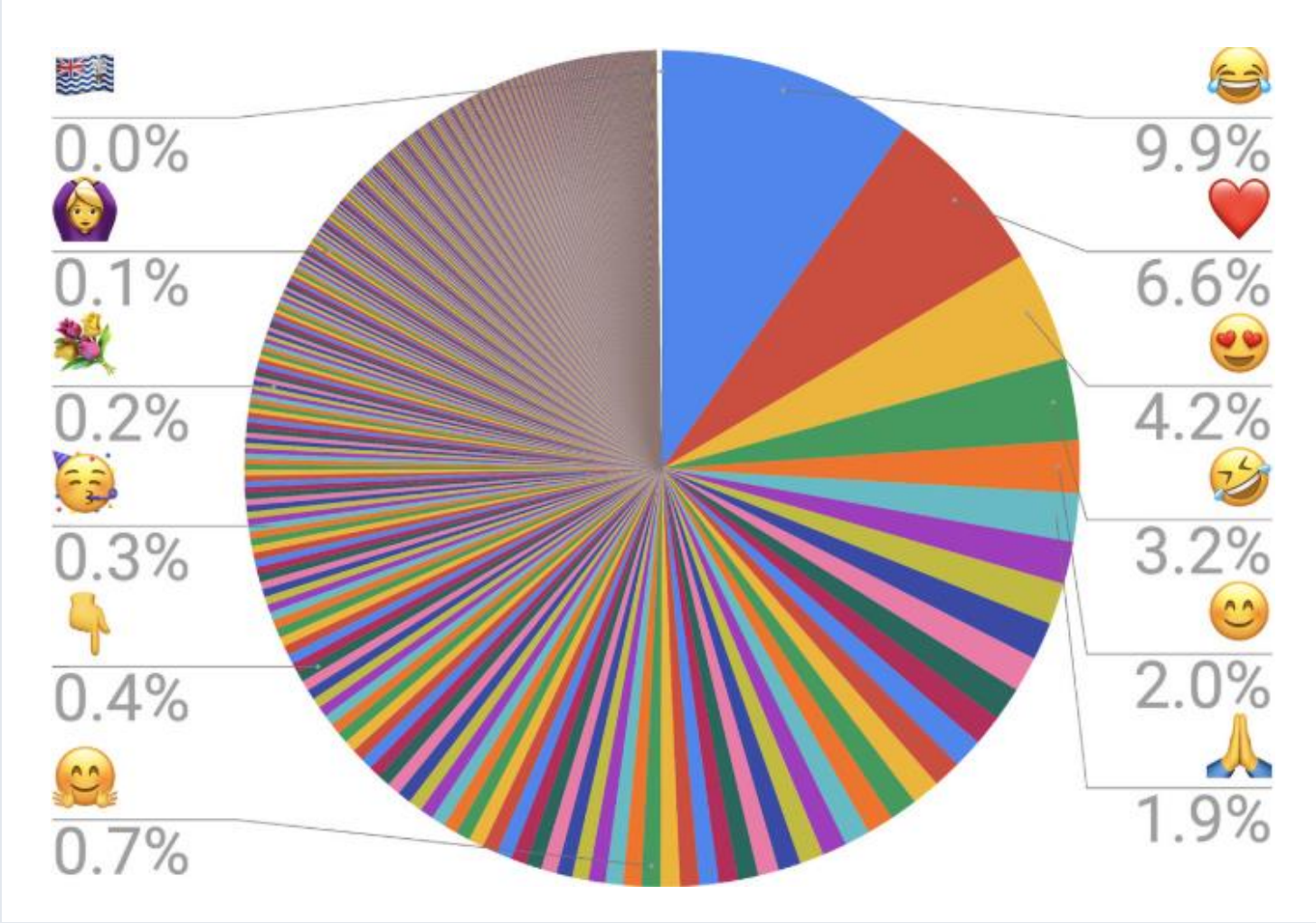
One Dataset, Visualized 25 Ways

An example of many different visualization types in action. See https://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/



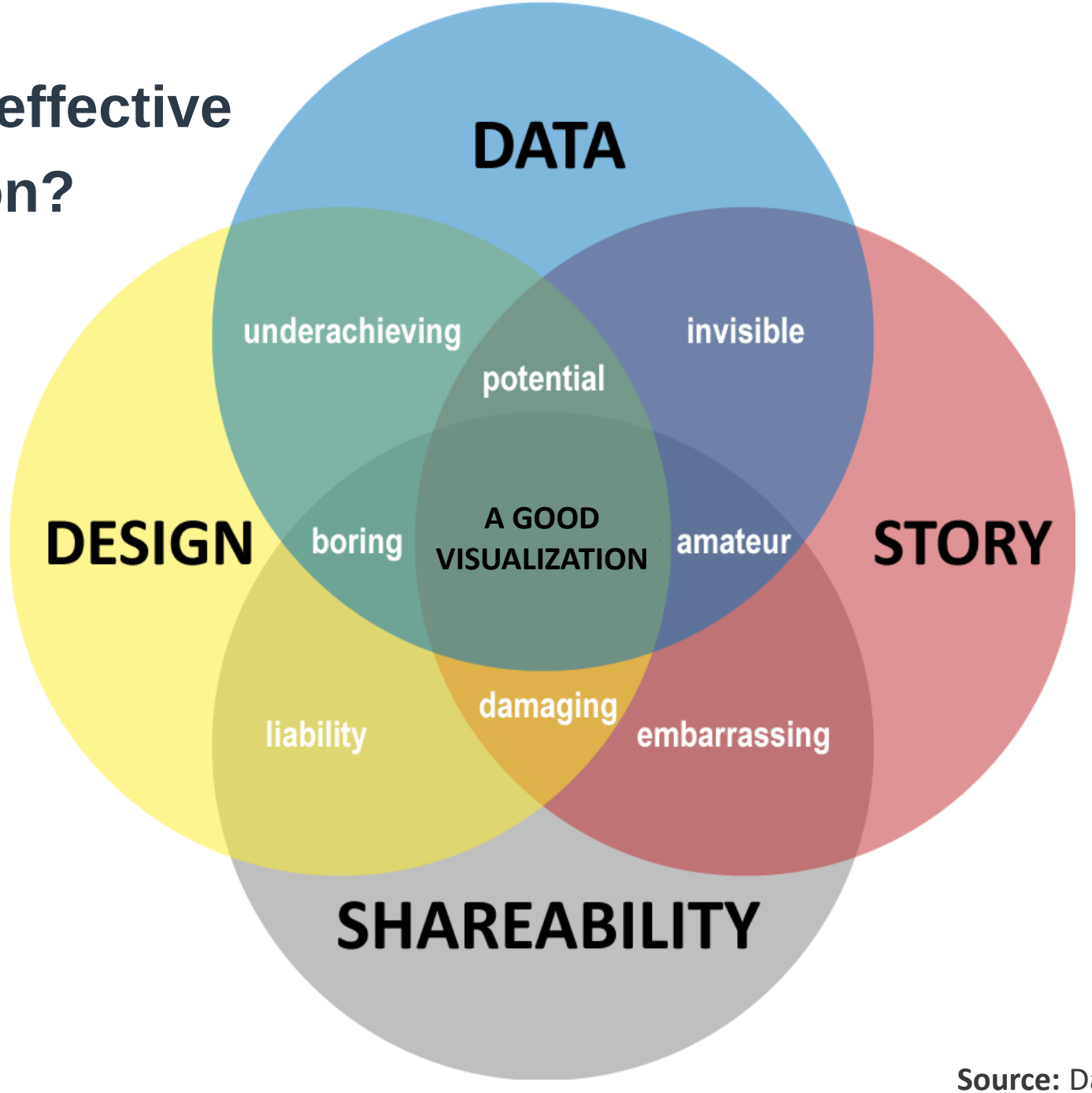
Are some chart types better than others?

Yes and no…



Source: viz.wtf 12/5/2019

What makes an effective data visualization?



Source: Daniel Zeevi (2013), Dashburst

What makes an effective data visualization?

Cole Knaflic's 6-step process:

1. Understand the context
2. Choose an appropriate visual display
3. Eliminate clutter
4. Focus attention where you want it
5. Think like a designer
6. Tell a story

Source: http://www.storytellingwithdata.com/

Ticket Trends: original graphic

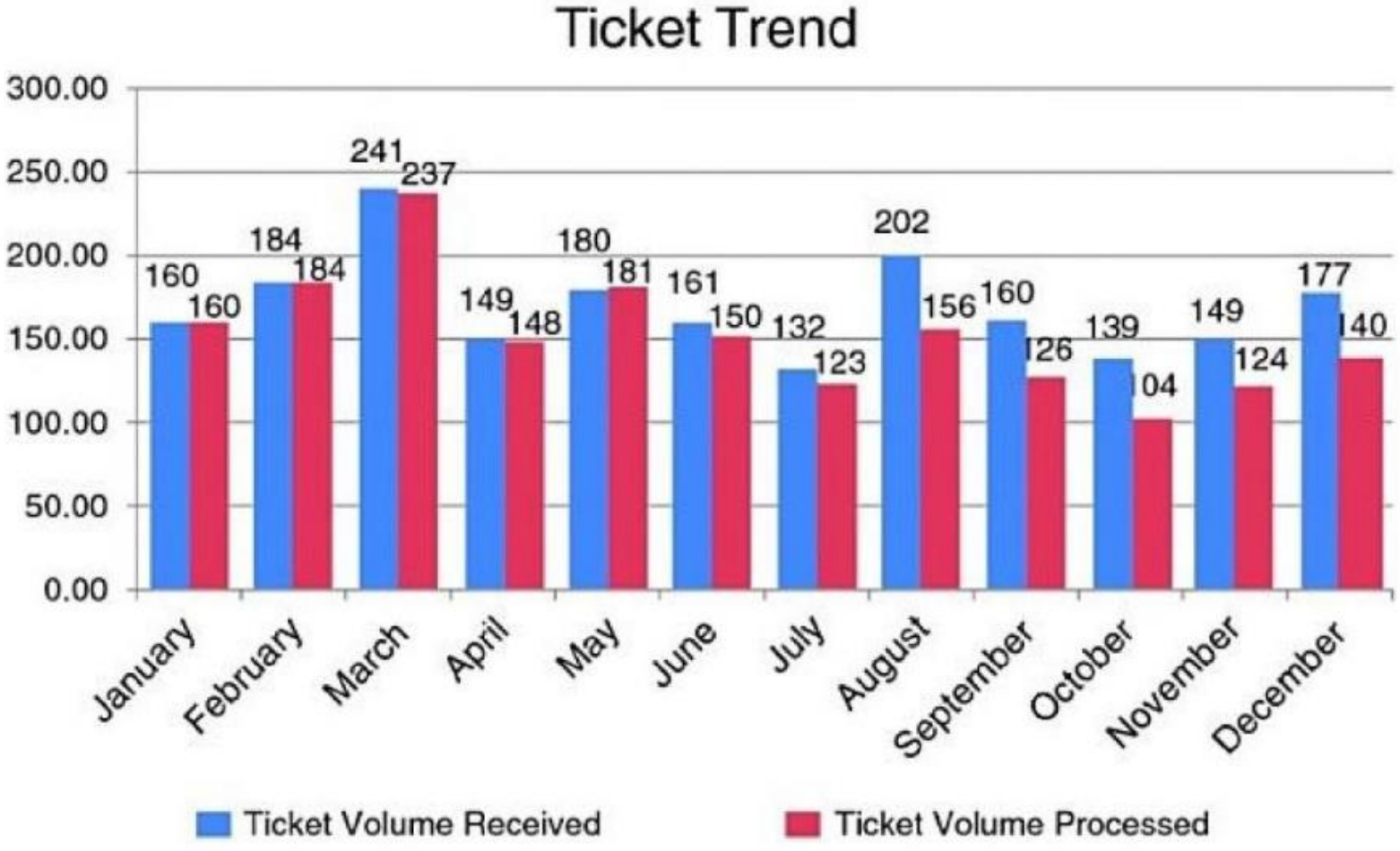


FIGURE 0.2 Example 1 (before): showing data

Source: Knaflic, Cole. Storytelling with Data: A Data Visualization Guide for Business Professionals, Wiley, © 2015.

Ticket Trends: revised graphic

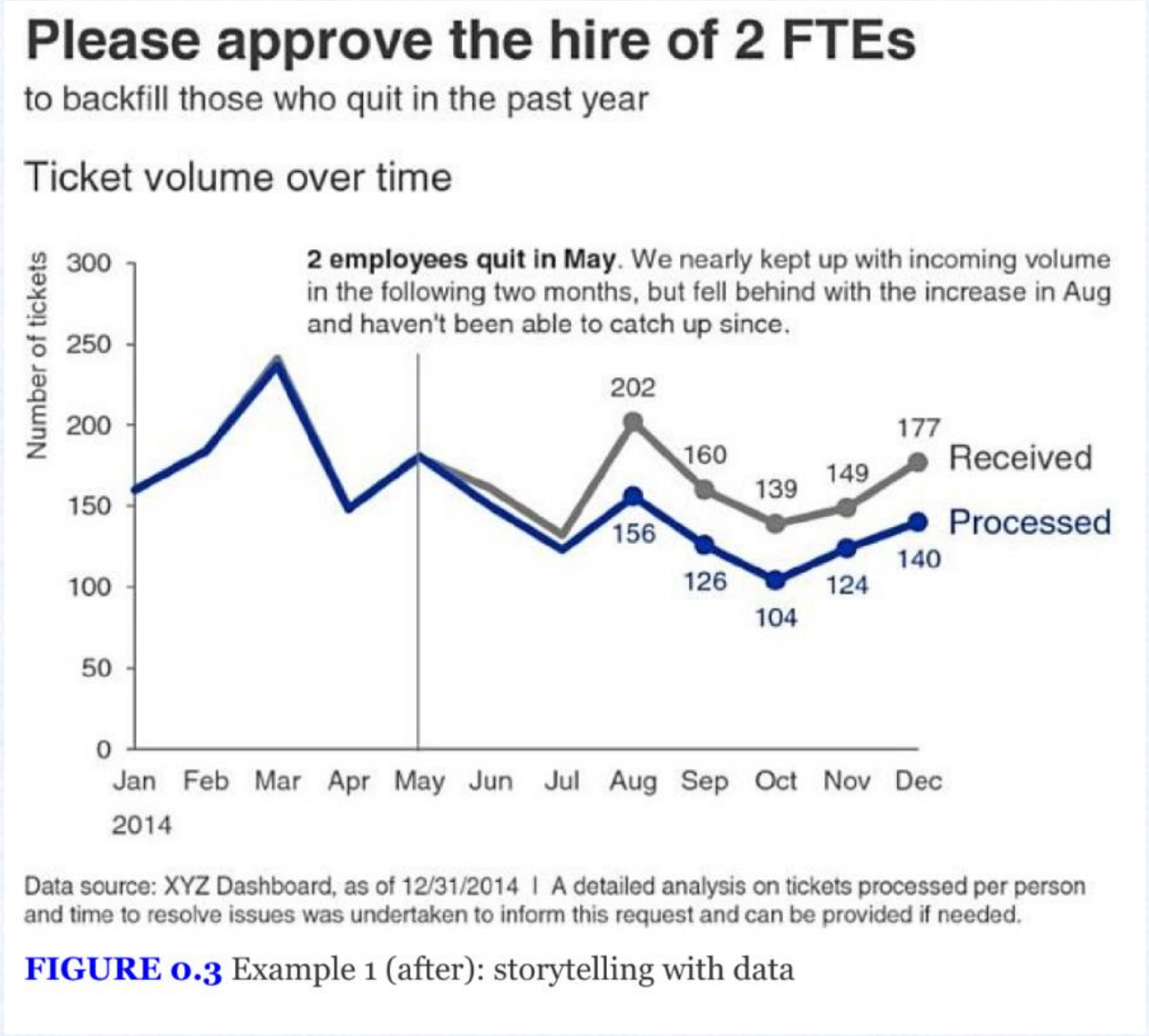
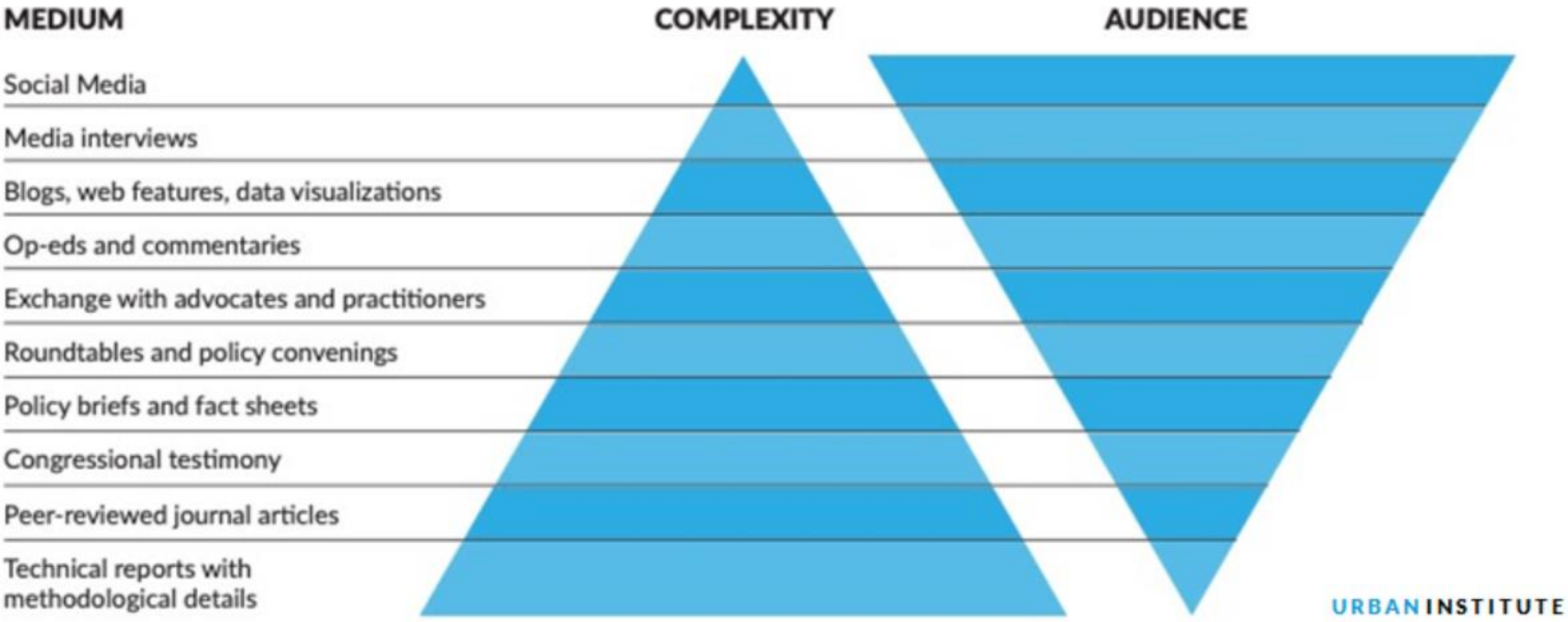


FIGURE 0.3 Example 1 (after): storytelling with data

Source: Knaflic, Cole. Storytelling with Data: A Data Visualization Guide for Business Professionals, Wiley, © 2015.

KNOW YOUR AUDIENCE



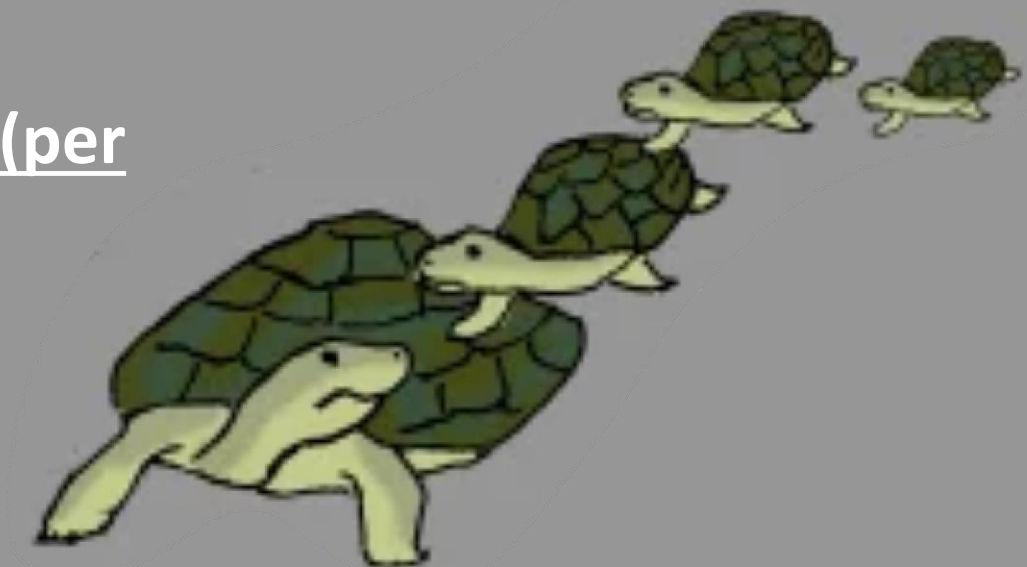
Source: https://www.urban.org/urban-wire/use-pyramid-philosophy-better-communicate-your-research

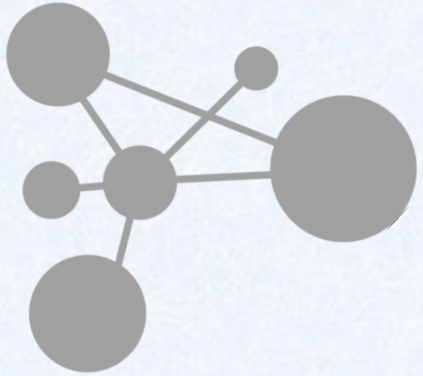
Data and data sets are not objective; they are creations of human design. We give numbers their voice, draw inferences from them, and define their meaning through our interpretations.

-Kate Crawford, 2013

Sample Visualizations

- **Seven endangered species that could (almost) fit in a single train carriage**, by Mona Chalabi
- **Your Contribution to the California Drought**, in the New York Times
- **What do colors mean and how do they appeal to buyers**? Designed by filwebasia, see US
- **Incarceration Rates (2016) in State Prisons (per 100,000)** by Chris Love





Visualization Demos in R and Tableau

Additional Resources

- Tableau Public Gallery: https://public.tableau.com/en-us/gallery
- Storytelling with Data: http://www.storytellingwithdata.com/
- FlowingData blog: https://flowingdata.com/
- Nightingale data visualization journal: https://medium.com/nightingale
- Kantar's Information is Beautiful awards: https://www.informationisbeautifulawards.com/
- Northwestern University's Visual Thinking Lab: https://visualthinking.psych.northwestern.edu/