Data Science II
- Storytelling with Data -

Prof. Dr. Eduard Kromer
Summer Semester 2024
University of Applied Sciences Landshut

Telling a Story with Data

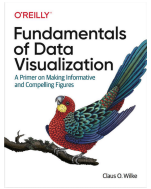
“We aren’t naturally good at storytelling with data.

In school, we learn a lot about language and math. On the language side, we learn how to put words together into sentences and into stories. With math, we learn to make sense of numbers. But it’s rare that these two sides are paired: no one teaches us how to tell stories with numbers. Adding to the challenge, very few people feel naturally adept in this space.”

–C.N. Knaflic - in Storytelling with Data

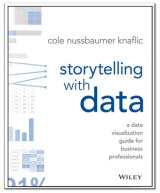
Telling a Story with Data

- the purpose of data visualization (in most cases) is communication
- your intention is to communicate insights about a data set to a specific audience
- for a successful communication you need to create a meaningful and exciting story
- if you don't provide a meaningful story your audience will make one up



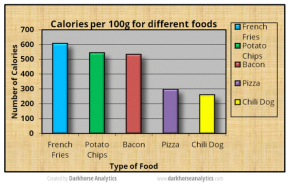
Telling a Story with Data

- technology has enabled us to amass great amounts of data and there is an increasing desire to make sense out of all this data
- meaningful visualizations and the ability to use those visualization to tell a story about your data is an important task that is increasingly in demand
- anyone can put data into a graphing application, like Excel, and create a graph, but without a clear path to follow this leads us in really bad directions: 3D, meaningless color, bad pie charts …



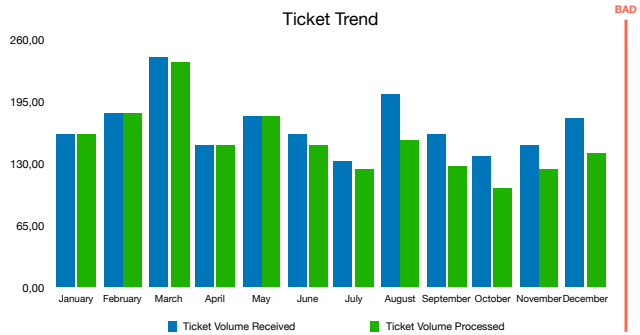
The Data-Ink Ratio

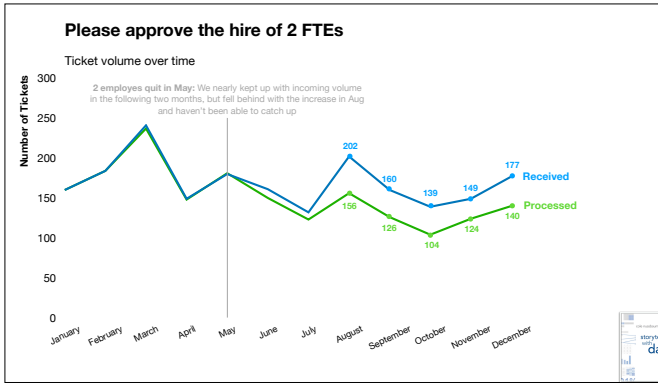
Remove backgrounds



Telling a Story with Data

- there is a story in your data
- your tools don't know what that story is
- it takes you to bring that story visually and contextually to life





Exploratory vs. Explanatory Analysis

- exploratory analysis: what you do to understand the data and figure out what might be interesting to highlight to others
- explanatory analysis: how to communicate your analysis to a specific audience
- you are probably tempted to show your exploratory analysis, but you should take the time to *turn data into information that can be consumed by a specific audience*

Who, What, and How

- Who: To whom are you communicating?
  - it is important to have a good understanding of who your audience is
  - What do you want your audience to know?
  - think about the relationship that you have with your audience
- What: What do you need your audience to know or do?
  - How to make what you communicate relevant to your audience?
  - How to form a clear understanding of why they should care about what you have to say?
  - you should always want your audience to know or do something

Who, What, and How

- How: How to present our data visually? (main focus of this course)
  - recall all the previous lectures of this course and the methods we discussed
  - don't show only the data that backs up your point and ignore the rest
    - it is misleading and paints a one-sided story
    - a discerning audience will poke holes in a story that doesn't hold up

Think Like a Designer

- visual affordances: indicate to your audience how to use and interact with your visualizations
  - highlight the important stuff
  - eliminate distractions
  - create a clear hierarchy of information

Highlight the Important Stuff

- it is critical to only highlight a fraction of the overall visual, since highlighting effects lose weight as the percentage that are highlighted increases
  - at most 10% of the visual design should be highlighted

Eliminate Distractions

- the decision of what to cut or de-emphasize can be even more important than what to include or highlight
- not all data are equally important
- when detail isn't needed, summarize
- ask yourself: Would eliminating this change anything?
- push necessary, but non-message-impacting items to the background

Create a Hierarchy of Information

- we can visually pull some items to the forefront and push other elements to the background, indicating to our audience the general order in which they should process the information we are communicating

Accessibility

- the concept of accessibility says that designs should be usable by people of diverse abilities (widely varying technical skills)
  - make it legible: use a consistent, easy-to-read font
  - keep it clean: visual affordances
  - use straightforward language: choose simple language over complex
  - remove unnecessary complexity: favor simple over complicated

Text is Your Friend

- thoughtful use of text helps ensure that your data visualization is accessible
  - use it to label, introduce, explain, reinforce, highlight, recommend and tell a story
- don't assume that two different people looking at the same data visualization will draw the same conclusion
  - if there is a conclusion you want your audience to reach, state it in words

Aesthetics

- be smart with color: use of color should always be an intentional decision
  - use color sparingly and strategically to highlight the important parts
- pay attention to alignment: organize elements on the page to create clean vertical and horizontal lines to establish a sense of unity and cohesion
- leverage white space: preserve margins; don't stretch your graphics to fill the space or add things simply because you have extra space

Summary

1. Understand the context
2. Choose an appropriate display
3. Eliminate clutter
4. Draw attention where you want it
5. Think like a designer
6. Tell a story

Choosing the Right Visualization Software

The Right Visualization Software

The best visualization software is the one that allows you to make the figures you need.

Literature