Make Better Charts

Kristin Briney
Data Services Librarian

https://xkcd.com/1798/
It’s not enough to collect data.

You also need to effectively convey it.
Learning Outcomes

1. Participants will be able to choose the right chart for their message and type of data
2. Participants will be able to make good design decisions to highlighting what’s important and eliminating excess information
Learning Outcomes

1. Participants will be able to choose the right chart for their message and type of data

2. Participants will be able to make good design decisions to highlighting what’s important and eliminating excess information

Focusing on the very basics for a good foundation
Choose the Right Chart
What is your message?
Same Data
Different Charts

http://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/
Same Data
Different Charts

http://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/
Same Data
Different Charts
Same Data
Different Charts

http://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/
Same Data
Different Charts

http://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/
What is your message?
What is your message?

Hint: make it the title of your chart
Let’s Break it Down by Your Type of Data

A single number
Comparison
Beating a benchmark
Survey results
Parts of a whole
Correlations
Change over time
Qualitative data [not covered]

Adapted from Evergreen, S. D. H. (2017). Effective data visualization: The right chart for the right data.
Research Says The Best Charts Are…

## Choose Your Chart

<table>
<thead>
<tr>
<th></th>
<th>Big number</th>
<th>Icon array</th>
<th>Pie chart</th>
<th>Bar/column chart</th>
<th>Side-by-side column chart</th>
<th>Slope graph</th>
<th>Back-to-back bar chart</th>
<th>Dot plot</th>
<th>Small multiples</th>
<th>Column chart with benchmark</th>
<th>Combo chart</th>
<th>Stacked bar/column chart</th>
<th>Number and icon</th>
<th>Histogram</th>
<th>Map</th>
<th>Scatterplot</th>
<th>Diagram</th>
<th>Line chart</th>
<th>Deviating bar chart</th>
<th>Do not visualize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single number</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beating a benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey results</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts of a whole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change over time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-Standard Charts
Introducing…

1. Big number
2. Icon array
3. Slope graph
4. Back-to-back bar chart
5. Dot plot
6. Small multiples
7. Column chart with benchmark line
8. Combo chart
9. Histogram
Big Number

CASES OF MEASLES IN THE LAST YEAR

58 cases occurred in just two **Brooklyn, NY** neighborhoods

Population: 250,000

There were 200 total cases of measles in the **United States**

Population: 317,000,000

Slope Graph

Sales are higher for nearly every fruit in the West region, with the exception of Strawberries.

Back-to-Back Bar Chart

Texas vs California Unemployment

Distribution of monthly unemployment rates for Texas & California, December, 1982, through December, 2012

Unemployment Rates

Source: Bureau of Labor Statistics via FRB St. Louis

http://www.exceluser.com/training/charts/excel-chart-058.htm
Kindergarten readiness increased between Fall and Spring.
Minimum entry requirement: 65% in all areas.

Small Multiples

LIFE EXPECTANCY AT BIRTH, 2000-2015


http://flowingdata.com/2017/01/24/one-dataset-visualized-25-ways/
Bar Chart with Benchmark Line

In United States, religious minorities more likely to have college degrees than Christian majority

% with higher education, by religion

- Hindus: 96%
- Jews: 75
- Muslims: 54
- Buddhists: 53
- Unaffiliated: 44
- Christians: 36

Note: Adults ages 25 years and older as of 2010 (or latest year available).
Source: Pew Research Center analysis.
See Methodology for more details.
“Religion and Education Around the World”

PEW RESEARCH CENTER

Combo Chart

https://support.google.com/docs/answer/190718?hl=en
Histogram

Distribution of Listicle Sizes for BuzzFeed Listicles

http://minimaxir.com/2015/02/ggplot-tutorial/
Activity 1

<table>
<thead>
<tr>
<th>Question</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>I use Pinterest</td>
<td>28%</td>
<td>30%</td>
<td>31%</td>
</tr>
<tr>
<td>I use LinkedIn</td>
<td>28%</td>
<td>24%</td>
<td>29%</td>
</tr>
<tr>
<td>I use Facebook</td>
<td>71%</td>
<td>72%</td>
<td>79%</td>
</tr>
<tr>
<td>I use Twitter</td>
<td>23%</td>
<td>24%</td>
<td>24%</td>
</tr>
<tr>
<td>I use Instagram</td>
<td>26%</td>
<td>29%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Strategies for Chart Improvement

1. Start with your message
2. Highlight only the important data
3. Pick good colors
4. Reorder visually
5. Eliminate junk
6. Add strategic labels
Improving My Own Figure

Improving My Own Figure

Data policies cover more data issues than IP policies

- If Researcher Leaves the University: 7% (IP) vs. 64% (Data Policy)
- Identities Data Owner: 76% (IP) vs. 62% (Data Policy)
- Requires Retention - Specific: 3% (IP) vs. 62% (Data Policy)
- Defines data: 10% (IP) vs. 61% (Data Policy)
- Defines Who Has Access to Data: 7% (IP) vs. 52% (Data Policy)
- Appoints Data Steward: 10% (IP) vs. 46% (Data Policy)
- Requires Retention - Vague: 7% (IP) vs. 10% (Data Policy)

Improving My Own Figure

Data policies cover more data issues than IP policies

<table>
<thead>
<tr>
<th>Condition</th>
<th>Data Under IP Policy</th>
<th>Standalone Data Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Researcher Leaves the University</td>
<td>7%</td>
<td>64%</td>
</tr>
<tr>
<td>Identifies Data Owner</td>
<td>76%</td>
<td>62%</td>
</tr>
<tr>
<td>Requires Retention - Specific</td>
<td>3%</td>
<td>62%</td>
</tr>
<tr>
<td>Defines data</td>
<td>10%</td>
<td>61%</td>
</tr>
<tr>
<td>Defines Who Has Access to Data</td>
<td>7%</td>
<td>52%</td>
</tr>
<tr>
<td>Appoints Data Steward</td>
<td>10%</td>
<td>46%</td>
</tr>
<tr>
<td>Requires Retention - Vague</td>
<td>7%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Chart Title

Start With Your Message

• Write your main message as your title
  • A subtitle can add further clarification

Data policies cover more data issues than IP policies

- Defines data
- Identifies Data Owner
- Appoints Data Steward
- Requires Retention - Specific
- Requires Retention - Vague
- Defines Who Has Access to Data
- If Researcher Leaves the University

Buffer 1 | Data Under IP Policy | Standalone Data Policy | Buffer 2
Highlight Only the Important Data

• Pick 1-3 things to highlight or pick a different chart
• Remember your message
Pick Good Colors

• Use color sparingly/to highlight
  • Grey is your friend
• Related things have related colors
  • Be consistent
• Avoid
  • Pairing red and green (colorblind)
  • Pink for women and blue for men
• Check: print in black and white!!!
Reorder Visually

- Others don’t care about original order
- Sort order visually
  - Caveats for time data, etc.
Eliminate Junk

- Delete legends
- Remove axis lines
- Remove or trim axis numbers
- Condense data categories
  - Eg. Positive and negative responses instead of 5-value Likert scale

Data policies cover more data issues than IP policies

- If Researcher Leaves the University
- Identifies Data Owner
- Requires Retention - Specific
- Defines data
- Defines Who Has Access to Data
- Appoints Data Steward
- Requires Retention - Vague
Add Strategic Labels

• Legend
  • Put series name right next to the data on the plot

• Axis numbers
  • Label data directly

Data policies cover more data issues than IP policies

<table>
<thead>
<tr>
<th>Issue</th>
<th>Data Under IP Policy</th>
<th>Standalone Data Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>If Researcher Leaves the University</td>
<td>7%</td>
<td>64%</td>
</tr>
<tr>
<td>Identifies Data Owner</td>
<td>76%</td>
<td>62%</td>
</tr>
<tr>
<td>Requires Retention - Specific Defines data</td>
<td>3%</td>
<td>62%</td>
</tr>
<tr>
<td>Defines Who Has Access to Data</td>
<td>10%</td>
<td>61%</td>
</tr>
<tr>
<td>Appoints Data Steward</td>
<td>7%</td>
<td>52%</td>
</tr>
<tr>
<td>Requires Retention - Vague</td>
<td>10%</td>
<td>46%</td>
</tr>
</tbody>
</table>
Improving My Own Figure

Data policies cover more data issues than IP policies:

- If Researcher Leaves the University: 76% under IP policy, 64% under standalone data policy.
- Identifies Data Owner: 7% under IP policy, 62% under standalone data policy.
- Requires Retention - Specific: 3% under IP policy, 62% under standalone data policy.
- Defines data: 10% under IP policy, 61% under standalone data policy.
- Defines Who Has Access to Data: 7% under IP policy, 52% under standalone data policy.
- Appoints Data Steward: 10% under IP policy, 46% under standalone data policy.
- Requires Retention - Vague: 7% under IP policy, 10% under standalone data policy.
Strategies for Chart Improvement

1. Start with your message
2. Highlight only the important data
3. Pick good colors
4. Reorder visually
5. Eliminate junk
6. Add strategic labels
Activity 2

https://support.google.com/docs/answer/190718?hl=en
Come Back for Data Viz 102!

• Friday, March 9 at 12PM in Library DH Lab
  • Discuss and improve your own visualization by working through the framework established in Data Visualization 101. Please bring a dataset with you to this workshop.
Attribution

• Content is available under a CC BY attribution license
• Please attribute to Kristin Briney, UW-Milwaukee

• Questions? Let me know at briney@uwm.edu