

Communicating Data Visualizations

Data visualization needs to effectively communicate the results of the data analysis, and demonstrate how you reached those results. The term for this is storytelling. Please watch the videos below for more information.

- The Beauty of Data visualization
https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization
- Are viewers expecting a Story in Your Data Visualization?
https://www.youtube.com/watch?v=XJqBcXZY9HY&ab_channel=storytellingwithdata
- How to declutter data visualization:
https://www.youtube.com/watch?v=X79o46W5pll&ab_channel=storytellingwithdata

Vocabulary

- visualization - transforming data into a graphical representation
- storytelling - using visualization to identify patterns and trends in data

What is an effective visualization?

According to David McCandless, a British data journalist, writer, and information designer, good data visualizations provide four qualities: function, form, integrity, and interest.

- Functional: the visualization shows the patterns and trends it was created for
- Form: visualization grabs and holds the attention of the target audience
- Integrity: How trustworthy and unchanged the data is
- Interest: interpreting the visualization provides new meaning and directions for target audience

Data visualization is often referred to as storytelling. Data Analysts must tell the story represented by the analysis and results. Take the time to identify the datasets that have relevance to the story. Pull out interesting patterns and trends, then identify outliers or patterns that do not fit the expectations of the story. When presenting, explain why the data or outliers are interesting, or unique or what correlations were found that was unexpected, and why that might be.

Visualization Design

When choosing a visualization design, there are 3 things to consider: Objective, audience, data. There was a question to be answered or a problem to be solved, that required data analysis. The visualization must demonstrate the analysis and the resolution. An audience will be viewing this resolution. What qualities does this visualization have to capture and hold the audience interested, or

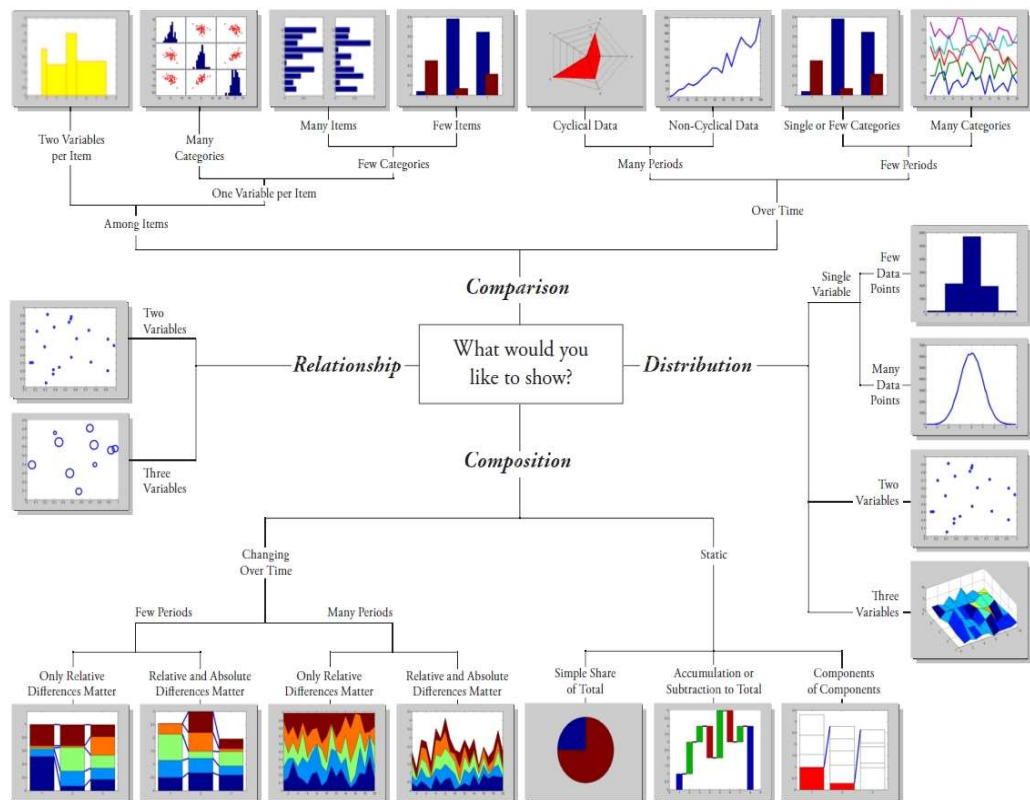
to influence the audience? The goal is to convey the results of the data analysis in a manner that highlights the important answers, informs the audience of patterns or trends, or influences them to make a decision. Consider providing subtexts and titles on the other information that came out of the analysis the audience should be aware of.

Use the right kind of model or chart. The chart is correct if it indicates relationships between values without the audience needing to take the time to decipher. It accurately represents quantities and/or comparison of quantities. Order values ascending or descending as appropriate, and point out how the audience should use the information. (Stephen Few, 2013, *Data Visualization for Human Perception*).

In general, the purpose of the story determines the type of chart

- Comparison: Column, bar, table, line, circular area chart
- Relationship: scatter or bubble chart
- Distribution: histogram, scatter, or 3D area
- Composition: stacked column or bar, stacked area, pie, waterfall

Chart Suggestions—A Thought-Starter



Modified with permission -Doug Hull
blogs.mathworks.com/videos
hull@mathworks.com 2009
www.ExtremePresentation.com
© 2009 A. Abela — a.v.abela@gmail.com

Design Tips on Charts and graphs

1. Always scale to zero, or the lowest available value in bar-based charts. Zero scale is optional in non-bar charts. The top of the chart is the highest value, not the highest available value. In bar charts, not having a 0-scale axis can exaggerate differences:

Chart shows an exaggerated difference

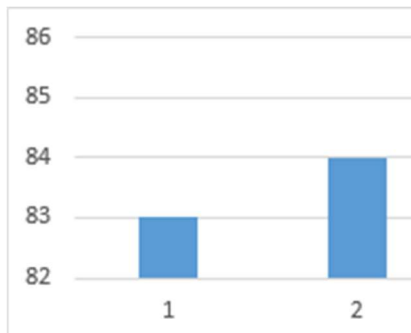
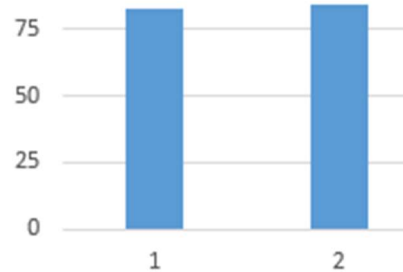
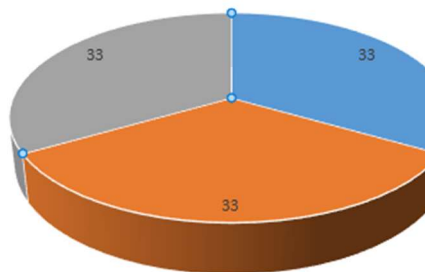


Chart accurately reflects the slight difference

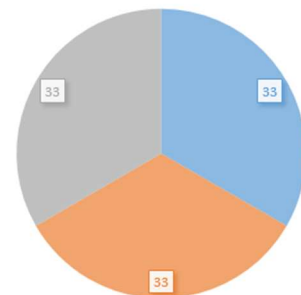


2. Use caution with 3D charts, as the 3D texture may give a false impression of importance. The charts below have the same values, but the 3D effect on the top chart gives false emphasis.

Example of Bad Pie chart



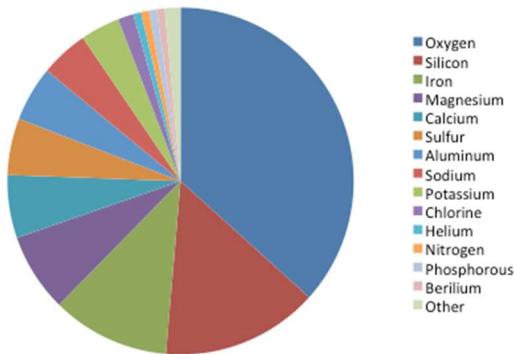
Corrected



3. Be sure to always readable labels for context. Not labeling an axis or bits of data could add unnecessary ambiguity. (see number 1, above) Yes - the chart could be read without it, but it would take more effort by the user, and that is to be avoided.

4. Use a chart type that accurately, and simply, tells your story

Example of incorrectly selected chart type



A better way to show proportion of the elements could be a bar chart, or a line chart, where color is not as important, and the elements can be seen against each other more clearly

Busy to read, and difficult to distinguish between so many colors, especially if converted to black and white

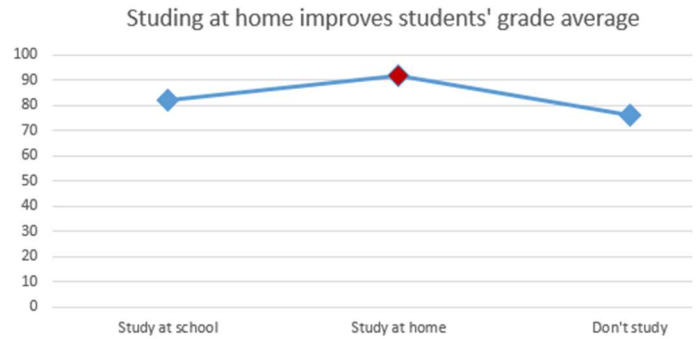
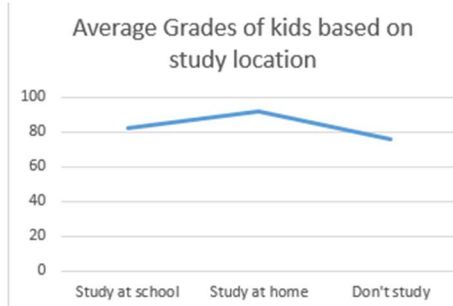
5. Consider adding color or thickness to a particular line or plot points to highlight the most important bits of information
6. Deflate numbers where possible. Example: 1,343,433 could be deflated to 1.34 for clarity and ease in plotting
7. Use a complimentary, yet limited color palette. Use less prominent colors on less prominent information. Use brighter colors only where information needs emphasized. consider how it would look in Black and white.
8. Don't fill with patterns. Not everyone sees as well as you, and it could obscure data labels.
9. Soften or hide gridlines all together, so they do not grab attention.
10. Consider using sans-serif fonts. Serif fonts add fancy decorative bits on the end. Sans-serif fonts are cleaner and more readable, especially for small text. Both fonts below are the same size, but the sans serif Trebuchet font is much cleaner in appearance than the serif Times New Roman.

This is in sans serif font.

This is in serif font.

11. Use bold or italics sparingly and only to emphasize important point or two
12. Don't make text too small to read
13. DON'T USE ALL CAPS AS IT BECOMES MORE DIFFICULT TO READ. ESPECIALLY WHEN COMBINED WITH SERIF, **BOLD, ITALICS, ALTERNATING COLORS, AND FONT SIZE**
14. Don't alternate colors for data points (**white on black ok for headings**, but not for data labels)
15. Title should reflect the meaning of the data. Chart on the left has a generic title. Chart on the right reflects the findings, and uses a highlighted plot point to draw the audience's attention.

Title on the left vague, title on the right implies the effect (improves)



Another example of Instead of a generic title on, title reflects the answer to the question

Instead of:

Speaker	Average time word repeated	Largest occurrence of repeated word	Most frequently used word	Least Frequently used word
Trump	6.9	233	Country	know
Biden	5.6	176	Thank	Americans

Consider instead:

Speaker	Average time word repeated	Largest occurrence of repeated word	Most frequently used word	Least Frequently used word
Trump	6.9	233	Country	know
Biden	5.6	176	Thank	Americans

Activity (Excel, Python, or Tableau)

Resources

Student-selected data-set from any open source (see link for suggested sources)

<https://tinyurl.com/yckuzmsr>

Beginner's Excel Charting tutorial: <https://tinyurl.com/jytvuvdf>

More Excel advanced tips and tricks: <https://tinyurl.com/2ke8dnrv>

Using Matplotlib with Python:

- <https://tinyurl.com/8mr9xmk8>
- <https://tinyurl.com/4vcuzpkj>

Tableau charting videos:

- <https://tinyurl.com/mr23kfcv>
- <https://tinyurl.com/495dpce5>

Steps

1. Select one data set that you can use to demonstrate 3 different answers with charts. Consider using aggregate functions (sum, average, min, max, count) to assist in finding patterns/trends.
2. Use the selected modeling tool (Excel, Matplotlib, or Tableau), create 3-5 different kind of charts to tell the story of your data set.
3. Prepare a presentation on the tool used, your data findings, as well as why you selected each type of chart to tell the data story (see Chart Suggestions - A Thought Starter above). Be prepared to present to your classmates.

Activity (Python)

Resources

Using Matplotlib with Python: <https://tinyurl.com/8mr9xmk8>

Steps

1. Download the sample files, and follow along with the tutorial
2. After completing both tutorials, try cleaning the sample files from the Excel activity by yourself

Activity (Tableau)

Resources

Tableau Public Training Creating your First Chart: <https://www.youtube.com/watch?v=L2D7e9APvTw>

Steps

Use tableau to create the charts/graphs for the sample data in either the Excel or Python

Extension:

Have students work in small groups to research a problem that needs to be solved, collect first hand data, as well as use open source data, and prepare a presentation that uses the data to highlight issues and concerns, and which might even point to a solution



Communicating Data Visualizations

An Expanding the Data Analytics Technician Pipeline Module



This material is based upon work supported by the National Science Foundation under Grant DUE 2055411. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.