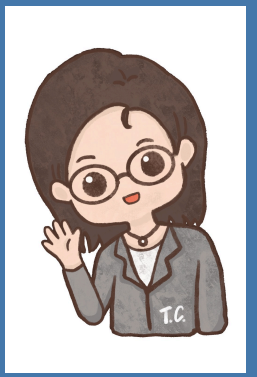
# MTH 208 Exploratory Data Analysis

## Lesson 9: Storytelling with Data

Ying-Ju Tessa Chen, PhD

Associate Professor
Department of Mathematics
University of Dayton

@ying-ju
ying-ju
ychen4@udayton.edu



# Learning Objectives

- Understand the context
- Choosing an Effective Visualization
- Apply principles from "Storytelling with Data" to real data sets.

This lesson is based on the concepts covered in the book: storytelling with data by Cole Nussbaumer Knaflic.

# Understand the context

Exploratory vs. Explanatory Analysis

**Exploratory:**

- understand the data and figure out what might be noteworthy or interesting to highlight to others.

**Explanatory:**

- turn the data into information that can be consumed by an audience.

Here, we focus on **explanatory** analysis and communication.

# Context of who, what, and how

### Who

To whom are we communicating?

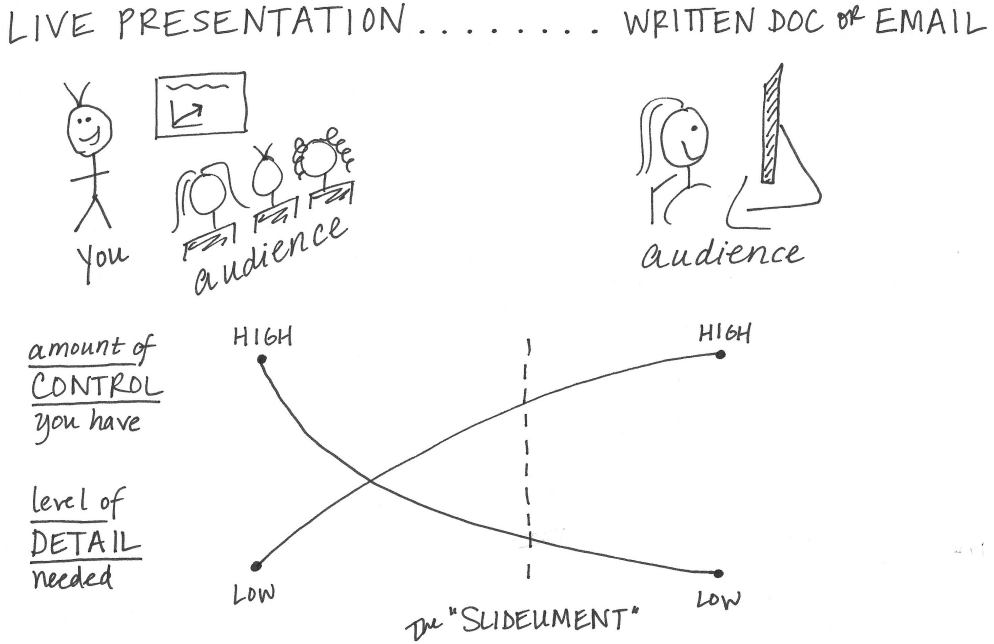
- Your audience: we must understand the audience for our presentation. We should avoid general audience, such as "internal and external stakeholders" or "anyone who might be interested." Sometimes, this means creating different communications for different audiences.
- Think about the relationship that we have with our audience. This helps structure our communication and whether and when to use data, and may impact the order and flow of the overall story we have.

# Context of who, what, how (Continued)

## What

What do we want our audience to know or do?

- How to make what we communicate relevant for our audience and form a clear understanding of why they should care about what you say.
- What is the mechanism that we are going to use to communicate with our audience. Filtering out the most effective medium of communication to interact with our audience.
- What tone do you want your communication to set?



# Context of who, what, how (Continued)

### How

How can we use data to help make our point?

- Only when we have clearly defined who our audience is and what we want them to know or do, we can turn to data and ask the question, "What data is available that will help make my point?" Data becomes supporting evidence for the story we'll create and communicate.
- What datasets/tables/metrics?
- What numbers are meaningful?
- Should we subset the data?
- How should we present it visually?

### Example

`Who:` The budget committee that can approve funding for continuation of the summer learning program.

`What:` The summer learning program on science was a success; please approve budget of $X to continue.

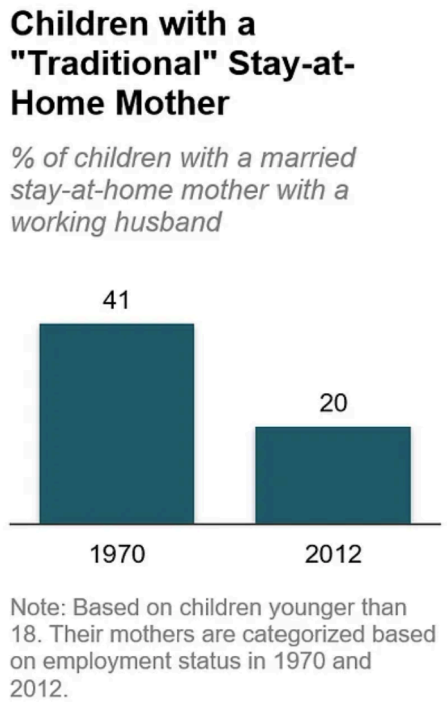
`How:` Illustrate success with data collected through the survey conducted before and after the pilot program.

# Choosing an Effective Visualization

How to adapt to different choices of the graphs given the context? - The choice of the graph, according to the situation and data, is fundamentally important for effective communication.

## Simple text

- When we have just a number or two to share, simple text can be a great way to communicate.




# Choosing an Effective Visualisation (Continued)

## Tables

- Tables are great for communicating to a mixed audience whose members will each look for their particular row of interest.
- If we need to communicate multiple different units of measure, this is also typically easier with a table than a graph.
- Borders should be used to improve legibility of your table

Heavy borders

| Group | Metric A | Metric B | Metric C |
|---|---|---|---|
| Group 1 | $X.X | Y% | Z,ZZZ |
| Group 2 | $X.X | Y% | Z,ZZZ |
| Group 3 | $X.X | Y% | Z,ZZZ |
| Group 4 | $X.X | Y% | Z,ZZZ |
| Group 5 | $X.X | Y% | Z,ZZZ |

Light borders

| Group | Metric A | Metric B | Metric C |
|---|---|---|---|
| Group 1 | $X.X | Y% | Z,ZZZ |
| Group 2 | $X.X | Y% | Z,ZZZ |
| Group 3 | $X.X | Y% | Z,ZZZ |
| Group 4 | $X.X | Y% | Z,ZZZ |
| Group 5 | $X.X | Y% | Z,ZZZ |

Minimal borders

| Group | Metric A | Metric B | Metric C |
|---|---|---|---|
| Group 1 | $X.X | Y% | Z,ZZZ |
| Group 2 | $X.X | Y% | Z,ZZZ |
| Group 3 | $X.X | Y% | Z,ZZZ |
| Group 4 | $X.X | Y% | Z,ZZZ |
| Group 5 | $X.X | Y% | Z,ZZZ |

# Choosing an Effective Visualisation (Continued)

### Graphs

In many cases, there isn't a correct visual display.

> **Most important question:** What do we need our audience to know?

Choose display that will enable to make that clear.

A simple approach to test this is to make a graphic and show it to a friend or colleague, then have them explain the following while processing the information:

- Where they focus
- What they see
- What observations they make
- What questions they have

PS. Since we have covered data visualization fundamentals, we won't emphasize different types of figures.

# Choosing an Effective Visualisation (Continued)

### To be avoided

- `Pie charts or Donut charts:` Replace them with horizontal bar charts
- `3D charts:` Never use 3D charts, only exception is if you are actually plotting a third dimension. We should avoid introducing unnecessary chart elements like side and floor panels and does weird things when plotting values.
- `secondary y-axis:` Think about whether following approaches will meet needs:
  - Don't show the secondary y-axis. Instead, label the data points that belong on this axis directly.
  - Pull the graphs apart vertically and have a seperate y-axis for each (both along the left) but leverage the same x-axis across both.

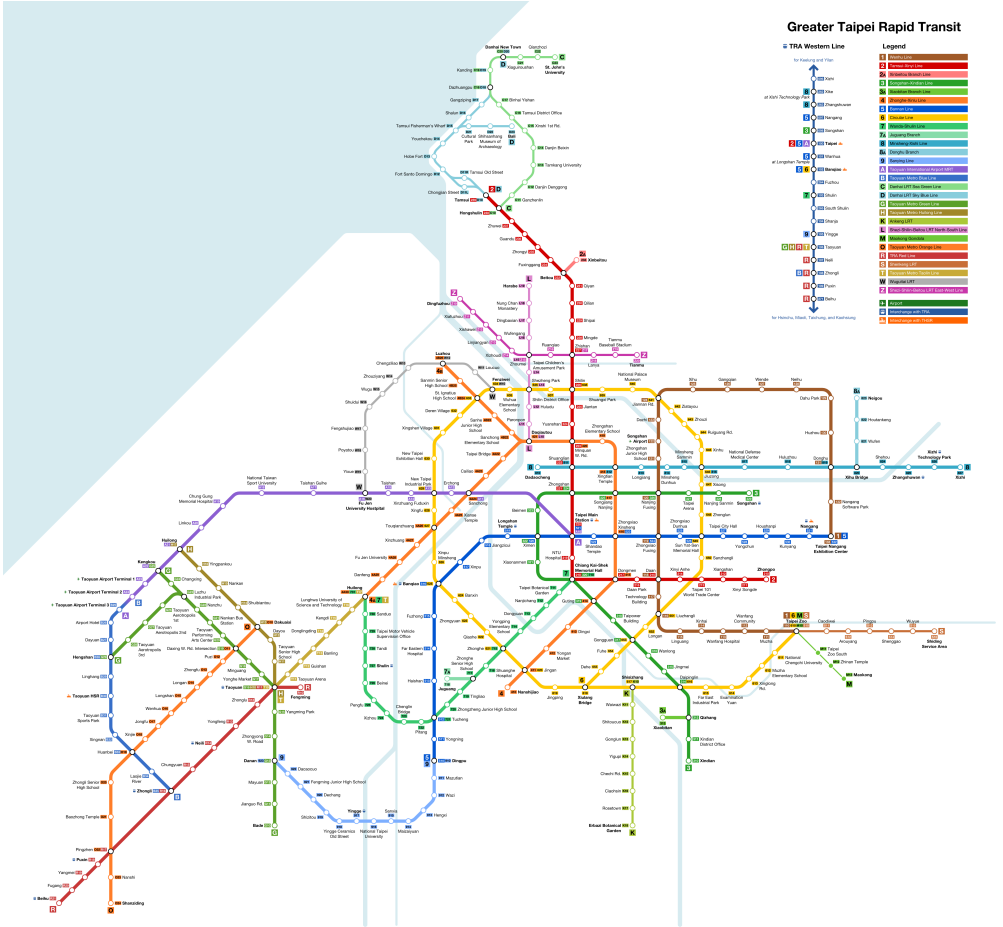
# Clutter is your enemy

Focus on three techniques:

**Decluttering your graph**

- **Why do we need to declutter our graphs?**
  - To reduce cognitive load, helping viewers who may be tired and cranky.
  - Allows viewers to focus on what truly matters.
  - Not all information presented is useful for your viewer.

# Clutter is your enemy (Continued)




# Clutter is your enemy (Continued)

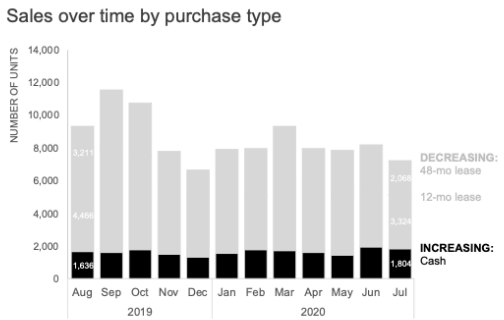
- **Cognitive Load**
  - `Consider our audience:`
  - They may be tired and cranky and prefer direct communication.
  - They can generally hold 5 +/- 2 items in their short-term memory at once.
  - `Remove elements that do not add value:`
  - Avoid shadows and 3D effects.
  - Simplify or eliminate legends and excessive colors.
  - Consider removing axis titles when they are not informative.

# Clutter is your enemy (Continued)

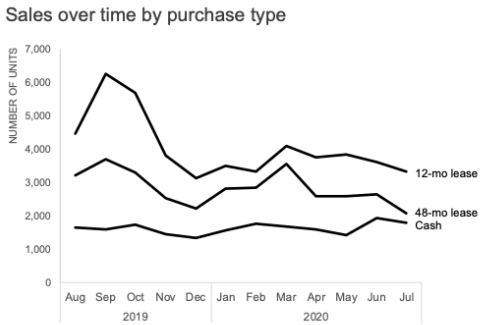
- **Ask Yourself**

> Does this element support the point I want to make about the data?

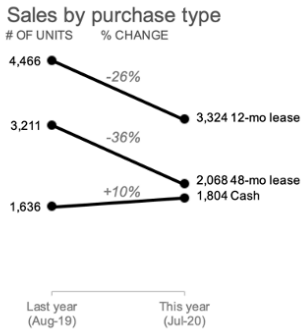
OPTION 1: STACKED VERTICAL BARS



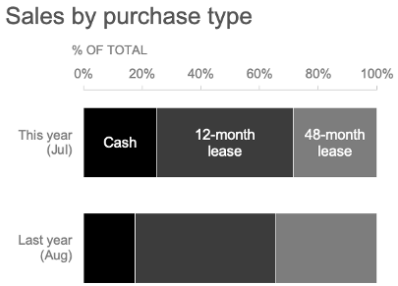

OPTION 2: LINES




OPTION 3: SLOPEGRAPH




OPTION 4: 100% STACKED BARS




Credit for the image: storytellingwithdata.com

# Clutter is your enemy (Continued)

### Annotating your graph

- `Guiding Your Viewer`
  - Annotations can effectively guide viewers through a visualization. Best practices for using annotations:
  - Use your titles and captions!
  - Titles should set the stage and focus attention on the key message.
  - "If there is a conclusion you want your audience to reach, state it in words." — Cole Nussbaumer Knaflic
- `Don't label everything`
  - Only label data points that are crucial to the discussion.
  - If comparing two items, make sure they are clearly labeled.
  - Use groupings or "super categories" to help viewers understand complex data.

# Clutter is your enemy (Continued)

### Highlighting Data

- `Preattentive attributes`
  - Color and contrast are known as preattentive attributes. They capture attention unconsciously before we even process the graph's content.

How many 3s are there in this image?


**Figure 4.2** Count the 3s example

# Clutter is your enemy (Continued)

## How about now?

Notice how quickly you can spot the 3s using color and contrast.

This demonstrates the effectiveness of preattentive attributes in focusing attention.

That's the power of preattentive attributes!

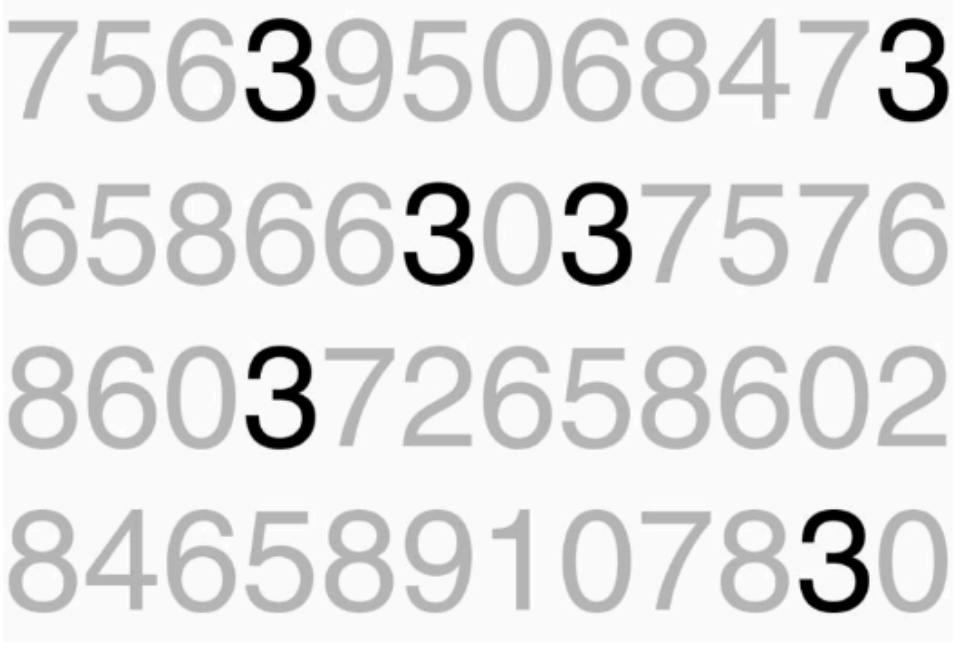



**Figure 4.3** Count the 3s example with preattentive attributes

# Tell the Story!

- `Start with something catchy:` Engage your audience immediately.
- `Share context:` Set the stage for your data by providing background.
- `Illuminate what matters:` Draw attention to the key findings of your analysis.
- `Add details:` Provide deeper insights and strengthen your narrative.
- `Identify limitations:` Enhance credibility by acknowledging the limits of your data.
- `Summarize:` Reinforce the key points, ensuring they are understood and remembered.
- `Reiterate the Message:` Leave a lasting impression with a clear takeaway message.

> Tell them what you talk about, talk about it, tell them what you talk about

# Presentation Tools: Enhancing Your Data Storytelling

- `Background and Colors:` Avoid using a white background for your slides. White backgrounds can be glaring and hard on the eyes, especially in dimly lit rooms or during long presentations.
- `Practice on the Slides:` Familiarity with your slides and how they transition will make your delivery smoother and more confident.
- `Become Data Translators:` Bridge the gap between complex data insights and your audience's understanding.
- `Listening is Communication:` Effective communication is as much about listening as it is about speaking. Encourage questions and feedback. Show that you value the audience's input by addressing their concerns and adapting your presentation to meet their interest levels and understanding.
- `Stories Follow a Clear Structure:` Organize your data presentation like a story with a clear beginning, middle, and end.
- `Practice with Visuals:` Use visuals strategically to enhance your storytelling.
- `Create Shared Understanding:` Ensure that everyone in the audience, regardless of their expertise, leaves with a common understanding of what the data shows.

## Engaging Tips

Here are additional tips to consider when preparing and delivering your data presentation:

- `Visual Contrast:` Use high contrast between text and background for readability.
- `Fonts:` Choose large, clear fonts. Sans-serif fonts like Arial or Helvetica are often easier to read in presentation formats.
- `Animations and Transitions:` Use these sparingly. Ensure they add value to your presentation and do not distract from the content.
- `Interactivity:` Depending on the setting, interactive elements like polls, live data queries, or interactive graphs can engage the audience and make your presentation more memorable.

# Case Studies

juiceanalytics

# References

The lectures of this course are based on the ideas from the following references.


- Exploratory Data Analysis by John W. Tukey
- A Course in Exploratory Data Analysis by Jim Albert
- The Visual Display of Quantitative Information by Edward R. Tufte
- Data Science for Business: what you need to know about data mining and data-analytic thinking by Foster Provost and Tom Fawcett
- Storytelling with Data: A Data Visualization Guide for Business Professionals by Cole Nussbaumer Knaflic