Data Visualization

Content from Kieran Healy, Data Visualization

SML201: Introduction to Data Science, Spring 2020
Michael Guerzhoy
What makes for good DataViz?

Charles Joseph Minard

• Presents information in a comprehensible, non-misleading, and memorable way
• A creative process
Good DataViz?

- Memorable
- Not misleading
- (Probably) comprehensible

*Monstrous Costs* by Nigel Holmes
Bad DataViz: Bad Taste

- Not as easy to read as it could be
- Tacky
- “Chart junk”
- Not so terrible
Bad DataViz: Bad Taste

Number of o-rings damaged depending on temperature: the figure in the report

A much better figure
Bad DataViz: Bad Data

- Looks nice
- Implies people were asked “Is it important to live in a democracy,” and we are plotting the percentage of “yes” responses
- Actually, we are plotting # of people who answered 10/10 to “On a scale of 1-10, how important is it to live in a democracy”
Bad DataViz: Bad Data

There is a decline, but this is not NYT material anymore, now that we measure the average rating.
Bad DataViz: Perception - Unfriendly design

Hard to figure out the values!
Bad DataViz: Uninspired Design

Difficult to see the trends for the different types
Bad DataViz: Honesty Issues

Same data, but the graph on the left is quite misleading
Preattentive search

- Compare how much color “pops out” compared to shape
- More data makes things much more difficult
  - Searching for the blue circle on the right is very difficult
Gestalt rules

- Humans readily perceive patterns (real or not) based on proximity, similarity, connection, continuity, closure, belonging to background/foreground, common movement.