

Name: \_\_\_\_\_

Consider the two examples from Tufte (1997) “Visual and Statistical Thinking: Displays of Evidence for Making Decisions.”

1. Cholera

- What are some of the strengths of the cholera visualization that helped scientists get to the root of the problem?
- Were there aspects of the visualization that clouded the story or made the solution harder to understand / communicate? Explain.

2. Challenger

- What are some of the aspects of the Challenger visualizations that made the story harder to understand / communicate?
- Were there aspects of the visualization that were creative or helpful? Explain.

Solution:

1. Cholera

- What are some of the strengths of the cholera visualization that helped scientists get to the root of the problem?
  - Snow thought carefully about how to present the data. Instead of simply looking at the data as counts or frequencies, he looked at the death spatially - on a map of the area.
  - Snow was able to **compare** different groups (brewers and employees at a work-house).
- Were there aspects of the visualization that clouded the story or made the solution harder to understand / communicate? Explain.

Some aspects of the graph can be scrutinized in terms of choices: Did the individuals die at the place on the map? Live at the place on the map? Which (types of) individuals were missing from the graph? Missing at random? What decisions did he make in creating the graph (axes, binning of histogram bars, time over which data are plotted, etc.) that change the story needing to be told?

2. Challenger

- What are some of the aspects of the Challenger visualizations that made the story harder to understand / communicate?
  - The engineers who understood the problem created tables and engineering graphs which were
    - \* Not visually appealing.
    - \* Not decipherable to the layman (e.g., “At about 50° F blow-by could be experienced in case joints”)
    - \* There was also no authorship (reproducibility!). **Figures should always have both accountability and reproducibility.**
  - The information provided included very relevant points (about temperature) and superfluous information unrelated to temperature. The univariate analysis was insufficient because the story the data were trying to tell was about the **bivariate relationship between temperature and o-ring failure**.
  - Missing data created an illusion of lack of evidence, when in fact, the true story was quite strong given the full set of information. (92% of the temperature data was missing from some of the most vital tables.)
  - Anecdotal evidence was misconstrued: SRM-15 at 57F had the most damage, but SRM-22 at 75F had the second most damage.
  - In the end, the shuttle launched on a day which was an extrapolation from the model suggested by the data. They had never launched a shuttle at temperatures of 26° – 29°F.
  - Tufte goes on to describe many ways which the final presentation by the engineers to the administrators was inadequate: disappearing legend (labels), chartjunk, lack of clarity depicting cause and effect, and wrong order.
- Were there aspects of the visualization that were creative or helpful? Explain.
 

Not really.