DATA 227

Captions and Tables

2022-12-01

Telling a Story and Making a Point 1

- "Most data visualization is done for the purpose of communication. We have an insight about a dataset, and we have a potential audience, and we would like to convey our insight to our audience. To communicate our insight successfully, we will have to present the audience with a clear and exciting story."
- "The need for a story may seem disturbing to scientists and engineers, who may equate it with making things up, putting a spin on things, or overselling results. However, this perspective misses the important role that stories play in reasoning and memory."

Wilke, Claus O. Fundamentals of data visualization: a primer on making informative and compelling figures. O'Reilly Media, 2019.

Telling a Story and Making a Point 2

- "We get excited when we hear a good story, and we get bored when the story is bad or when there is none. Moreover, any communication creates a story in the audience's minds."
- "If we don't provide a clear story ourselves, then our audience will make one up. In the best-case scenario, the story they make up is reasonably close to our own view of the material presented. However, it can be and often is much worse. The made-up story could be 'this is boring,' 'the author is wrong,' or 'the author is incompetent.' "

What is a story?

- "A story is a set of observations, facts, or events, true or invented, that are presented in a specific order such that they create an emotional reaction in the audience. The emotional reaction is created through the build-up of tension at the beginning of the story followed by some type of resolution towards the end of the story."
  - Opening–Challenge–Action–Resolution
  - Lead–Development–Resolution
  - Lead–Development (where the lead gives away the main point up front and the subsequent material provides further details)
  - Action–Background–Development–Climax–Ending

Data Visualization as a Story

- "Instead, I want to discuss how we can bring data visualizations into the story arc. Most importantly, we need to realize that a single (static) visualization will rarely tell an entire story. A visualization may illustrate the opening, the challenge, the action, or the resolution, but it is unlikely to convey all these parts of the story at once. To tell a complete story, we will usually need multiple visualizations."

Understanding Meaning

- "It is entirely possible to follow all the recommendations I have provided throughout this book and still prepare figures that confuse. When this happens, you may have fallen victim to two common misconceptions."
  - "First, that the audience can see your figures and immediately infer the points you are trying to make;"
  - "Second, that the audience can rapidly process complex visualizations and understand the key trends and relationships that are shown."
- "Neither of these assumptions is true. **We need to do everything we can to help our readers understand the meaning of our visualizations and see the same patterns in the data that we see.**"

A Caption's Purpose

"A data visualization is not a piece of art meant to be looked at only for its aesthetically pleasing features. Instead, its purpose is to convey information and make a point. To **reliably achieve this goal** when preparing visualizations, we have to place the data into context and provide accompanying titles, captions, and other annotations."

Titles

- One critical component of every figure is the figure title. **Every figure needs the title.**
- The job of the title is to accurately convey to the reader what the figure is about, what point it makes.
- The first part of the caption is always the title, not a description of the contents of the figure.
- A title does not have to be a complete sentence, though short sentences making a clear assertion can serve as titles.

Captions vs. Paragraphs 1

- "Captions should be concise but comprehensive. They should describe the data shown, draw attention to important features contained within the figure, and *may sometimes also include interpretations of the data*. Figures are typically read from the bottom up, so captions go below the figure and are left-justified."

The Writing Center, University of North Carolina at Chapel Hill, "Figures and Charts," https://writingcenter.unc.edu/tips-and-tools/figures-and-charts/.

Captions vs. Paragraphs 2

- This may be a matter of personal preference, but to me:
  - A caption should contain facts–describing the data, drawing attention to the data features, etc.
  - The paragraph might include facts, but also should include interpretations of the data.

Captions vs. Paragraphs: Example

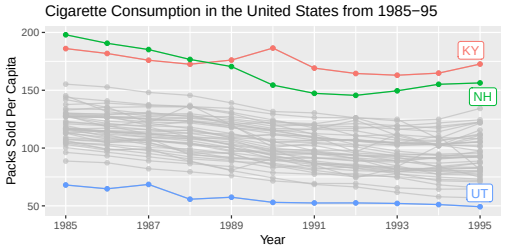


Figure 1: Cigarette Consumption in the United States from 1985-1995. This scatterplot displays the trends in packs of cigarettes sold per capita across the 48 contiguous states in the US for 1985-1995. Most of the states show a linear, decreasing relationship. The states with the highest packs per capita sales are Kentucky and New Hampshire, whereas the state with the lowest packs per capita sales is Utah.

As shown in Figure 1, cigarette sales per capita trended downwards in the 48 contiguous states in the US from 1985-1995. This downward trend is potentially due to public health campaigns advertising the negative health effects of cigarette smoking. The states with the highest packs per capita sales are Kentucky (which contains about half of the US's tobacco farmland) and New Hampshire (which has no sales tax, relative to a very high sales tax in all surrounding states). The state with the lowest per capita sales is Utah (which has a very high number of Mormon residents who might abstain from tobacco as per the Mormon Church's official handbook).

Charts and Accessibility

1. If the data in a chart, graph or map is crucial to the content of a Web page, then you must provide a text description of the image. In some cases, a numeric table replicating the chart data could provide additional accessibility.
2. Supplement color-coding of charts with texture, differences in line style, text in graphs or different shades of color to improve accessibility for colorblind users. Charts should be readable in black and white.
3. Don't convert tables of data into images—use an actual data table instead.

Penn State Accessibility, "Charts and Accessibility," https://accessibility.psu.edu/images/charts/.

Alt Text for Data Visualization

"You probably can't write text that conveys the entire meaning of a chart. But, that doesn't mean that you shouldn't try."

Amy Cesal, Nightingale Society, "Writing Alt Text for Data Visualization," Medium, July 23, 2020, https://medium.com/nightingale/writing-alt-text-for-data-visualization-2a218ef43f81.

- Alt text (also called alt tags or alternative text) is a written description of an image that accompanies the image to help convey meaning.
- Good alt text helps more people (visually impaired, people using screen readers) to understand the content of the image.

How to Write Alt Text for Data Visualization 1

- Adding alt text is better than not doing anything at all if there's meaningful information in the image.
- Add a long description if you can add HTML, to more fully convey the meaning.
- Supplement with a link to the raw data, so curious readers can access the data in their own preferred program.
- Keep your alt text short. Alt text is read linearly by screen readers, which means that people can't go back a word if they missed something. For SEO purposes, Google cuts off after a certain amount of characters.

How to Write Alt Text for Data Visualization 2

alt= “**Chart type** of **type of data** where **reason for including chart**”

Include a **link to data source** somewhere in the text

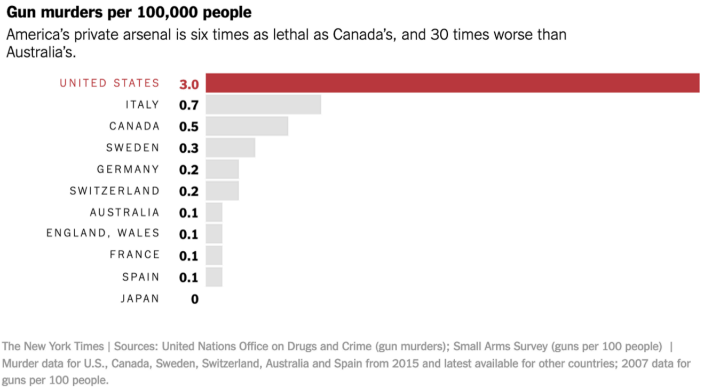
How to Write Alt Text for Data Visualization 3

- **Chart type**–It's helpful for people with partial sight to know what chart type it is and gives context for understanding the rest of the visual.
  - Example: Line graph
- **Type of data**–What data is included in the chart? The x and y axis labels may help you figure this out.
- Example: number of bananas sold per day in the last year

How to Write Alt Text for Data Visualization 4

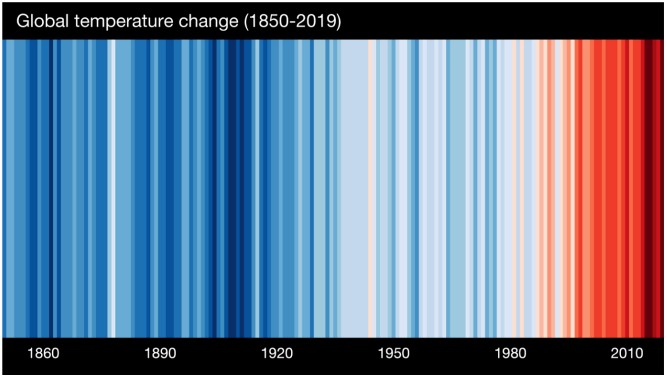
- **Reason for including the chart**–Think about why you're including this visual. What does it show that's meaningful. There should be a point to every visual and you should tell people what to look for.
  - Example: the winter months have more banana sales
- **Link to data or source**
  - *Don't include this in your alt text*, but it should be included somewhere in the surrounding text. People should be able to click on a link to view the source data or dig further into the visual. This provides transparency about your source and lets people explore the data.
  - Example: Data from the USDA

Alt Text Examples 1



- alt = "Bar Chart of gun murders per 100,000 people where America's murder rate is 6 times worse than Canada's, and 30 times than Australia."

Alt Text Examples 2



- alt = "Colored stripes of chronologically ordered temperatures where they increase in red to show the warming global temperature."

Adding Alt Text

- In HTML code: Add alt text alt=“ ” to any image that adds meaning inside the image tag . Add a long description longdesc=“ ” when possible.
- On Twitter: When you add an image, use the add description button or +alt. Twitter has full instructions.
- On Instagram: Instagram hides the alt text functionality, but there is documentation on how to add alt text. Choose a filter and edit the image, then tap Next. Tap Advanced Settings at the bottom of the screen. Tap Write Alt Text.
- In Microsoft Office Products: Microsoft Office has a variety of ways that you can add alt text depending on which program and what version you are using. However, they have extensive documentation on how to add alt text.

Tables vs. Data Visualizations

- Tables are a unique form of visualizing data because, unlike many charts, they are not usually intended to give a quick, visual representation of data.
- Instead, tables are useful when you want to show the exact values of your data or estimates.
- *They are not the best solution if you want to show a lot of data or if you want to show the data in a compact space*, but a well-designed table can help your reader find specific numbers and discover patterns and outliers.

Schwabish, Jonathan A. "Ten guidelines for better tables." Journal of Benefit-Cost Analysis 11, no. 2 (2020): 151-178.

Guidelines for Tables (Fundamentals of Data Visualization)

Some key rules for table layout are the following:

1. Do not use vertical lines.
2. Do not use horizontal lines between data rows. (Horizontal lines as separator between the title row and the first data row or as frame for the entire table are fine.)
3. Text columns should be left aligned.
4. Number columns should be right aligned and should use the same number of decimal digits throughout.
5. Columns containing single characters are centered.
6. The header fields are aligned with their data, i.e., the heading for a text column will be left aligned and the heading for a number column will be right aligned.

Vertical and Horizontal Lines

1. Do not use vertical lines.
2. Do not use horizontal lines between data rows.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan*	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

Creating Separation

- Horizontal lines as separator between the title row and the first data row or as frame for the entire table are fine.
- If we feel that visually separating table rows is necessary, then alternating lighter and darker shading of rows tends to work well without creating much clutter.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan*	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

Text Columns

3 Text columns should be left aligned.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan*	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

Number Columns

4. Number columns should be right aligned and should use the same number of decimal digits throughout.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan*	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

More Guidelines for Tables (Ten Guidelines for Better Tables)

- Rule 1: Offset the headers from the body.
- Rule 5: Select the appropriate level of precision.
- Rule 7: Remove Unit Repetition
- Rule 8: Highlight Outliers

Offset the Headers from the Body

- Make your column titles clear. Try using boldface type or lines to offset them from the numbers and text in the body of the table. It should be clear that the headers are not data values but categories or labels.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan*	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

Select the Appropriate Level of Precision

- Precision to the fifth decimal place is rarely necessary. Consider a balance between necessary precision and a clean, spare table.
- But don't use too few digits, either. Reporting per capita GDP growth as whole numbers masks important variation across countries.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	350000	1950
Lord of the Rings	J.R.R. Tolkien	3	480000	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480000	1968
Wheel of Time	Robert Jordan*	14	4900000	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1080000	1997

Remove Unit Repetition

- Repeating the symbol throughout the table is overkill and adds clutter.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7 books	345535 pages	1950
Lord of the Rings	J.R.R. Tolkien	3 books	481103 pages	1954
The Earthsea Cycle	Ursula K. Le Guin	6 books	480503 pages	1968
Wheel of Time	Robert Jordan*	14 books	4900036 pages	1990
A Song of Ice and Fire	George R.R. Martin	5 books	1770000 pages	1991
Harry Potter	J.K. Rowling	7 books	1084170 pages	1997

Highlight Outliers

- We might want to highlight outlier values by making the text boldface, shading it with color, or even shading entire cells–but be careful to follow accessibility guidelines!

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

Final Table

Table 1: A table containing various characteristics of several popular fantasy series, including their titles, authors, the number of books in the series, the number of words in the series, and the year the first book in the series was published.

Series	Author	Books	Words	Year
Chronicles of Narnia	C.S. Lewis	7	345535	1950
Lord of the Rings	J.R.R. Tolkien	3	481103	1954
The Earthsea Cycle	Ursula K. Le Guin	6	480503	1968
Wheel of Time	Robert Jordan [1]	14	4900036	1990
A Song of Ice and Fire	George R.R. Martin	5	1770000	1991
Harry Potter	J.K. Rowling	7	1084170	1997

[1] Robert Jordan died before the final installment of his series was completed, so author Brandon Sanderson was brought in to help finish it (it was was subsequently published in three volumes).

Sources

- Amy Cesal, Nightingale Society, "Writing Alt Text for Data Visualization," Medium, July 23, 2020, https://medium.com/nightingale/writing-alt-text-for-data-visualization-2a218ef43f81.
- Penn State Accessibility, "Charts and Accessibility," https://accessibility.psu.edu/images/charts/.
- Schwabish, Jonathan A. "Ten guidelines for better tables." Journal of Benefit-Cost Analysis 11, no. 2 (2020): 151-178. The Writing Center, University of North Carolina at Chapel Hill, "Figures and Charts," https://writingcenter.unc.edu/tips-and-tools/figures-and-charts/.
- Wilke, Claus O. Fundamentals of data visualization: a primer on making informative and compelling figures. O'Reilly Media, 2019.