



INSTITUTE FOR DEFENSE ANALYSES

## **Visualizing Data: I don't Remember That Memo, but I Do Remember that Graph**

Heather Wojton Project Leader

Matthew Avery  
Andrew Flack  
Brian Vickers

September 2020

Approved for Public Release.  
Distribution Unlimited.

IDA Document NS D-14363

Log: H 2020-000362

INSTITUTE FOR DEFENSE ANALYSES  
4850 Mark Center Drive  
Alexandria, Virginia 22311-1882



The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

#### About This Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-19-D-0001, Task BD-09-2299, BD-9-2299 TestSci App for the Office of the Director, Operational Test and Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

#### Acknowledgments

The IDA Technical Review Committee was chaired by Mr. Robert R. Soule and consisted of Dr. Dean Thomas, Dr. Jonathan Bell, Dr. Kristina Carter, Dr. Edward Beall, and Dr. Holly Szumila-Vance from the Operational Evaluation Division.

#### For more information:

Heather Wojton, Project Leader  
[hwojton@ida.org](mailto:hwojton@ida.org) • (703) 845-6811

Robert R. Soule, Director, Operational Evaluation Division  
[rsoule@ida.org](mailto:rsoule@ida.org) • (703) 845-2482

#### Copyright Notice

© 2020 Institute for Defense Analyses  
4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 [Feb. 2014].

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-14363

## **Visualizing Data: I don't Remember That Memo, but I Do Remember that Graph**

Heather Wojton Project Leader

Matthew Avery  
Andrew Flack  
Brian Vickers



# Executive Summary

---

Graphs and plots are critical tools for conveying information in briefings or reports. Understanding what makes a visual display clear, appealing, and convincing helps analysts build more compelling products. The attached slides accompany a one-hour course developed by IDA on general concepts and best practices for data visualization.

## A. Goals and Challenges of Data Visualization in Operational Testing

The Director, Operational Test and Evaluation (DOT&E), uses data visualization in a variety of contexts, including test planning, test evaluation, emerging results briefs (ERBs), modeling and simulation, and conveying information to decision-makers. Given the scope of duties and systems under evaluation, it is difficult to give step-by-step advice that addresses all considerations. This training describes goals to keep in mind when creating visualizations, principles to apply across plots and graphs, and guidance on styling graphical elements.

## B. Goals of Visual Design

Data visualizations should be driven by two goals.

### 1. Visualizations should have a purpose and clear point

Before creating a visual, analysts should ask themselves whether it is necessary and, if it is, what information they want the visual to convey. This process may take time, shifting from prototyping and exploratory looks at data early in the process to clearer and more pointed figures related to statistical models later in the process. Final figures should use the ways in which the data are presented and the annotations are applied to emphasize their points.

### 2. Visualizations should be tailored to your audience

The best format and appropriate level of detail will depend on your audience. For example, if a statistician or subject matter expert is trying to evaluate a test and draw conclusions, they may expect to see specific data points and overall trends. However, if the same data are being presented to a decision-maker, it may

be better to emphasize summary values and levels of uncertainty.

Additionally, visualizations may be seen by individuals from a variety of backgrounds. Solicit feedback from a range of peers, subject matter experts, and other stakeholders to ensure that the message is being conveyed clearly and broadly.

## C. Principles of Data Visualization Design

Keep four design principles in mind when building visualizations:

- Choose a sensible display that incorporates a viewer’s expectations and variables that connect back to the real world.
- Guide the viewer with consistent colors, linetypes, and other variable encodings throughout the product.
- Simplify and break apart pieces of complex models to make them more understandable.
- Make sure each graph has a clear takeaway and use the visual’s details to emphasize this message.

## D. Graphical Elements

Additionally, improve graphical elements like titles, annotations, axes, and legends by keeping the following guidance in mind:

- Titles should specify the takeaway from the figure, instead of just describing what data are presented (e.g., “False alarm rate is 70 percent higher with less experienced operators”).
- Captions (smaller text below or to the side of a figure) should give details about the figure so that readers do not have to search the main text.
- Add annotations (in-figure text) to make graphs more interpretable (e.g., place legend labels next to encodings to make it easier to identify each encoding).
- Avoid dual-axes whenever possible.
- Use more accurate encodings such as position and length as opposed to less accurate encodings like volume, curvature, and shading.
- Match color palettes to the type of data being displayed. Use a sequential palette for continuous data moving from a low to high level; use a diverging palette when the data contain a meaningful middle point; and use a qualitative palette for categorical data.
- Whenever possible, use color palettes that preserve important characteristics when printed in grayscale or seen by colorblind viewers.

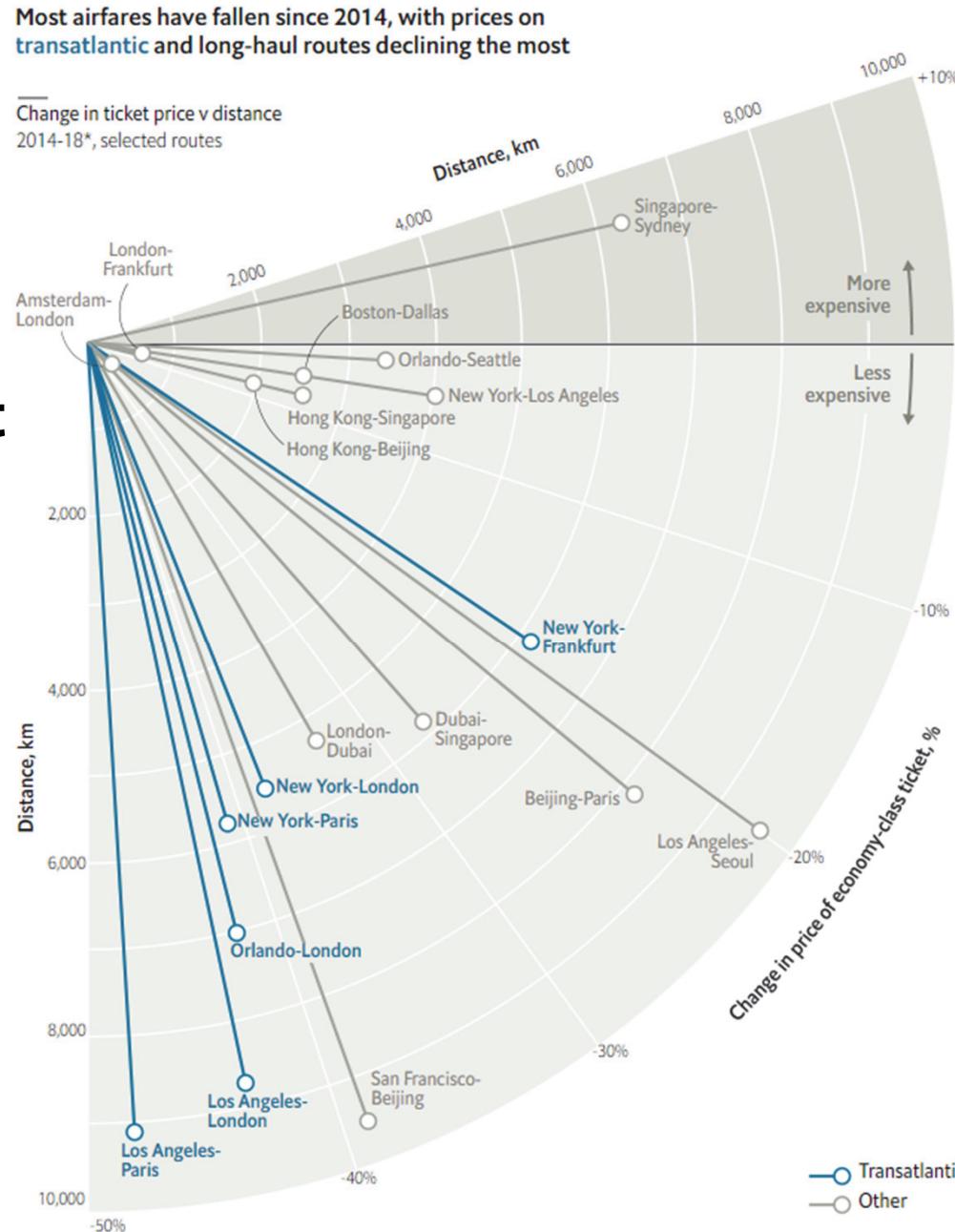




# Visualizing Data: I Don't Remember That Memo, but I Do Remember That Graph

October 25, 2019

Brian Vickers  
Andrew Flack  
Matt Avery



Institute for Defense Analyses  
4850 Mark Center Drive • Alexandria, Virginia 22311-1882

\*Comparing equivalent quarters

Source: "Why Ticket Prices on Long-Haul Flights Have Plummeted," *The Economist*, Dec 8, 2018

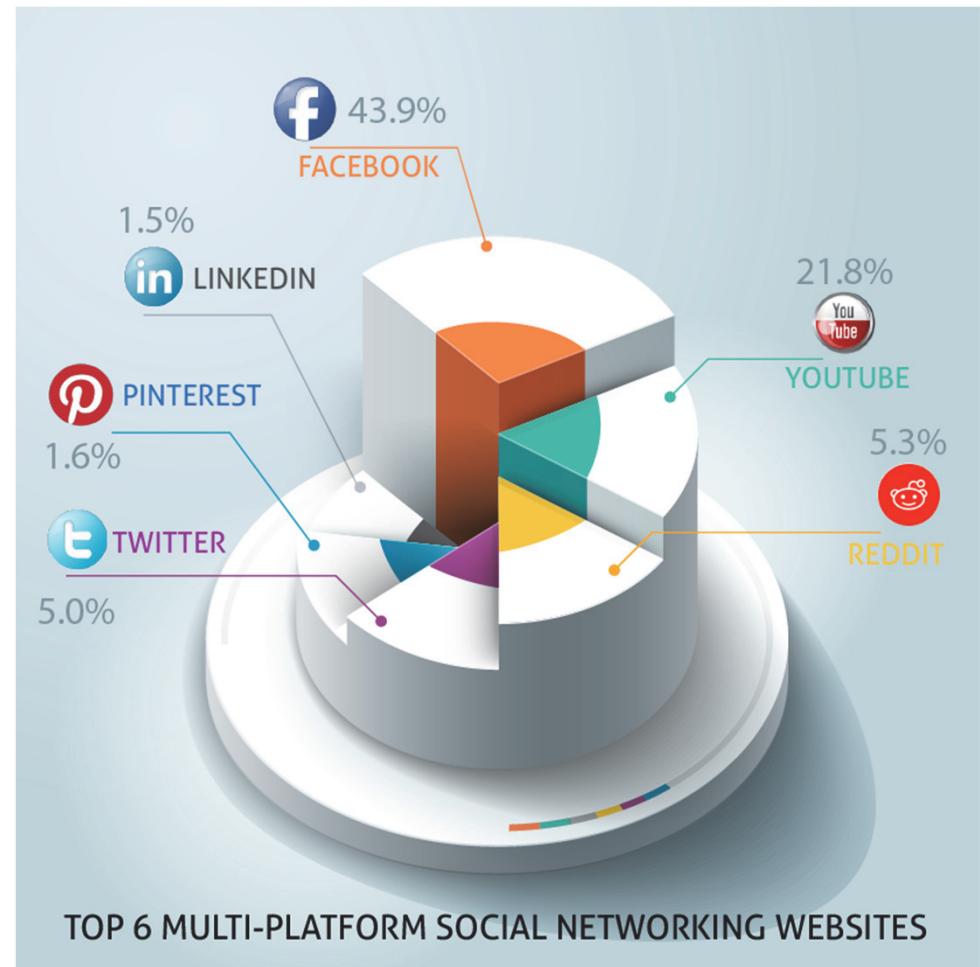


What are your takeaways?

What is wrong?

How would you fix it?

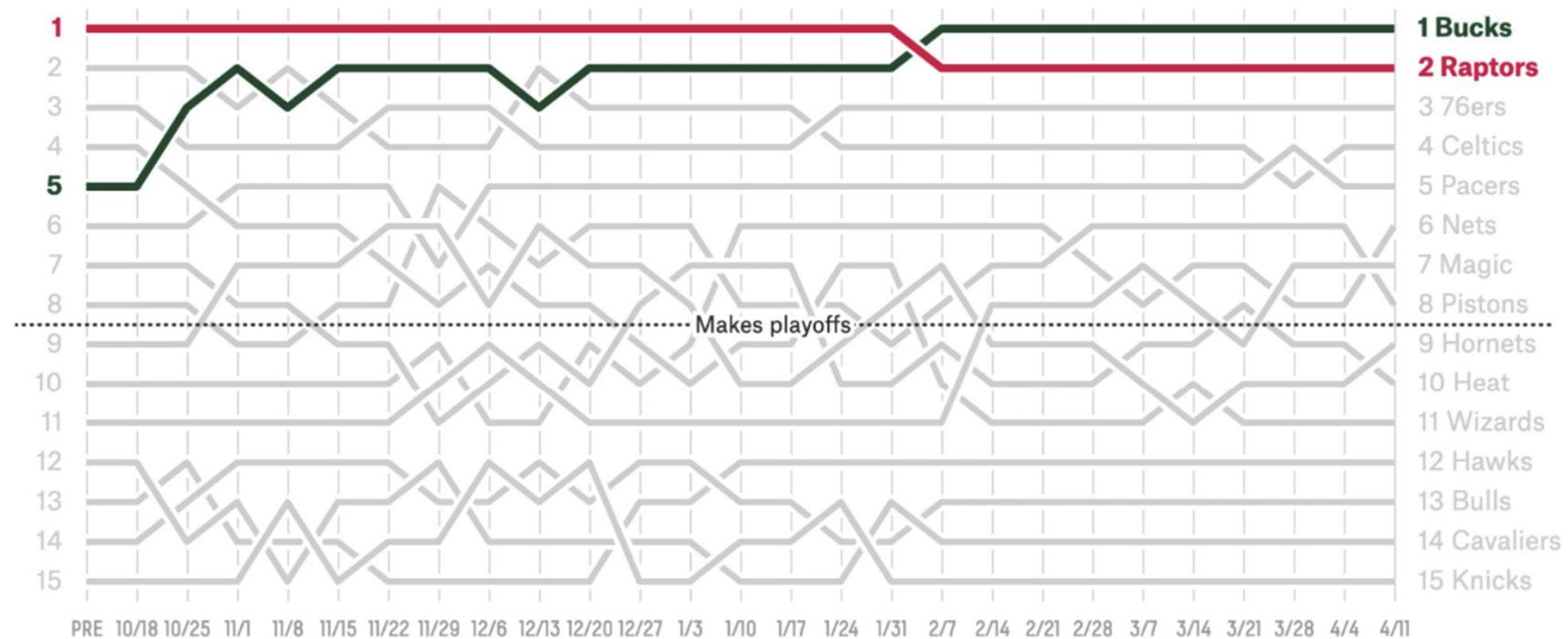
A lot of this is intuitive



Source: Lao (2019)

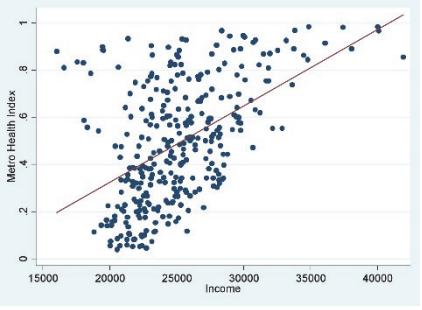
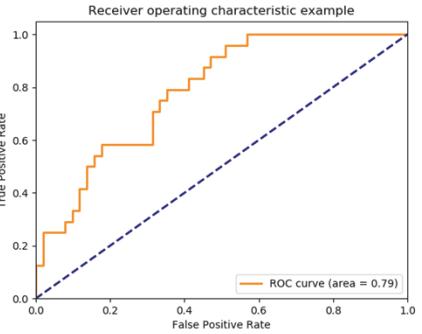
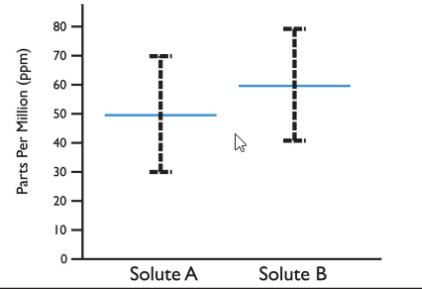
## The Bucks and Raptors jockey for No. 1

Projected playoff seeds according to FiveThirtyEight's NBA projections for Eastern Conference teams, by week



Source: "The Story of the NBA Regular Season in 9 Charts," *FiveThirtyEight*, April 11, 2019

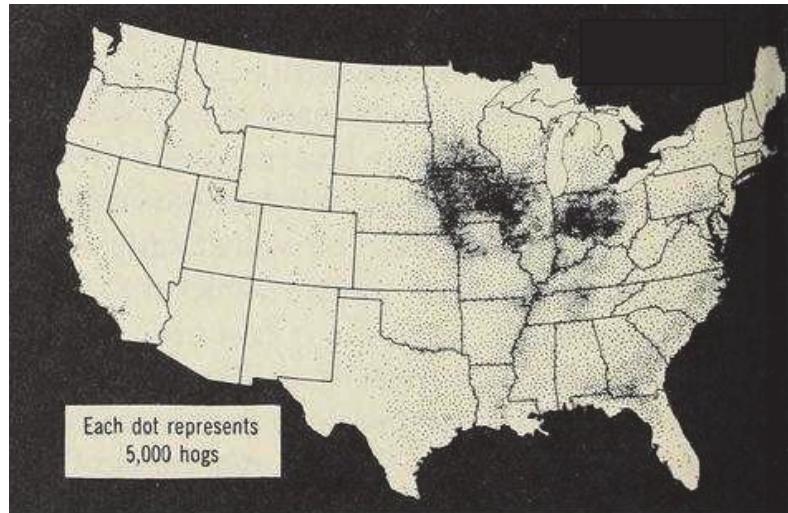
# Think about your audience

| Audience                              | General Interests                               | Example   |
|---------------------------------------|---|---|
| Statisticians, subject matter experts | Gauging analyses and conclusions; the details   |  A scatter plot with 'Income' on the x-axis (ranging from 15,000 to 40,000) and 'Metro Health Index' on the y-axis (ranging from 0 to 1). Numerous blue data points show a positive correlation, with a red regression line drawn through them.  |
| Program managers, consultants         | Set of reliable takeaways; what to focus on/fix |  A Receiver Operating Characteristic (ROC) curve plot. The x-axis is 'False Positive Rate' and the y-axis is 'True Positive Rate', both ranging from 0.0 to 1.0. A dashed diagonal line represents a random classifier. An orange step-like line represents the ROC curve, which starts at (0,0), goes up to approximately (0.1, 0.2), then to (0.2, 0.5), (0.3, 0.7), (0.4, 0.8), (0.5, 0.9), and finally to (1.0, 1.0). A legend indicates 'ROC curve (area = 0.79)'. |
| The Director                          | Information for decision-making                 |  A box plot comparing the concentration of two solutes. The y-axis is 'Parts Per Million (ppm)' ranging from 0 to 80. Two horizontal blue lines represent the median values for Solute A (~50 ppm) and Solute B (~58 ppm). Vertical black dashed lines with caps indicate the range (whiskers) for each. A cursor arrow points towards the Solute B whisker.   |

Sources: [1](#), [2](#), [3](#)

[Note: The figures shown here are for illustrative purposes and do not display program data]

# People read more from graphs than you intend



nemfrog

"Each dot represents 5,000 hogs." *World Geography*. 1948.

lemondemon

untapped infinite hog supply in the ocean

sketchmagetech

Each state is lined with an impenetrable wall.of swine. We are trapped

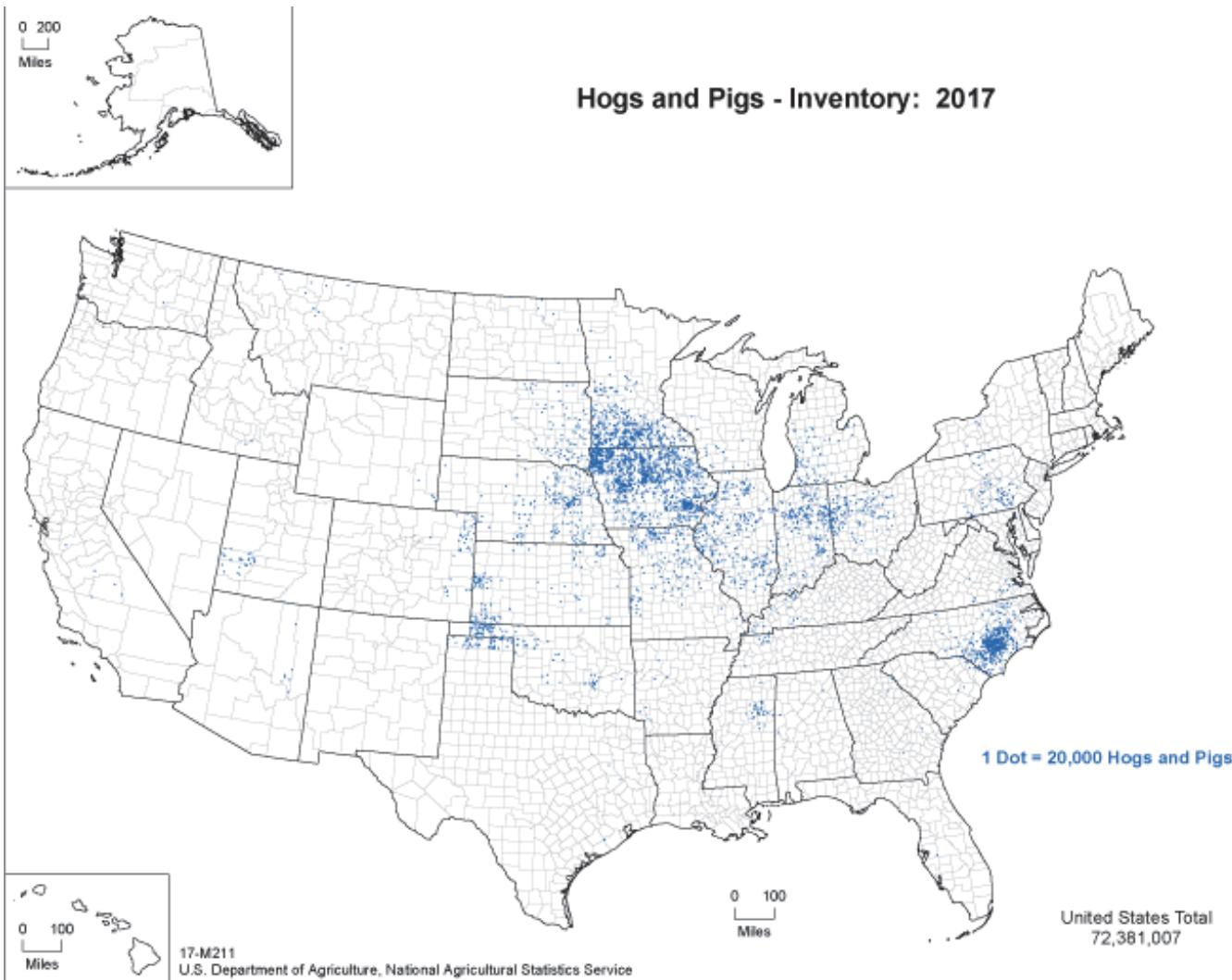
2kittensinacup

We've lost canada and mexico to the hogs already

Source: nemfrog

- Audiences may take away unintended messages
- To ensure clarity, get feedback from multiple sources

# The revised version is much clearer



Source: U.S. Department of Agriculture

# Topics for this presentation

- Deciding to include a figure
- Tables vs. figures
- General concepts
- Importance of context
- Graphical elements
- Colors, shapes, etc.
- Experimental examples
- Let's make some fixes (*additional example*)

# **Do I NEED A FIGURE?**

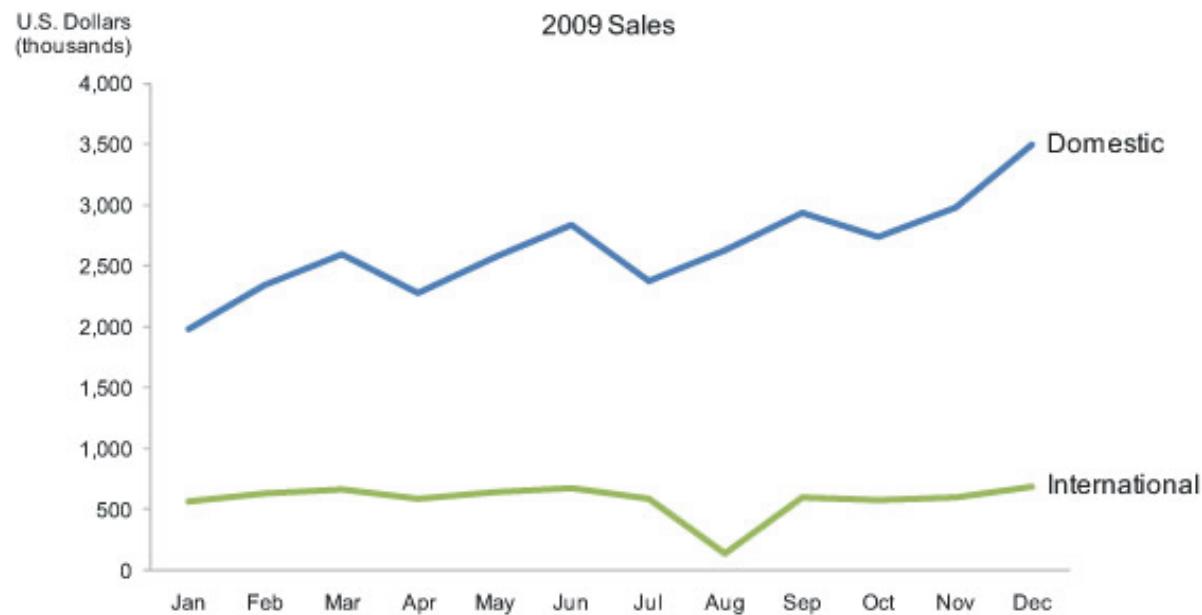
# Do I need a figure?

- If you can make a point in a sentence or two, you (usually) don't need a figure
- Unless
  - Important takeaway
  - Convey (un)certainty

**CAN'T I JUST USE A TABLE?**

# Data communicated visually are interpreted differently than data communicated as text

| 2009 Sales (thousands of U.S. \$) |       |       |       |       |       |       |       |       |       |       |       |       |        |
|-----------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Region                            | Jan   | Feb   | Mar   | Apr   | May   | Jun   | Jul   | Aug   | Sep   | Oct   | Nov   | Dec   | Total  |
| Domestic                          | 1,983 | 2,343 | 2,593 | 2,283 | 2,574 | 2,838 | 2,382 | 2,634 | 2,938 | 2,739 | 2,983 | 3,493 | 31,783 |
| International                     | 574   | 636   | 673   | 593   | 644   | 679   | 593   | 139   | 599   | 583   | 602   | 690   | 7,005  |
| Total                             | 2,557 | 2,979 | 3,266 | 2,876 | 3,218 | 3,517 | 2,975 | 2,773 | 3,537 | 3,322 | 3,585 | 4,183 | 38,788 |



Visual perception is fast and efficient

Source: Stephen Few

# Use tables for exact numbers; appendix/inline otherwise

## Tables

- ✓ Convey specifics
- ✓ #s comparable across dimensions
- ✓ Can create new metrics
- ✓ Allow text data
- ✓ Familiar to everyone
- ✓ Point comparisons easy
- ✗ Trend comparisons difficult
- ✗ Fixed (tabular) formatting
- ✗ Identifying exceptions or outliers is difficult

## Figures

- ✓ Convey summaries
- ✗ Hard to compare color/sizes
- ✗ Must “ballpark” differences
- ✗ Data restrict chart types
- ✗ Some types get confusing
- ✗ Point comparisons harder
- ✓ Trend comparisons easy
- ✓ Extremely flexible formatting
- ✓ Identifying exceptions or outliers is easy

**Tables serve a purpose, but they are not suitable in all cases**

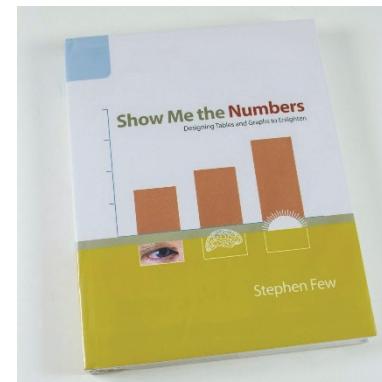
# You can augment tables

- Sparklines can help track trends
- Color scales allow quicker concept mapping
- Book: Steven Few's *Show Me the Numbers*

|          | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 Trend  |
|----------|------|------|------|------|------|---|
| System A | 70   | 73   | 78   | 77   | 88   |  |
| System B | 72   | 73   | 73   | 76   | 70   |  |

| Sales Rep      | Sales Amount \$   | Budget X  |
|----------------|-------------------|---|
| Totals         | 14,778,612,124.42 | 18,355  |
| Deborah Halmon | 693,494,875.76    |  |
| Brenda Gibson  | 734,561,064.70    |  |
| Ronald Milam   | 226,775,743.58    |  |
| Kathy Clinton  | 103,932,257.17    |  |
| John Greg      | 671,882,870.41    |  |
| Amanda Honda   | 1,586,303,262.30  |  |
| John Davis     | 44,604,576.56     |  |
| Dwight Howard  | 533,718,914.69    |  |

| City            | Jan | Feb | Mar | Apr | May | Jun |
|-----------------|-----|-----|-----|-----|-----|-----|
| Barstow         | 80  | 84  | 84  | 97  | 95  | 98  |
| California City | 78  | 86  | 84  | 96  | 98  | 102 |
| Cinco           | 83  | 86  | 86  | 97  | 95  | 103 |
| Hesperia        | 78  | 85  | 87  | 98  | 97  | 102 |
| Lancaster       | 78  | 85  | 86  | 99  | 95  | 101 |
| Mojave          | 82  | 85  | 86  | 98  | 96  | 99  |
| Palmdale        | 81  | 84  | 85  | 97  | 95  | 101 |
| Ridgecrest      | 81  | 87  | 87  | 97  | 96  | 98  |
| Rosamond        | 82  | 86  | 88  | 99  | 97  | 101 |
| Santa Clarita   | 79  | 85  | 87  | 95  | 96  | 103 |



Sources: [1](#), [2](#), [3](#)

Sparkline additional information [1](#), [2](#)

**WHAT GENERAL CONCEPTS SHOULD I  
KEEP IN MIND?**

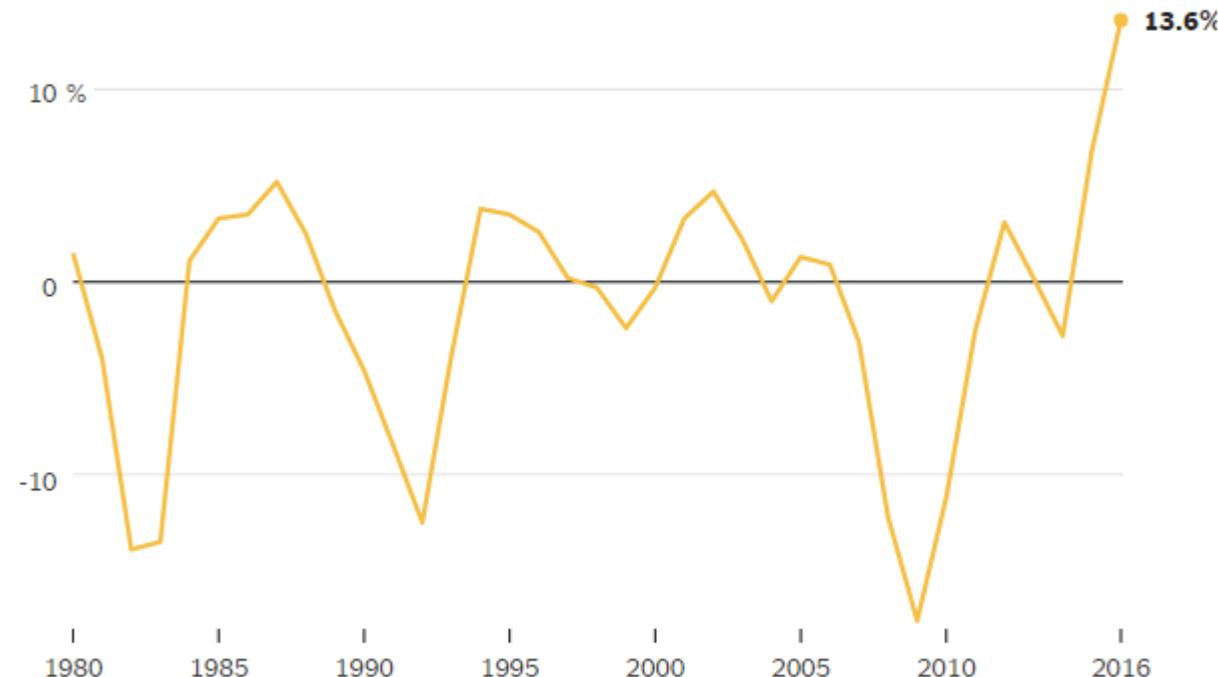
# Keep these four principles in mind

- Choose a sensible display
  - Consider viewer expectations and data types
- Guide the viewer
  - Consistency aids understanding
  - Link 16 is blue in Fig 1; make it blue in all figs
- Simplify or break down complex models
  - Can pull pieces into separate figures
- Clear “so what?”
  - Graphic should make a point

# Rates of traffic fatalities are spiking like never before!

## A Surge in Fatalities

Change in U.S. vehicle deaths over successive two-year periods.



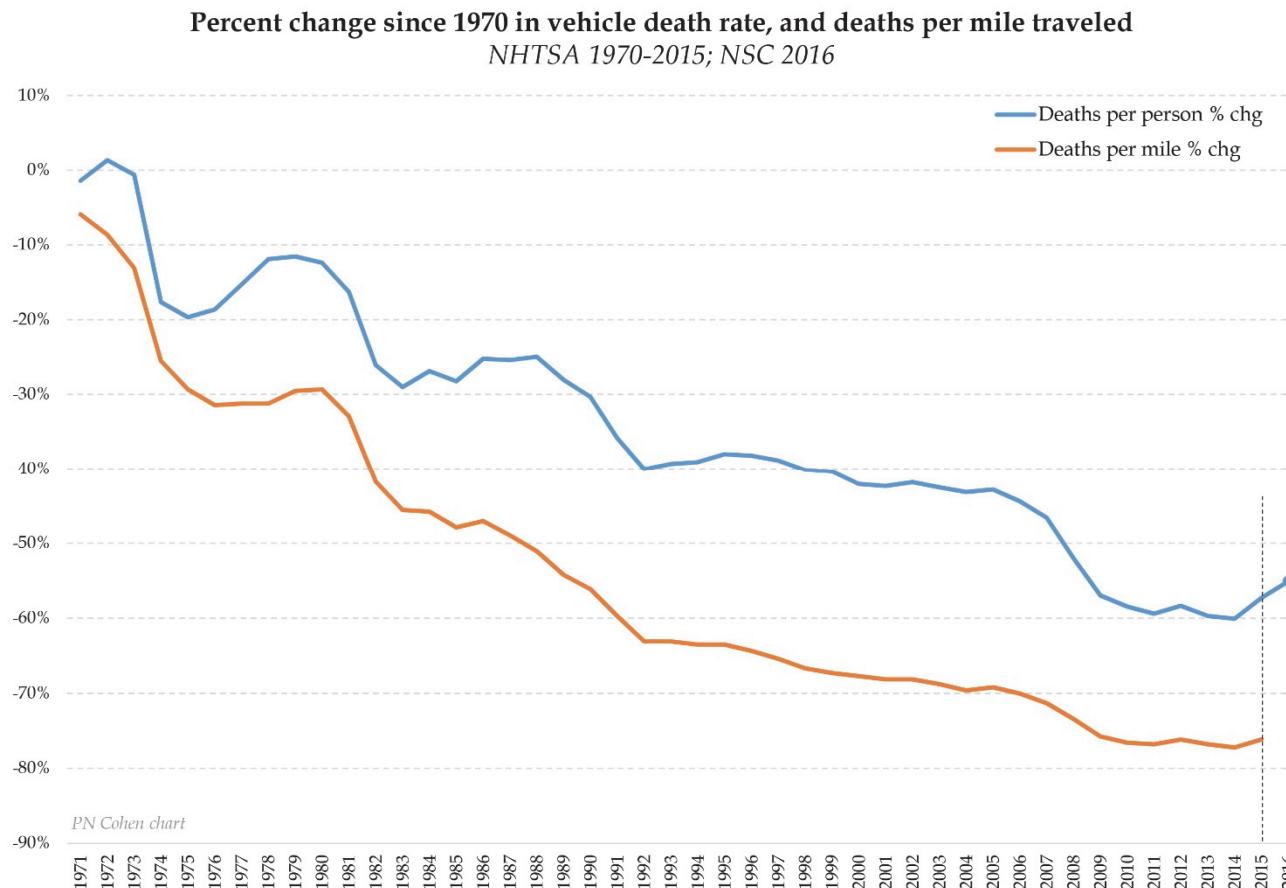
Deaths decline with recessions (such as 2007–9) because driving declines.

Source: National Safety Council

[Choose a sensible display]

Source: David Leonhardt (NYT)

# ...except overall traffic fatalities have dropped significantly over the same time frame



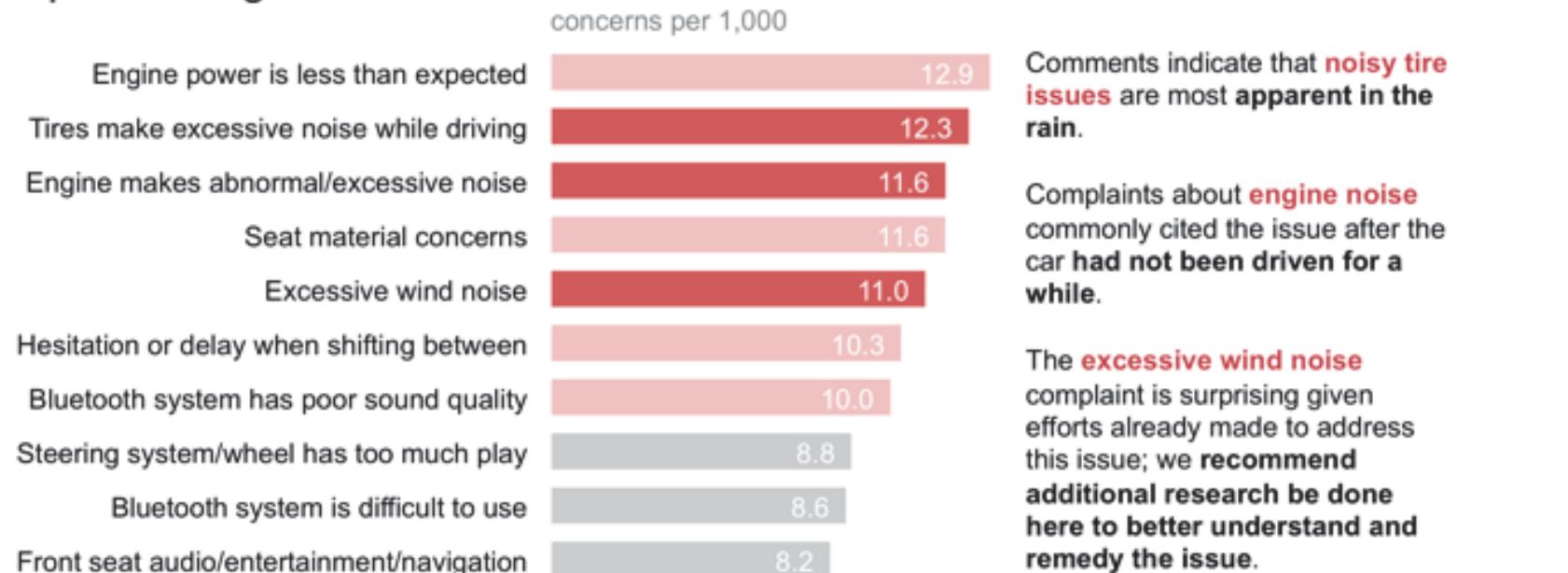
[Choose a sensible display]

Source: Philip Cohen

# Strategic use of color and annotations tells the viewer where to look

Of the top design concerns, three are noise related.

## Top 10 Design Concerns



[Guide the viewer]

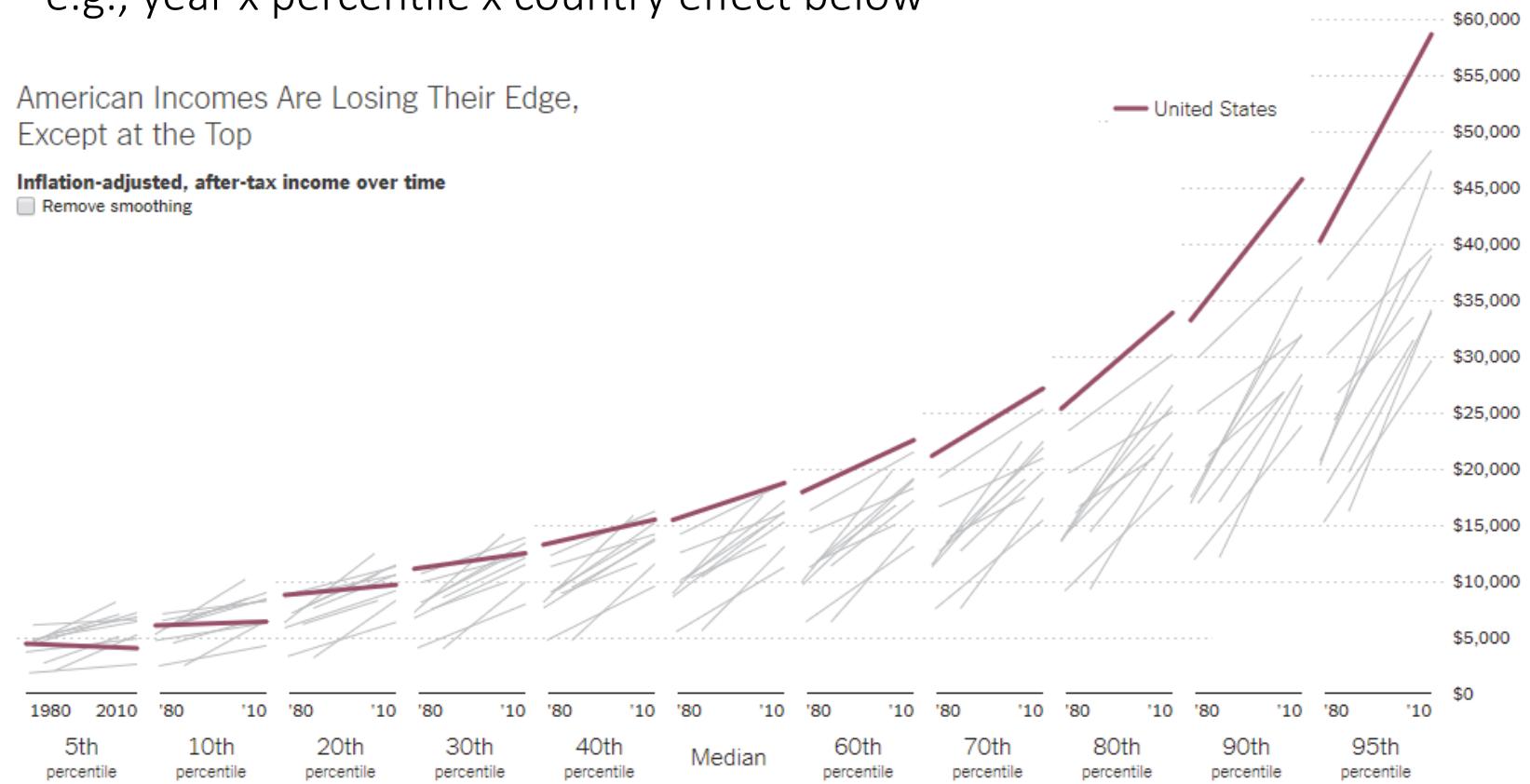
Source: Storytelling with Data

# Simplify complicated models

- Focus on your effect or trend of interest
  - e.g., year x percentile x country effect below

American Incomes Are Losing Their Edge,  
Except at the Top

Inflation-adjusted, after-tax income over time  
 Remove smoothing

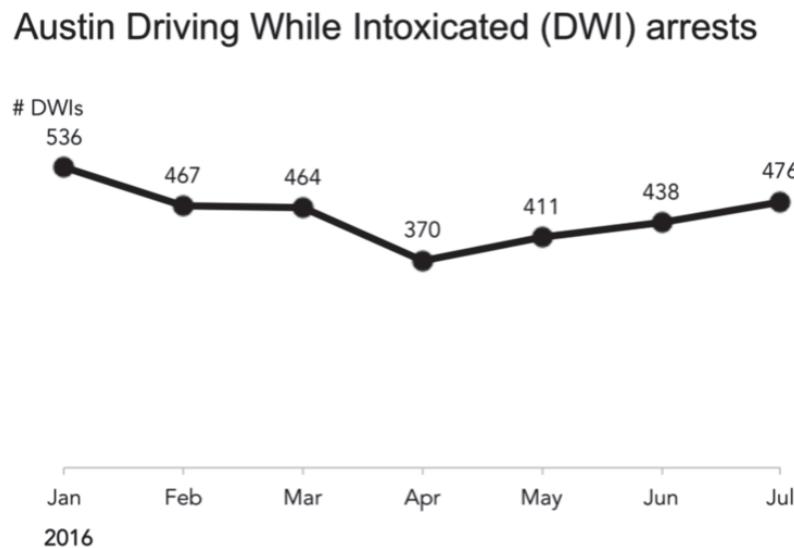


[Simplify]

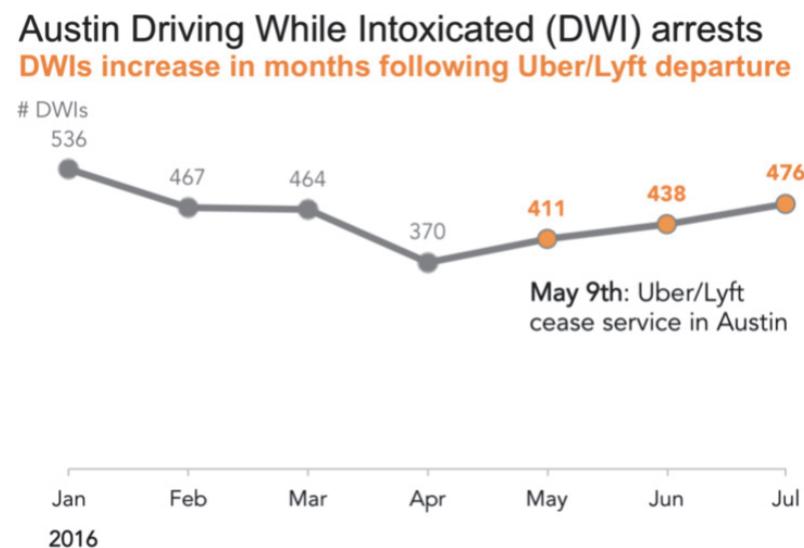
Source: NY Times

# Takeaway is clear; viewer is led to relevant data

Original Version



Better: Added a clear “So What”



[Clear “so what?”]

Source: Storytelling with Data



**BUT AREN'T THERE  
DIFFERENT CONTEXTS?**

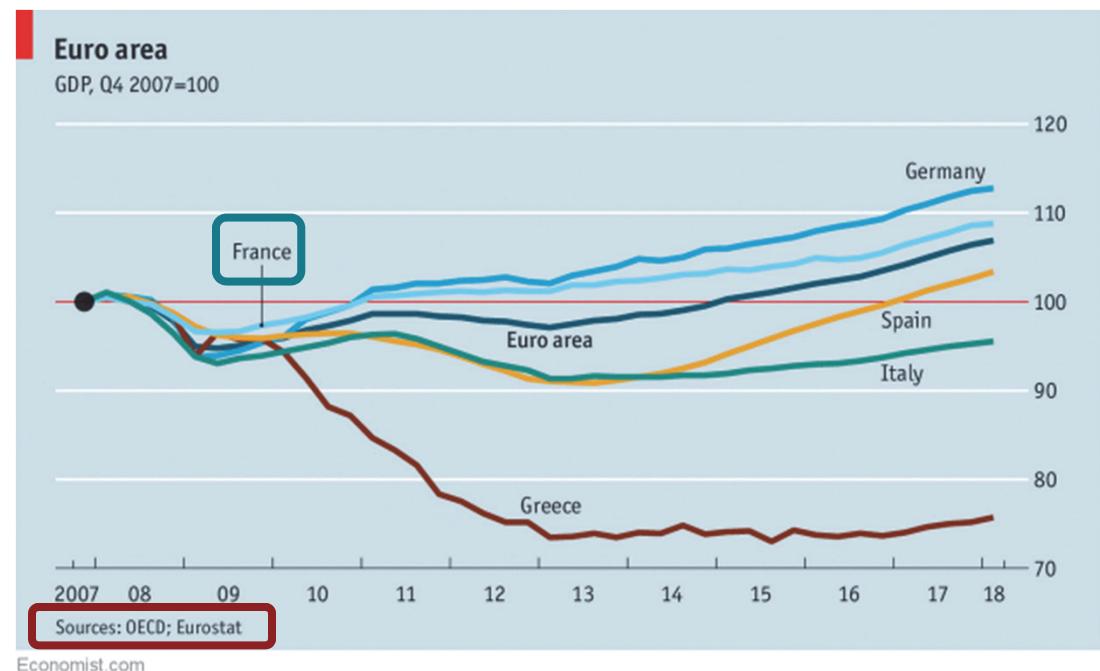
# In the end, it's about storytelling

- We'll differentiate a few different goals
  - Single-chart meeting
  - Make one point vs. make one chart
  - Reports
  - Emerging results briefs

# HOW SHOULD I THINK ABOUT GRAPH ELEMENTS?

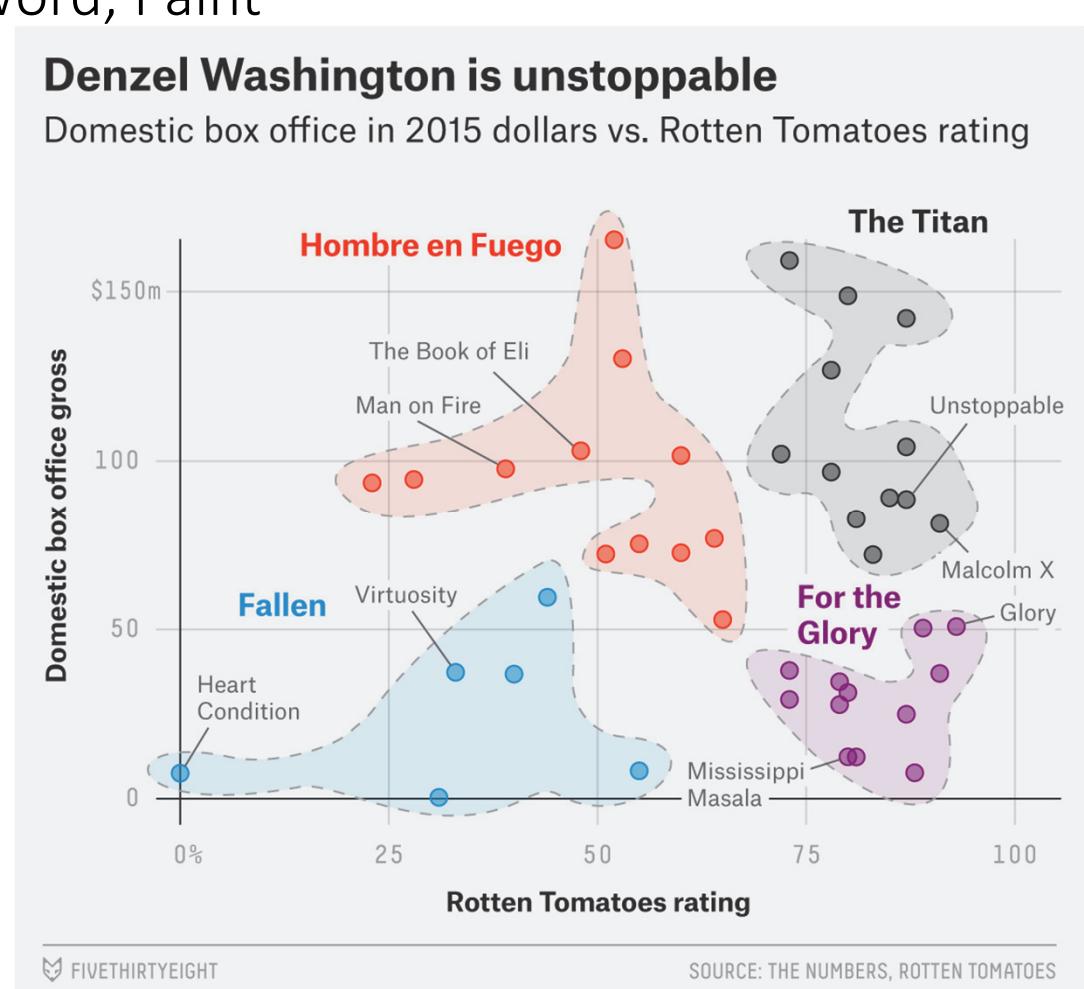
# Send clear messages with chart text

- Title should be the takeaway
  - “False alarm rate is 70% higher with less experienced operators.”
- **Captions** can give readers details so they don’t have to dig through the paper
  - “Error bars represent standard error.”
- **Annotations** are simpler than legends



# Annotations can point out outliers, special cases

- You don't need to annotate everything
- Can edit in PowerPoint, Word, Paint
  - Aim for reproducibility



# If you're an R user...ggplot2: Text vs. label

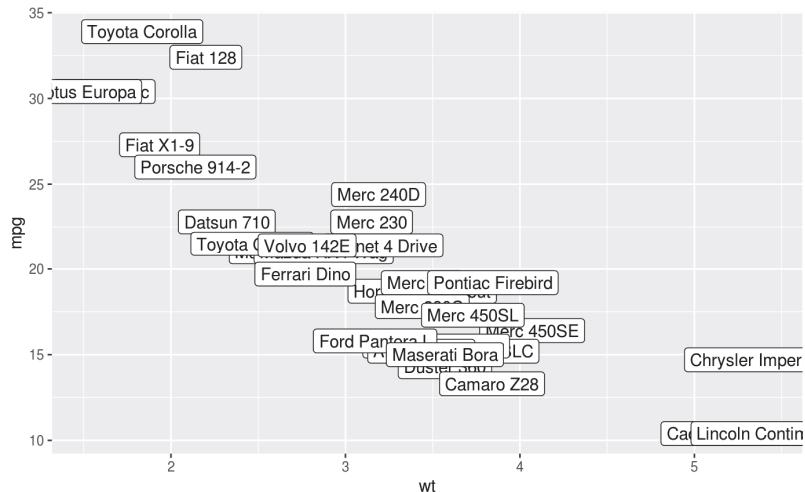
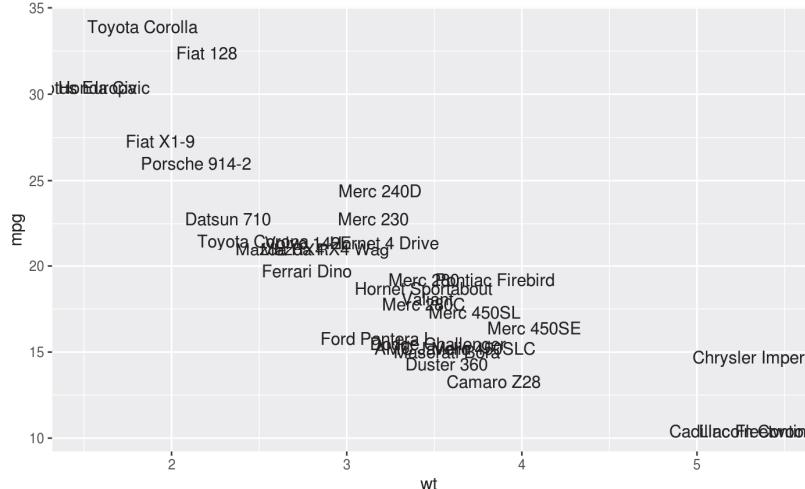
- geom\_text(...):

- Each point labeled,  
no background

- geom\_label(...):

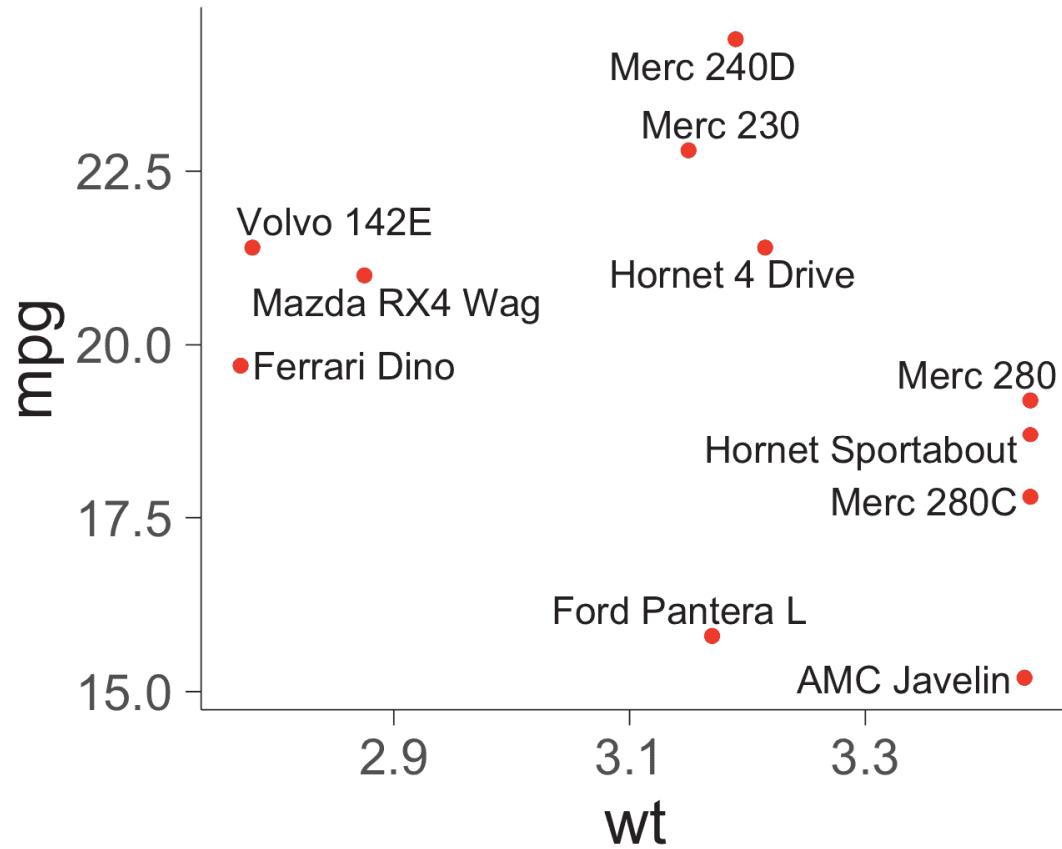
- Each point labeled,  
flexible background

```
check_overlap = TRUE  
Removes random  
overlapping points
```



# The ggrepel package offers more advanced labeling options

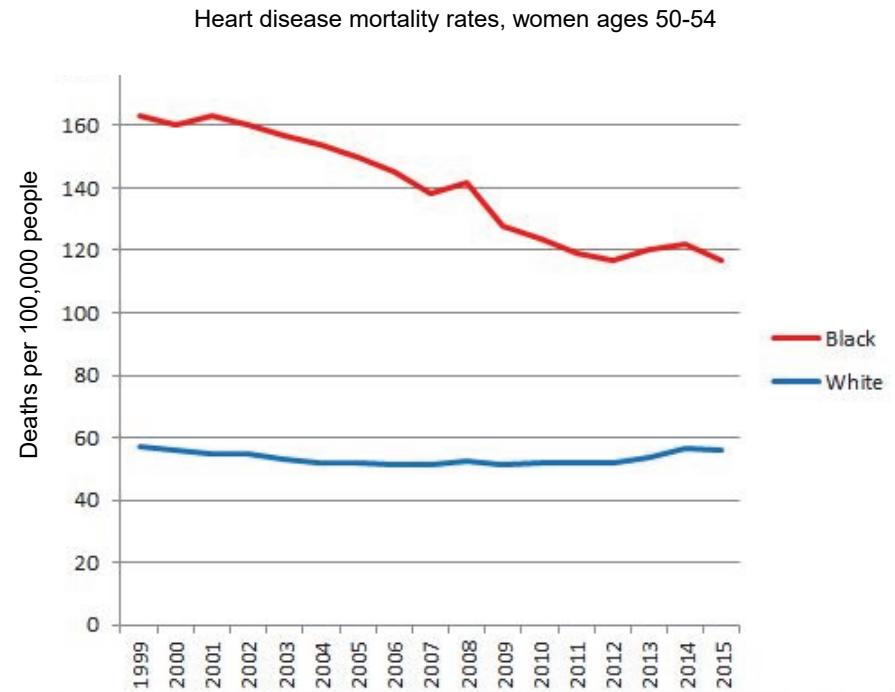
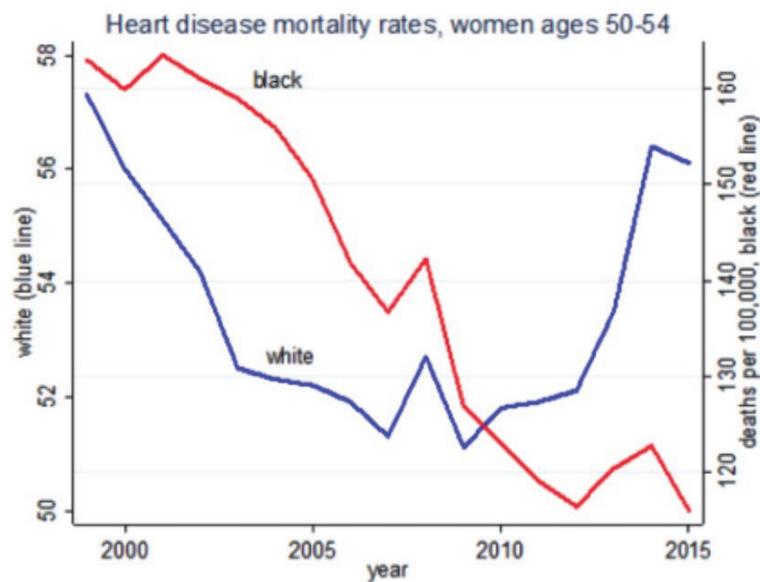
- `geom_text_repel(...)`:



# Make BIG, relatable legends and labels

- In presentations, titles and legend text should be bigger than you originally think necessary
  - Ensures readability throughout an entire room
- Use axes that “make sense”
  - Value of 0 important?
  - Can values be negative (e.g., weights)?
- Avoid dual-axes (misleading)

The figure on the right, plotted on a single y-axis, tells a very different story than the figure on the left

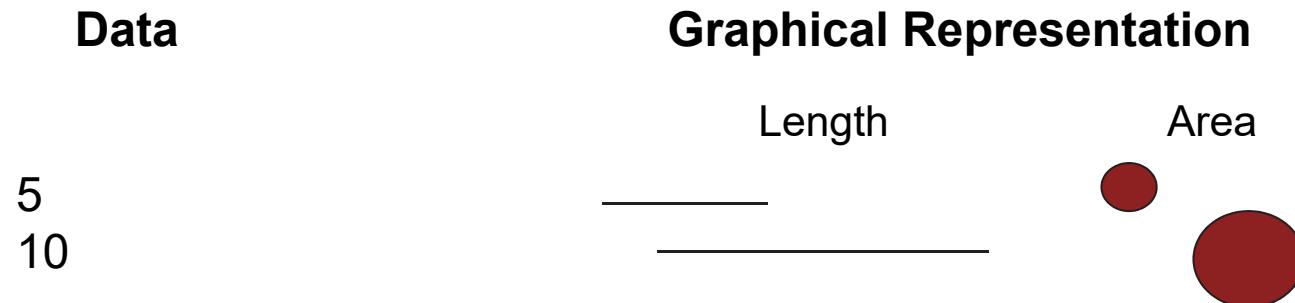


[Source](#): Case and Deaton (2017)

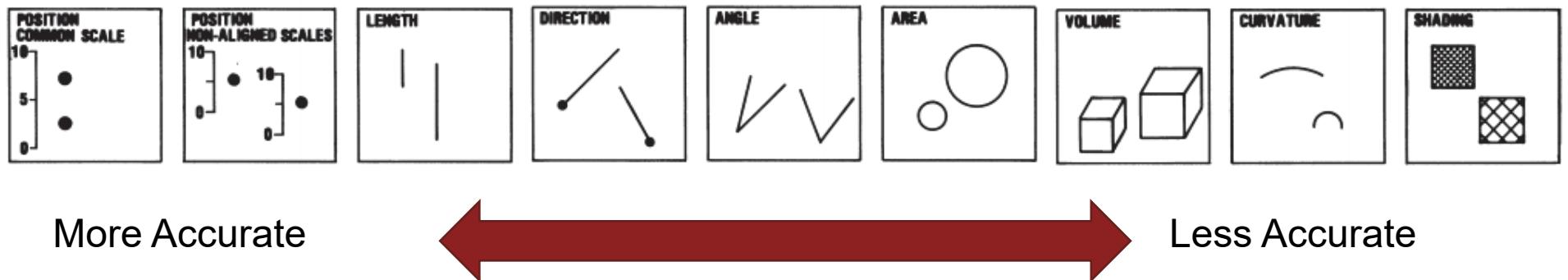


I THOUGHT I WAS GOING TO LEARN  
ABOUT COLORS, SHAPES, ETC.?

# “Encodings” map data to graphical representations



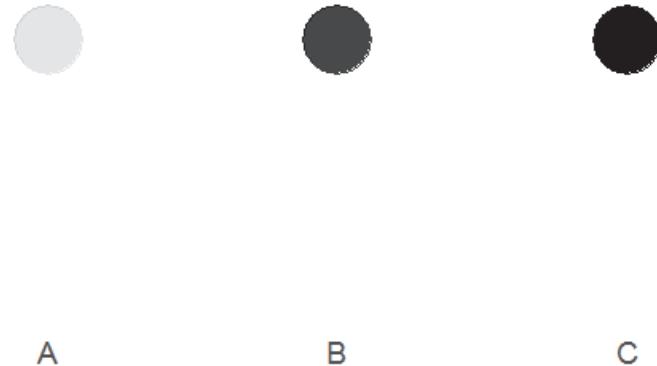
- Some encodings make it easier to **accurately** read data values



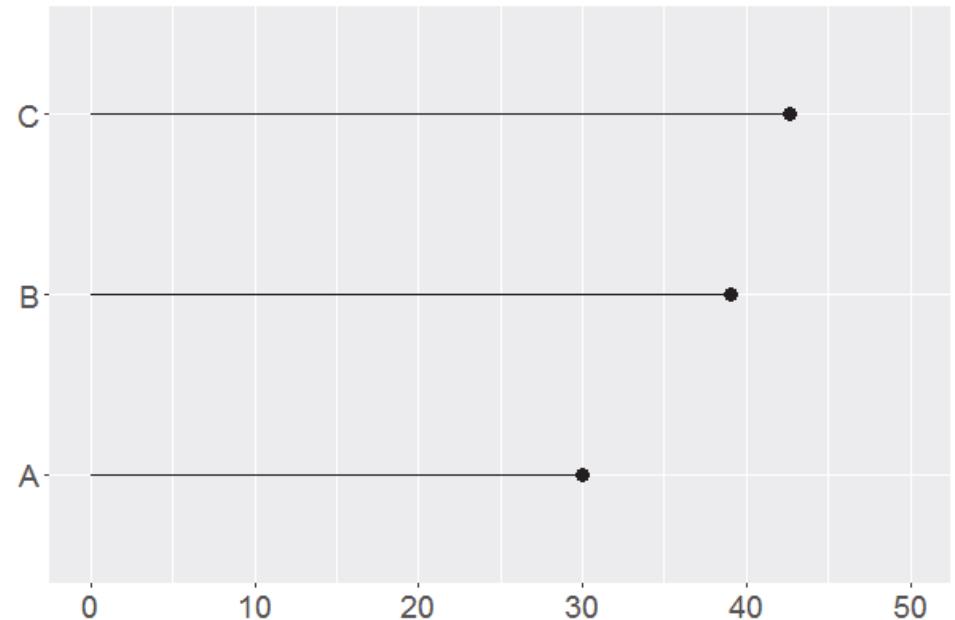
[Source](#): Cleveland and McGill (1984)

# What are the values of B and C relative to A? 20% larger? 50%?

**Shading**



**Position (common scale)**

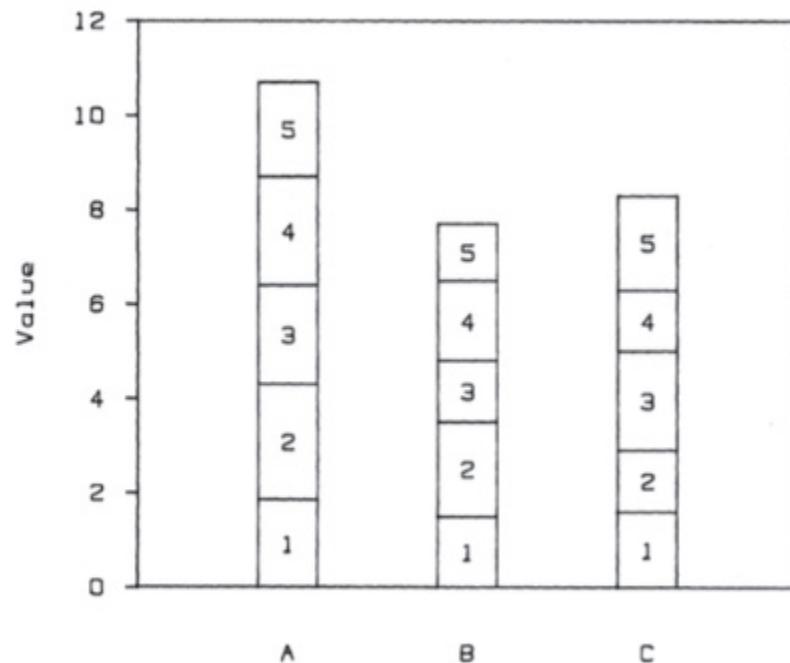


Using **shading** as an encoding (left), it is difficult to tell that the value of C is greater than the value of B, and it's impossible to estimate the values of B and C relative to A.

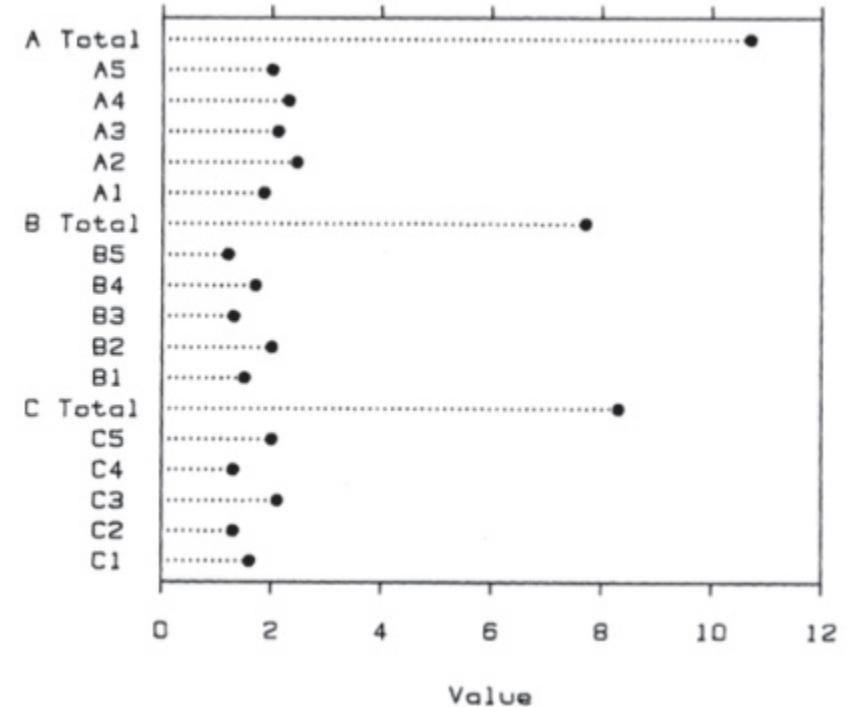
Using **position along a common scale** (right), it is clear that C is greater than B, and we can accurately estimate that B and C are about 30% and 42% greater than A, respectively.

# Which segment is larger, A4 or A2? C3 or C5?

Length



Position (common scale)



Using **length** as an encoding (left), it is difficult to compare the relative sizes of the bar segments. For example, A4 and A2 appear to be about the same size.

Using **position along a common scale** (right), it is clear that A2 is greater than A4.

# Conflicting encodings can cause confusion

The author intended to use **position along a common scale** to encode the sales values (e.g., a bar graph), but the **area** of the logos is also changing.

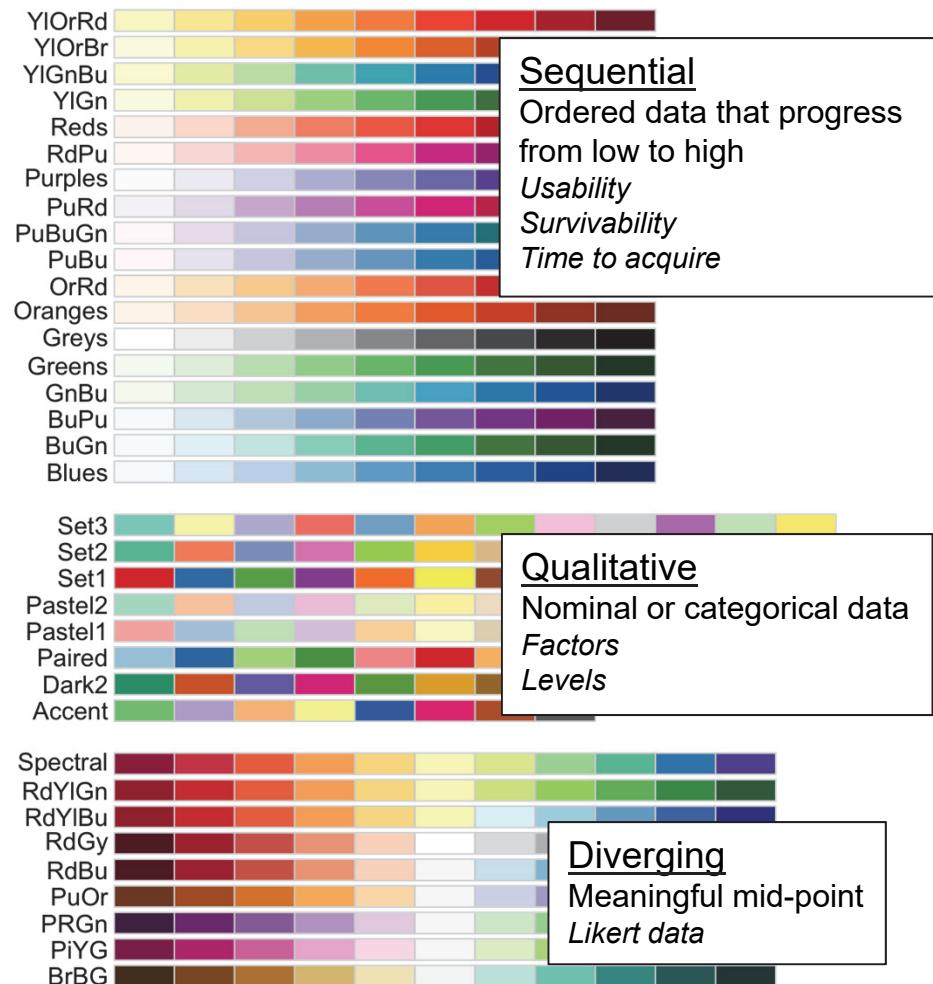


Burger King sales are three times more than Starbucks. The logo is 3x taller, but 9x larger! This may be confusing, or even misleading.

[Source: Business Insider](#)

# Colors have a whole science behind them

- Tie color to variable types
  - Common mistake: Using sequential colors for qualitative data
- Be cautious of default Excel and PowerPoint palettes
  - Printability (grayscale)
  - Colorblindness



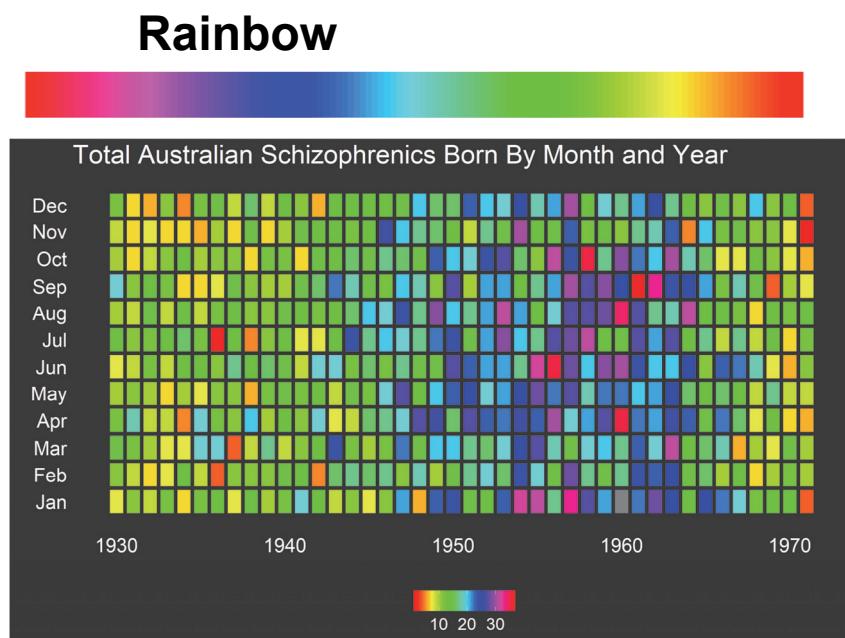
Source: Color Brewer

# Initial considerations when choosing palettes

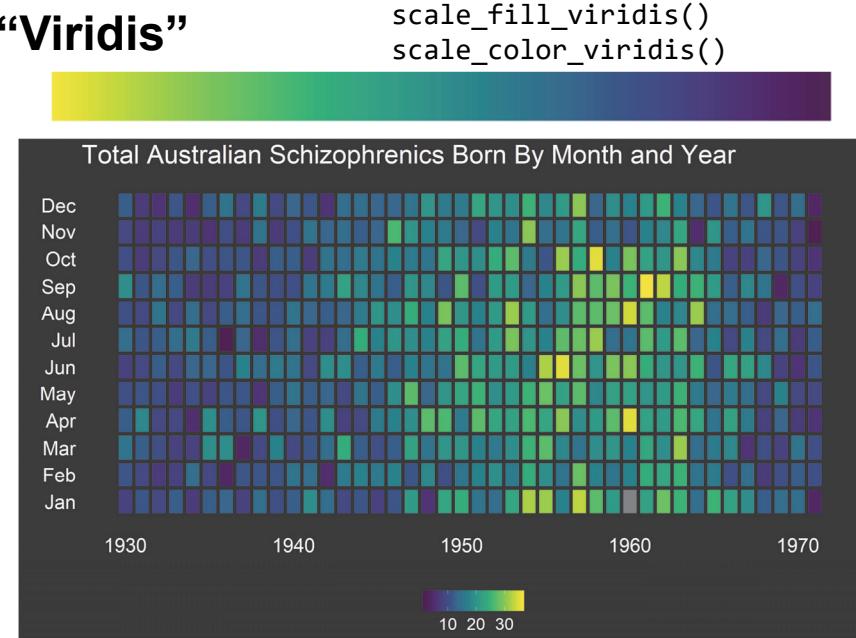
Colorful: Makes differences easily visible

Perceptually uniform: Close values have similar colors, consistent across range

Pretty (yes, it counts): Helps engage viewers



**“Viridis”**



Colorful, but lack of uniformity makes numbers/changes difficult to track

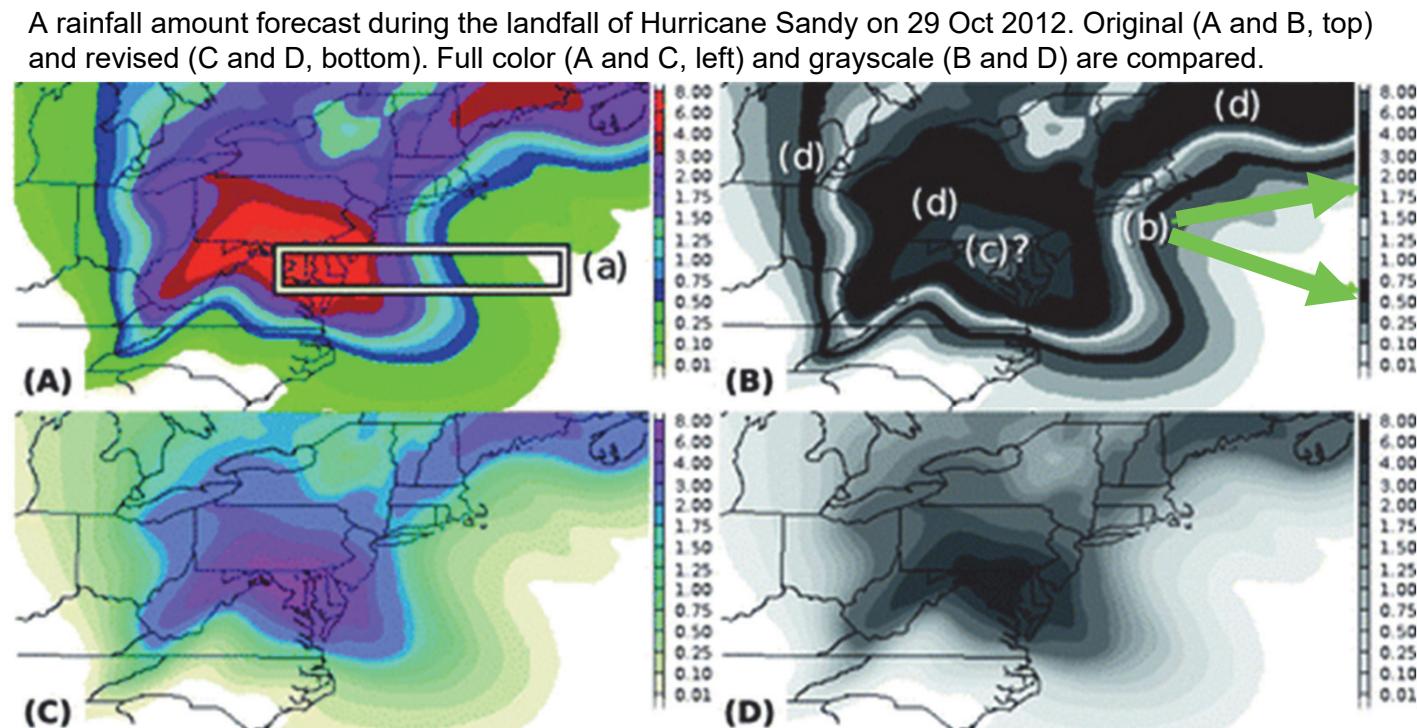
Adherence to principles makes it easy to visualize changes, understand trends

[Source: Viridis color palettes](#)

# Differences in lightness may not be preserved in grayscale

- Avoid “bands” of similar lightness in grayscale; scales should be continuous

Original  
from NOAA:  
*NOT grayscale-friendly*

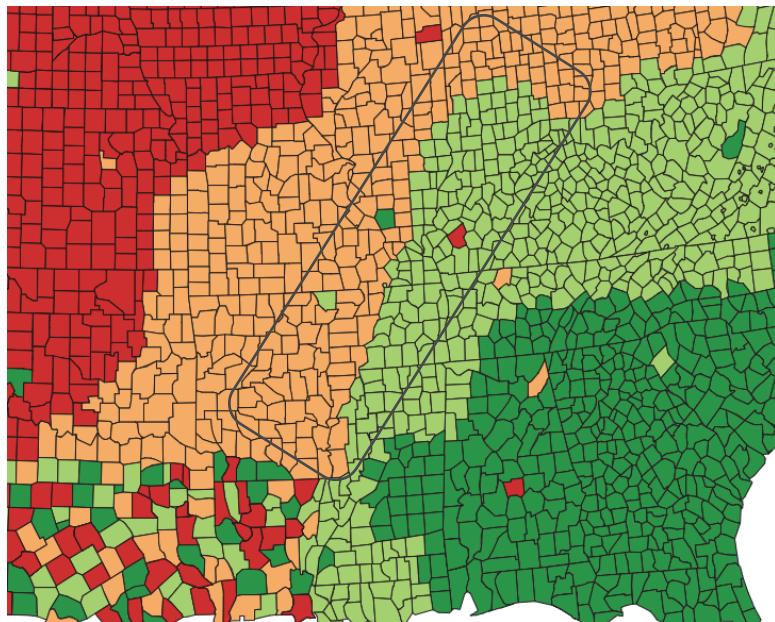


Modified:  
Grayscale-friendly

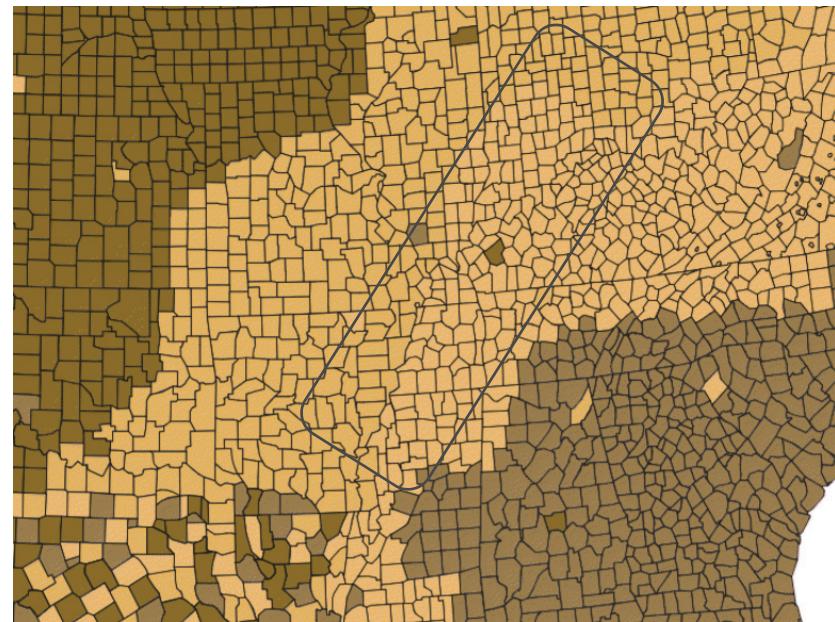
Source: NOAA, Rainfall forecast during Hurricane Sandy

# Red/green color blindness can make some palettes unusable

Normal



(deutanopia)  
Colorblind Simulation

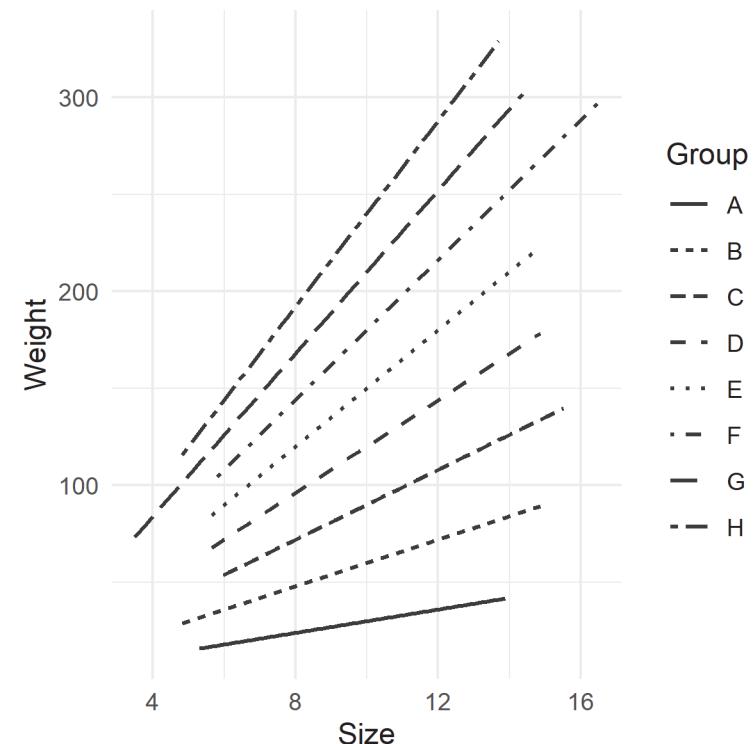


**Rule of thumb:** Avoid red and green unless supported by another encoding or annotation (e.g., added numbers in stoplight charts)

[Source: Color Brewer](#)  
[Source: Color Blind Simulator](#)

# It's hard to see more than two or three line types

- Using different line types is an effective way to distinguish between categories, especially when printing in grayscale
- ... but more than two or three types gets unwieldy



# Use “small multiples” for multiple comparisons AKA “faceting” for ggplot2 users

- More space to label clearly
- Allows comparison within and across groups
  - “Small multiples, reveal, all at once, a scope of alternatives, a range of options.” – Tufte

**Regional Support for Same-Sex Marriage**

% favoring same-sex marriage, 2003-2014

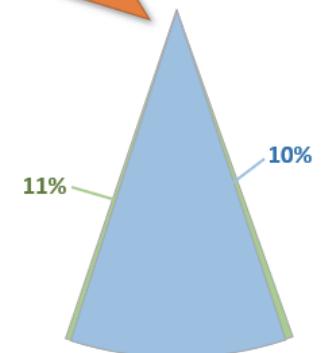
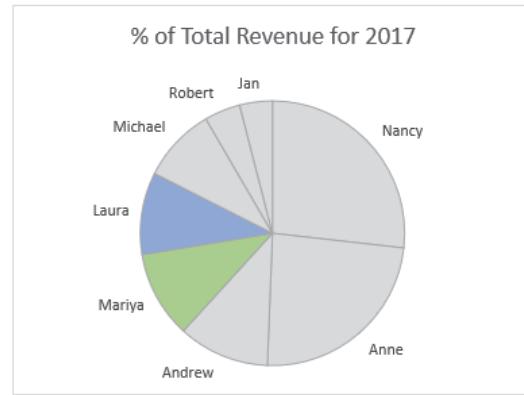


Use common style and axes across panels

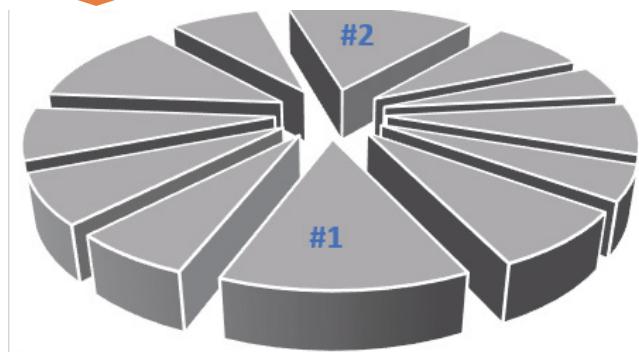
# Angles make comparisons difficult

- Angle differences are difficult to discern
  - Use bar or other chart
- Especially in 3D

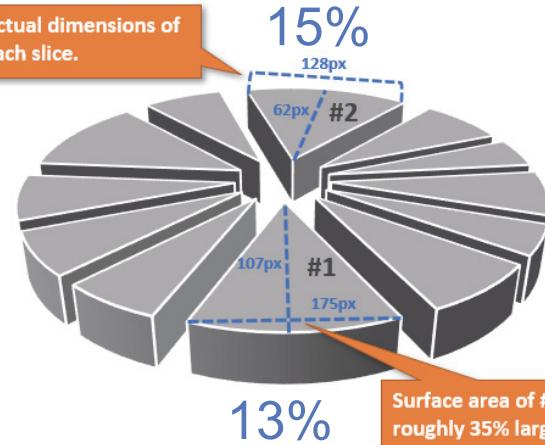
Compare slices by stacking them. The **size difference** is still very small even when enlarging the shapes.



Which slice represents the larger data point?



Actual dimensions of each slice.

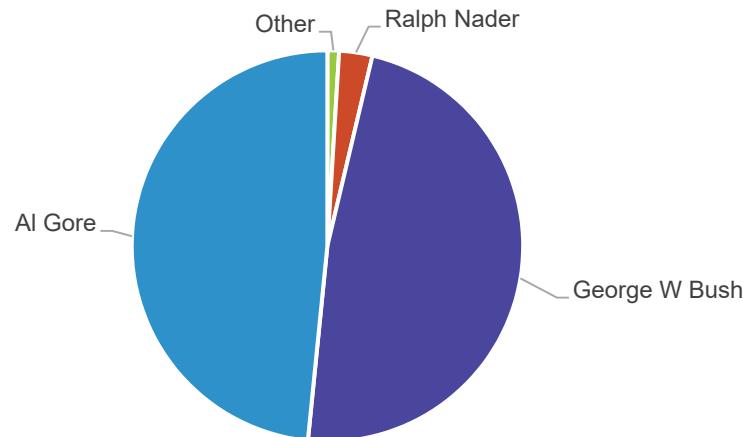


Source: Excel Campus

# Pie charts aren't all bad; it depends on the takeaway

- Two plots of the same data, but with a different takeaway
- Pie chart is appropriate on the left, but not on the right

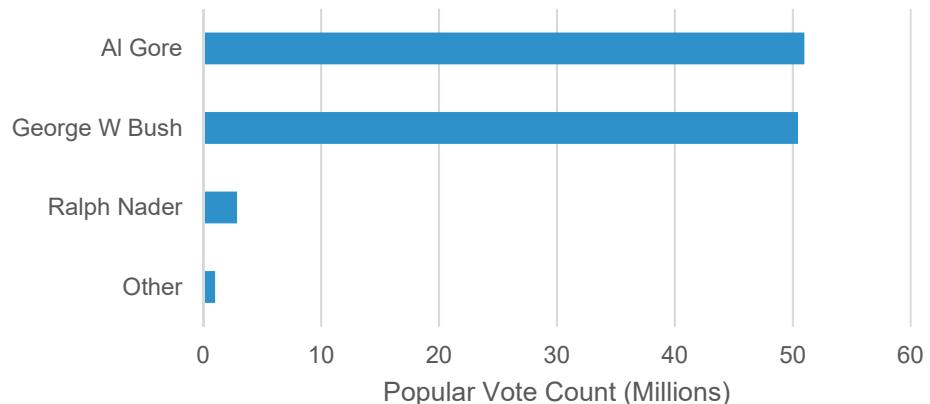
Two candidates accounted for the vast majority of the popular vote



Given this takeaway:

- Specific values **not** important
  - “Bush and Gore received a lot more votes than Nader”
- Ordering **not** important
- Focus on “part-to-whole” comparison

Gore edged Bush in popular vote, with Nader a distant third



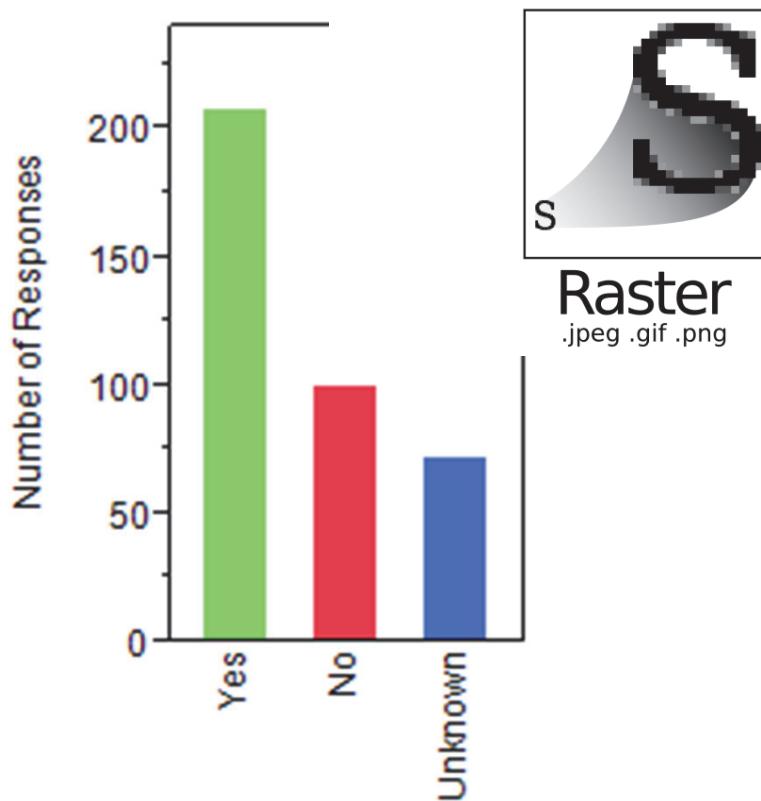
Given this takeaway:

- Specific values **are** important
  - “It was close, but Gore received ~52M votes, Bush ~51M”
- Ordering **is** important
  - Gore won the popular vote

# Vector graphics have higher quality overall

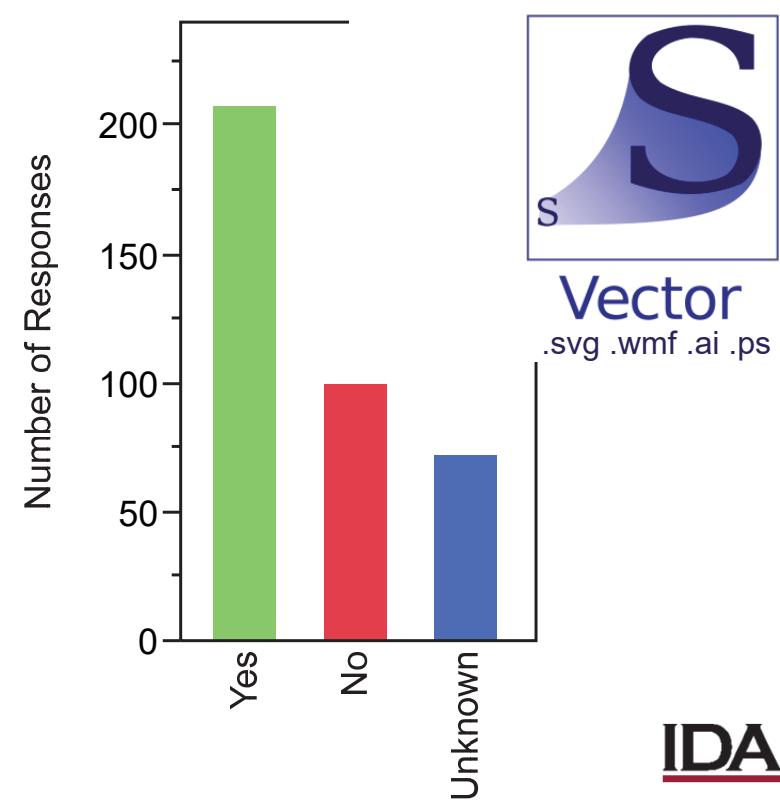
- Raster: “Pixels”

- Higher quality images have larger file sizes
- .png allows transparency



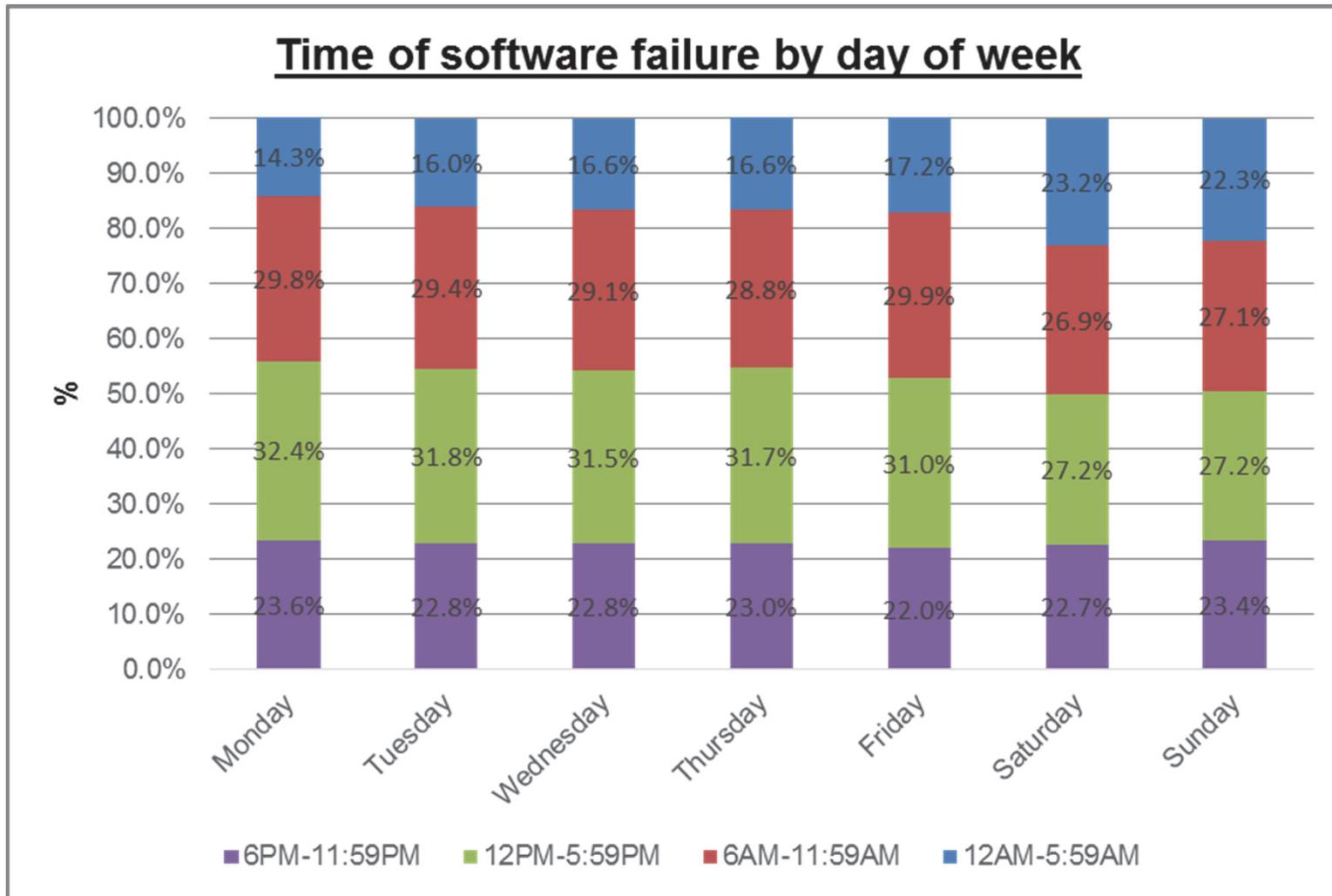
- Vector: “Strokes”

- Resizes easily



# LET'S MAKE SOME FIXES

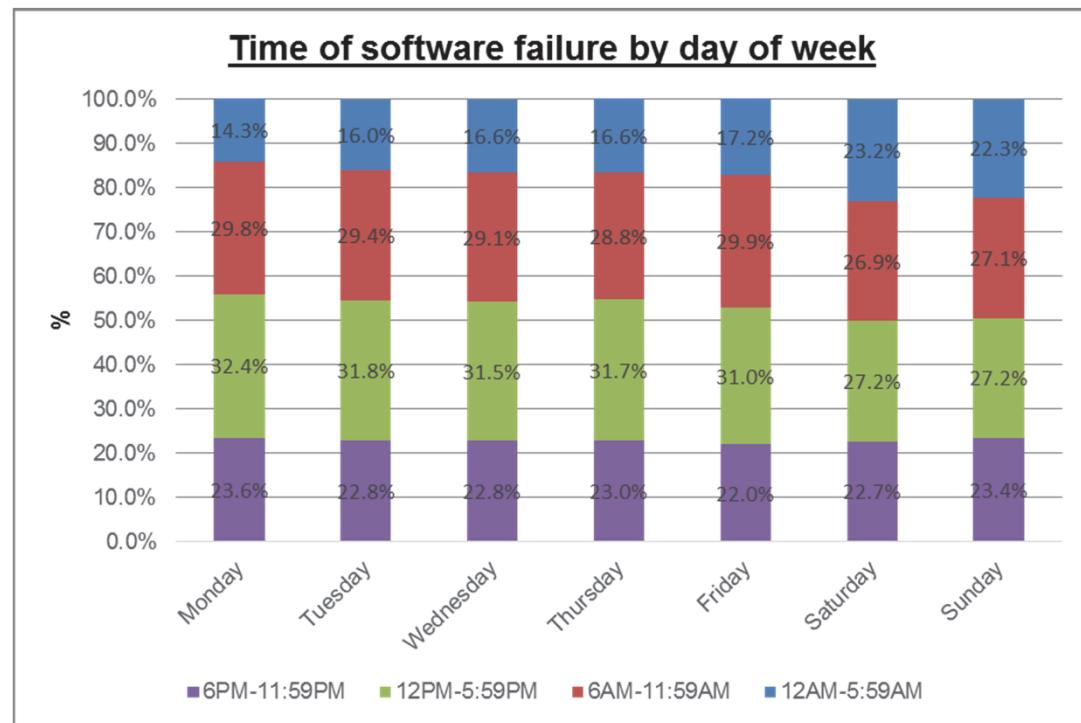
# Our redesign candidate – a stacked bar chart using Excel defaults



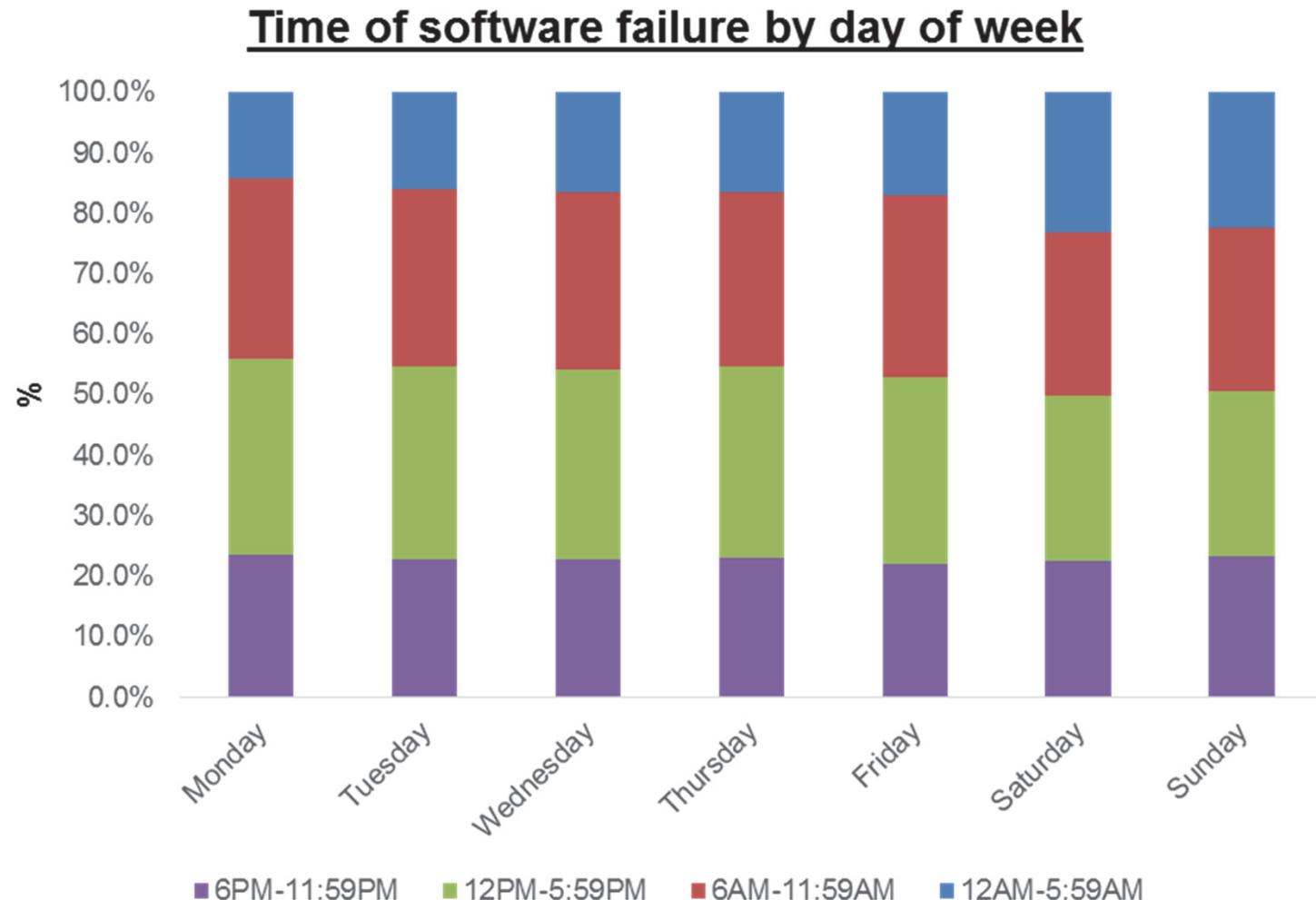
Source: Storytelling with Data

# The original figure has room for improvement

- Questions we can answer for the viewer
  - What is the context?
  - % of what?
  - Are the time bins meaningful?
- Elements we may consider changing
  - Non-data elements
    - o Axis labels
    - o Legend
    - o Title
  - Design
    - o Colors
    - o Legend placement
  - Guidance for viewer

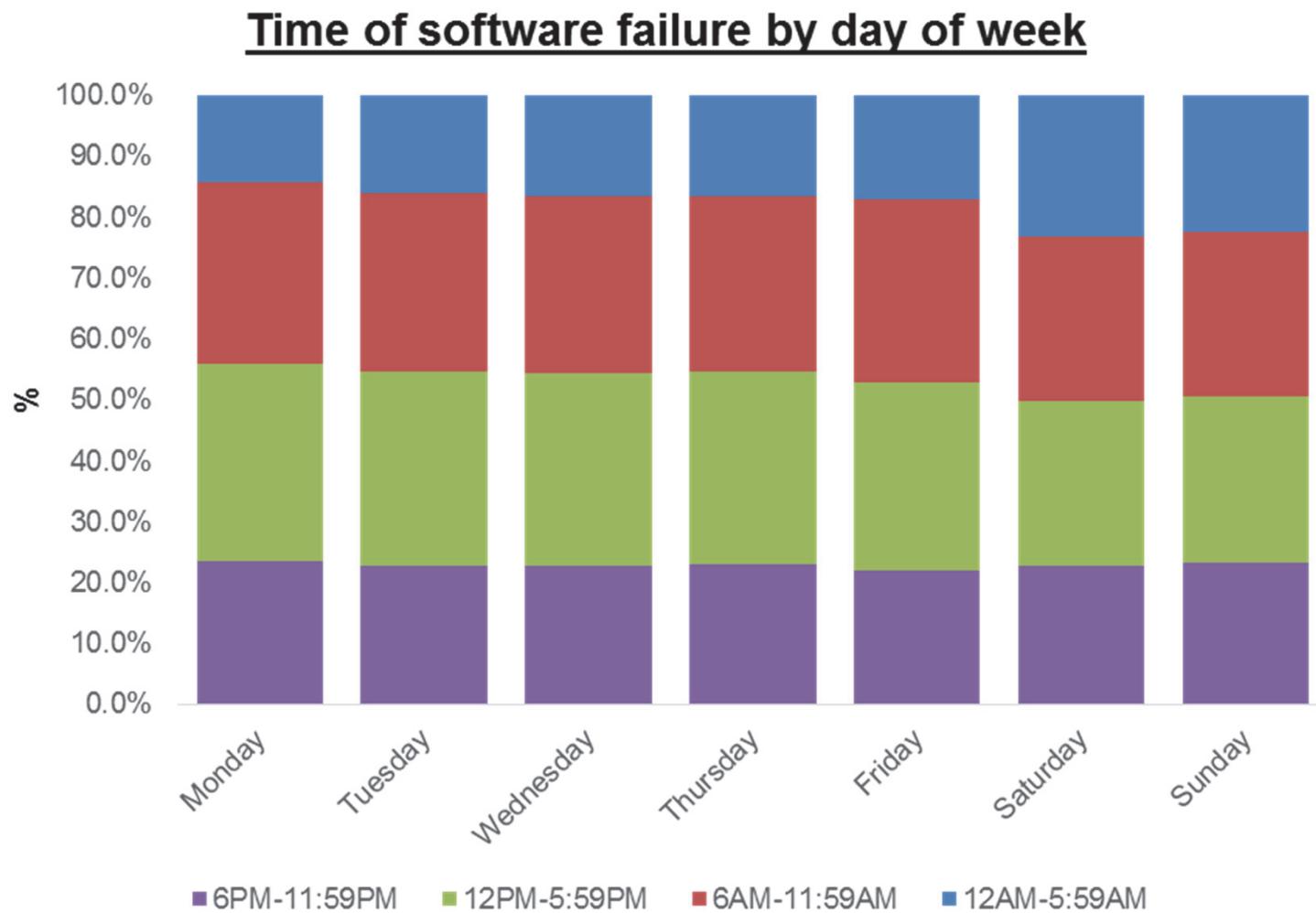


# Remove border, data labels



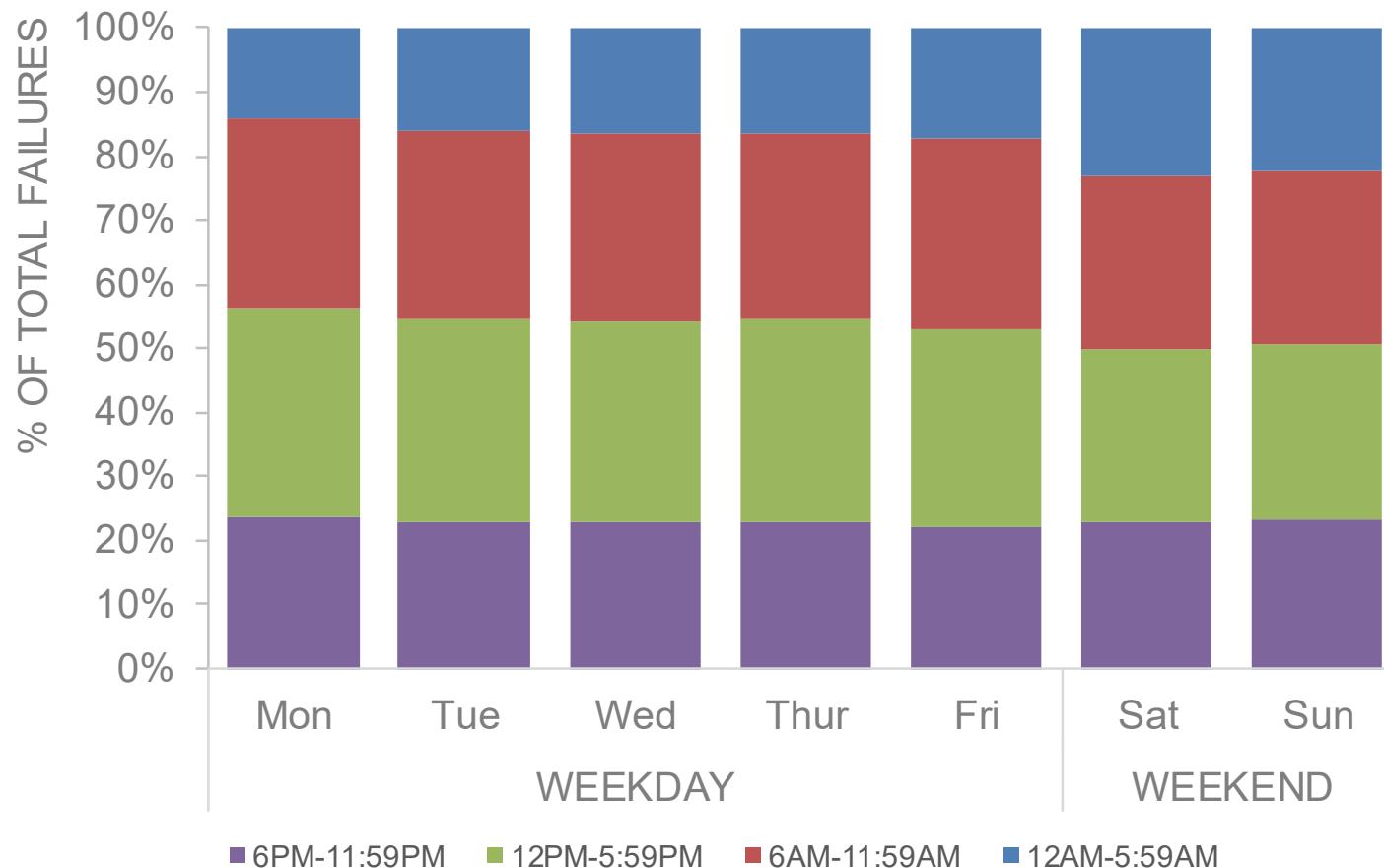
Removing visual clutter helps to focus the attention of the viewer

# Smaller gaps between bars enables more accurate comparison of segments



# Clarify axis labels, left-align title

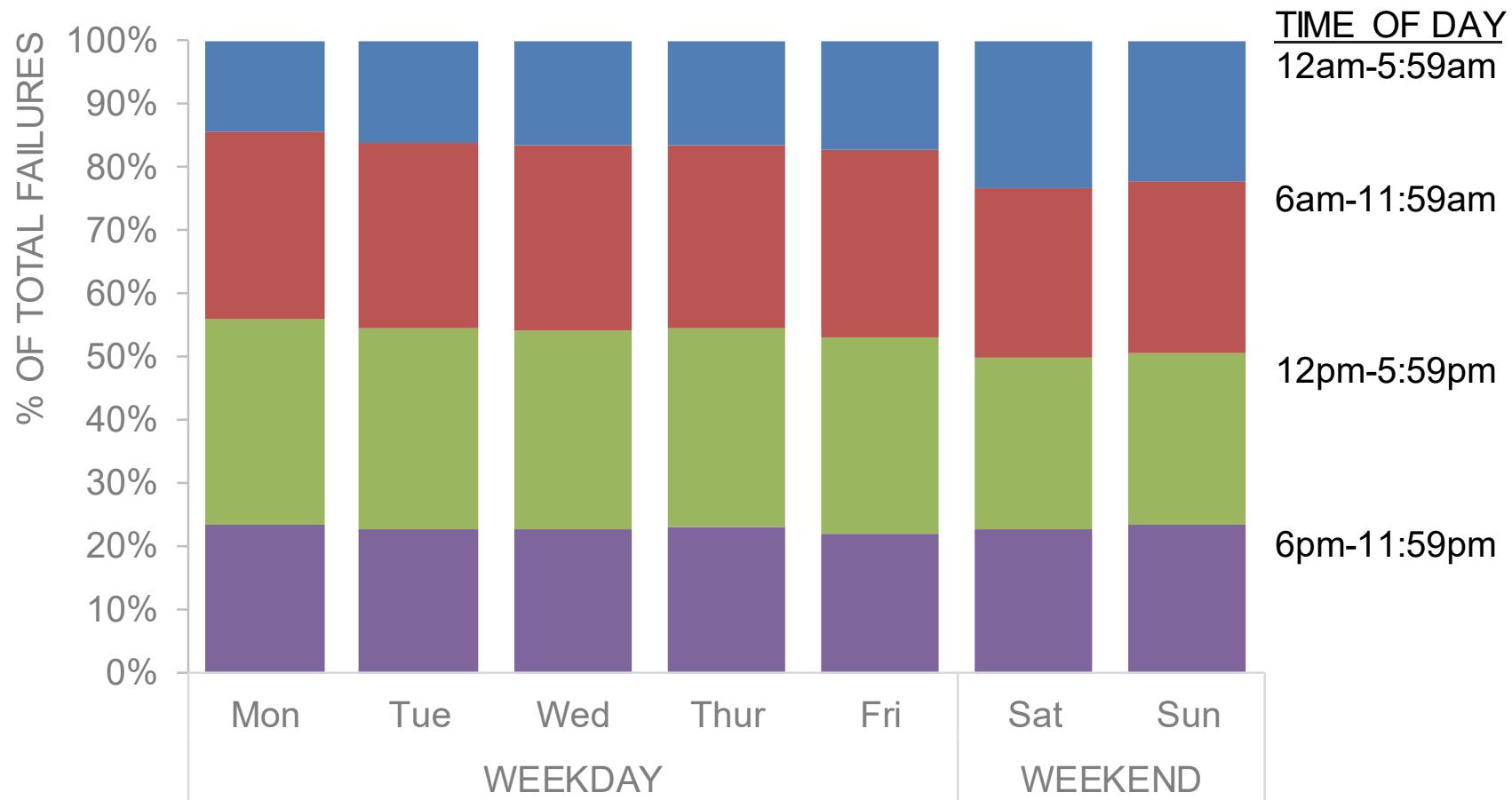
## Software failure time by day



Our eye naturally follows a “Z” pattern across and down the page, so left-aligning the title gives it more impact

# Remove legend, use direct label instead

## Software failure time by day

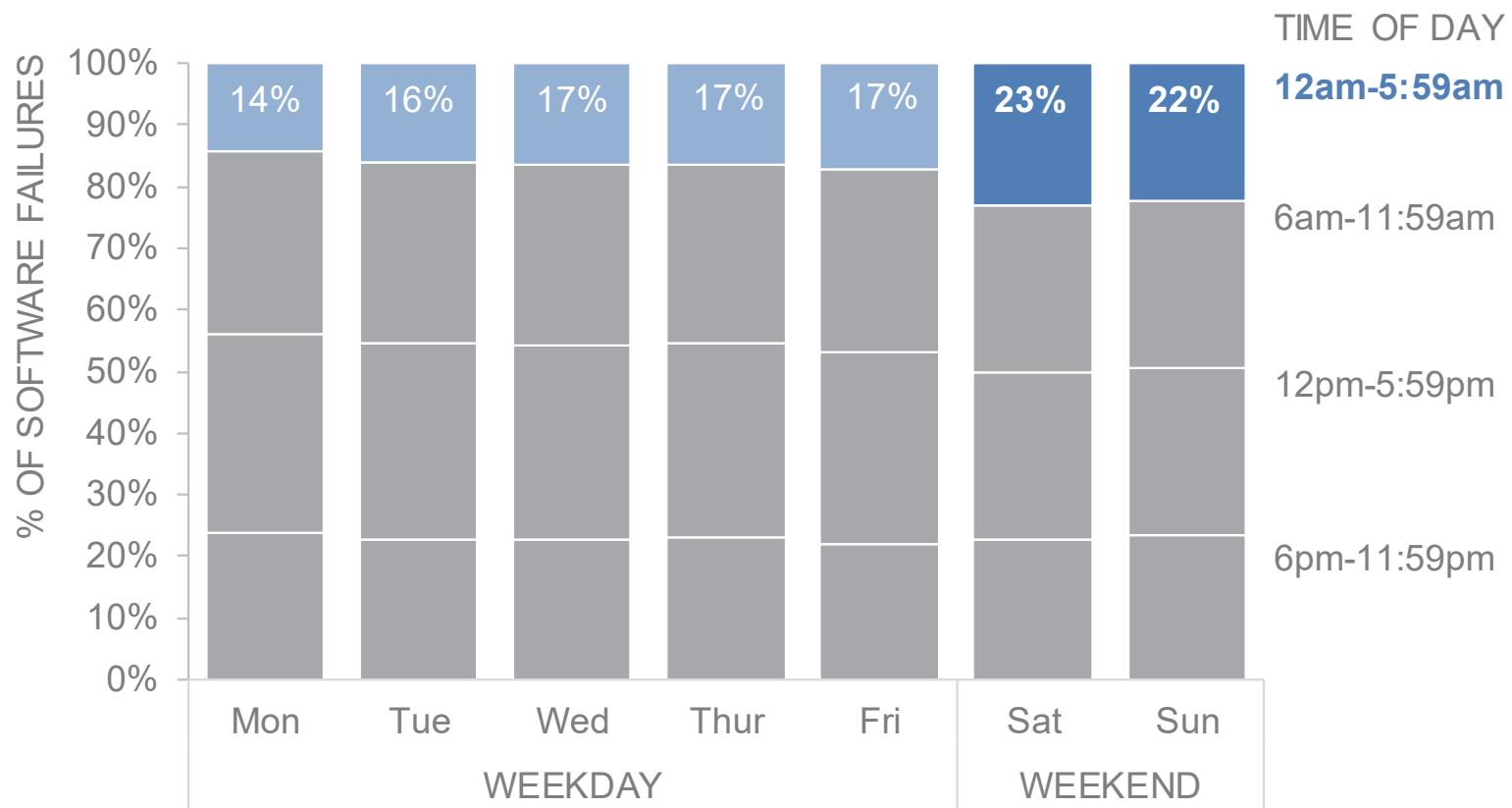


Directly labeling the bar segments rather than using a legend lowers the burden for the viewer

# Guide viewer with color, add a takeaway

When software failures occur

Early morning failures are more likely to occur on a weekend



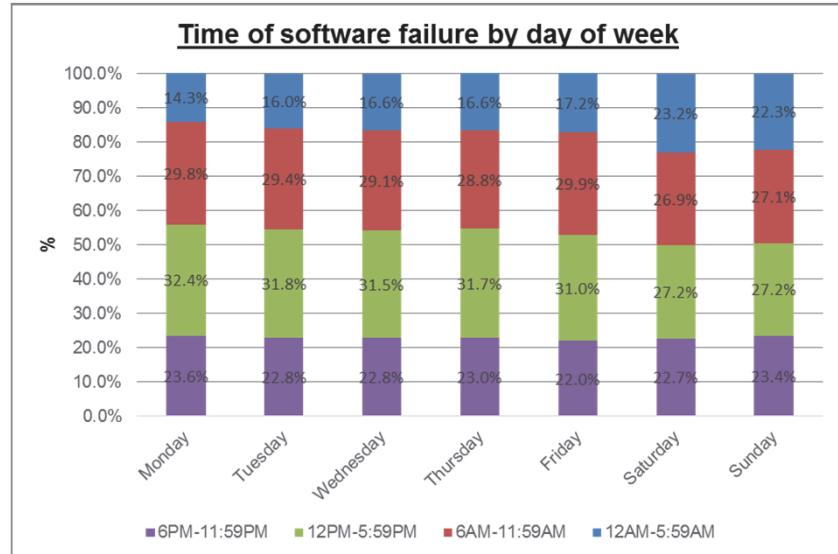
Data source: Fake System IOT&E (Fake Statistics Reports, Vol. 67, No. 1, January 31, 2018)

Coloring only the segments of interest (using the same color as the takeaway) tells the viewer exactly where to focus their attention

# Checking our work

## Before

- Visual clutter
- No “so what?”

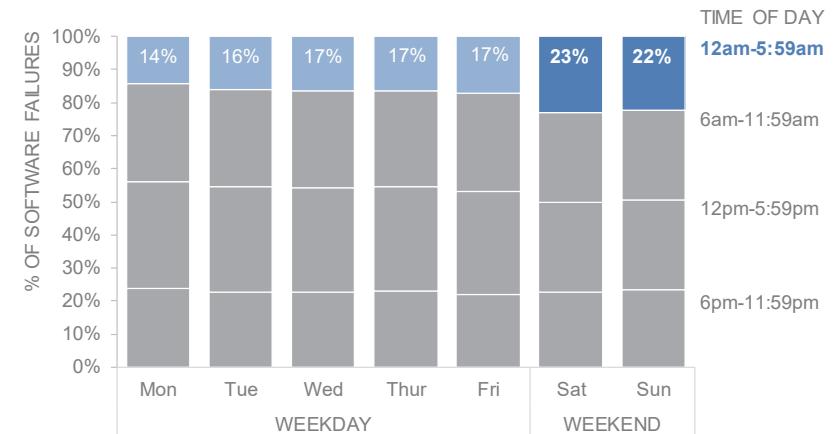


## After

- Reduced clutter
- A clear takeaway
- Guidance for the viewer

### When software failures occur

Early morning failures are more likely to occur on a weekend



Data source: Fake System IOT&E (Fake Statistics Reports, Vol. 67, No. 1, January 31, 2018)



# QUESTIONS?

Follow up with us at:  
Brian Vickers, [bvickers@ida.org](mailto: bvickers@ida.org)  
Andrew Flack, [aflack@ida.org](mailto: aflack@ida.org)  
Matt Avery, [mavery@ida.org](mailto: mavery@ida.org)



|  |  |                                     |  |  |
|--|--|-------------------------------------|--|--|
| <b>REPORT DOCUMENTATION PAGE</b>   |  |                                     | <i>Form Approved<br/>OMB No. 0704-0188</i>                                     |  |
| <p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b></p> |  |                                     |  |  |
| <b>1. REPORT DATE (DD-MM-YYYY)</b><br>09-2020  | <b>2. REPORT TYPE</b><br>IDA Publication |                                     | <b>3. DATES COVERED (From - To)</b>  |  |
| <b>4. TITLE AND SUBTITLE</b><br><br>Visualizing Data: I Don't Remember That Memo, but I Do Remember That Graph   |  |                                     | <b>5a. CONTRACT NUMBER</b><br>HQ0034-19-D-0001                                 | <b>5b. GRANT NUMBER</b><br>_____                               |
|  |  |                                     | <b>5c. PROGRAM ELEMENT NUMBER</b><br>_____                                     |  |
| <b>6. AUTHOR(S)</b><br><br>Matthew R. Avery (OED); Andrew C. Flack (OED); Brian D. Vickers (OED);  |  |                                     | <b>5d. PROJECT NUMBER</b><br>BD-09-2299  | <b>5e. TASK NUMBER</b><br>229990                               |
|  |  |                                     | <b>5f. WORK UNIT NUMBER</b><br>_____   |  |
| <b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b><br>Institute for Defense Analyses<br>4850 Mark Center Drive<br>Alexandria, Virginia 22311-1882   |  |                                     | <b>8. PERFORMING ORGANIZATION REPORT NUMBER</b><br>NS D-14363<br>H 2020-000362 |  |
| <b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b><br>Director, Operational Test and Evaluation<br>1700 Defense Pentagon<br>Washington, DC 20301   |  |                                     | <b>10. SPONSOR/MONITOR'S ACRONYM(S)</b><br>DOT&E                               |  |
|  |  |                                     | <b>11. SPONSOR/MONITOR'S REPORT NUMBER</b>                                     |  |
| <b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b><br>Approved for Public Release. Distribution Unlimited.   |  |                                     |  |  |
| <b>13. SUPPLEMENTARY NOTES</b><br>Project Leader: Heather Wojton   |  |                                     |  |  |
| <b>14. ABSTRACT</b><br>IDA analysts strive to communicate clearly and effectively. Good data visualizations can enhance reports by making the conclusions easier to understand and more memorable. The goal of this seminar is to help you avoid settling for factory defaults and instead present your conclusions through visually appealing and understandable charts. Topics covered include choosing the right level of detail, guidelines for different types of graphical elements (titles, legends, annotations, etc.), selecting the right variable encodings (color, plot symbol, etc.), advice on practical implementations, and determining whether to include a chart at all. Most of the time, there's no single "right" answer, so this presentation will include audience discussion to examine the trade-offs associated with different options.  |  |                                     |  |  |
| <b>15. SUBJECT TERMS</b><br>Data visualization; statistics   |  |                                     |  |  |
| <b>16. SECURITY CLASSIFICATION OF:</b>   |  | <b>17. LIMITATION OF ABSTRACT</b>   | <b>18. NUMBER OF PAGES</b>   | <b>19a. NAME OF RESPONSIBLE PERSON</b><br>Heather Wojton (OED) |
| <b>a. REPORT</b><br>Unclassified   | <b>b. ABSTRACT</b><br>Unclassified       | <b>c. THIS PAGE</b><br>Unclassified | Unlimited  | 65   |
| <b>19b. TELEPHONE NUMBER (include area code)</b><br>(703) 845-6811   |  |                                     |  |  |