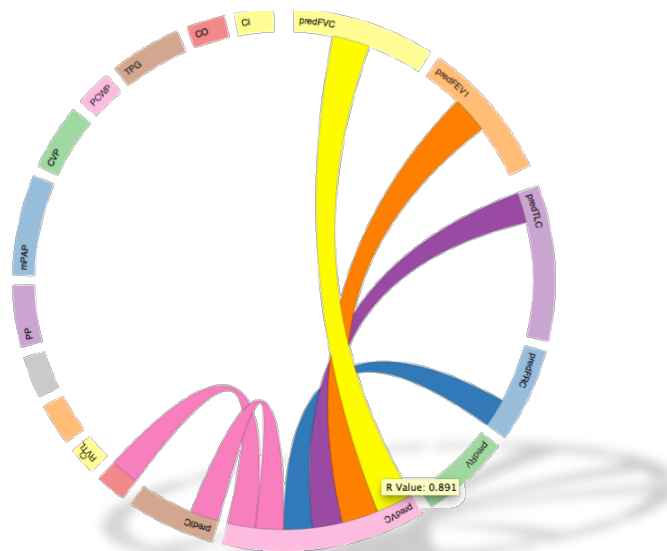


Module #1c

Data Visualization Lies

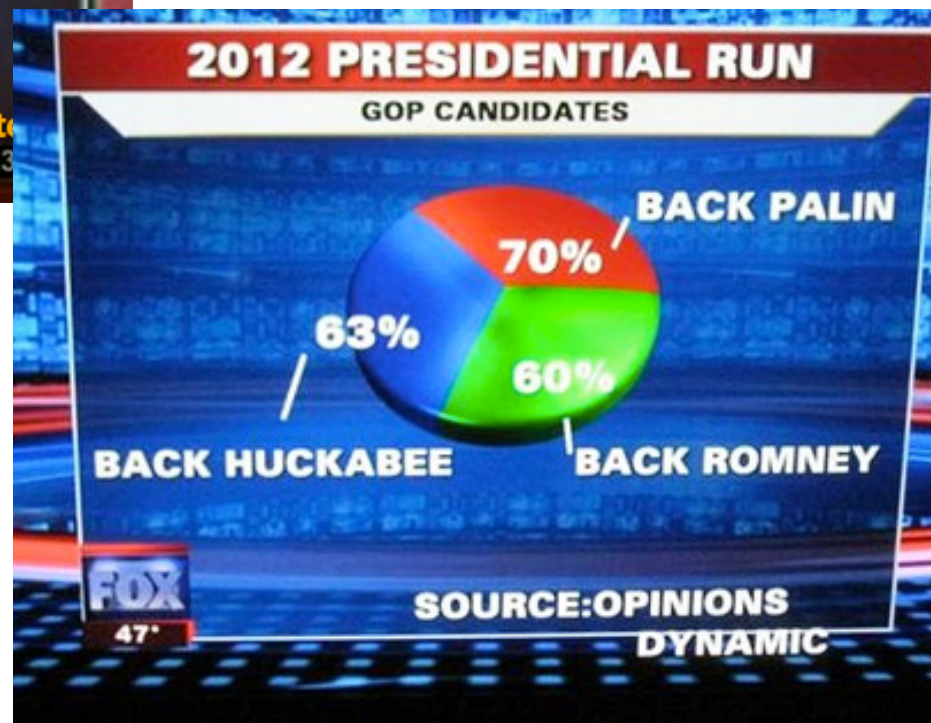




CONSIDERATIONS – **VIS LIES**

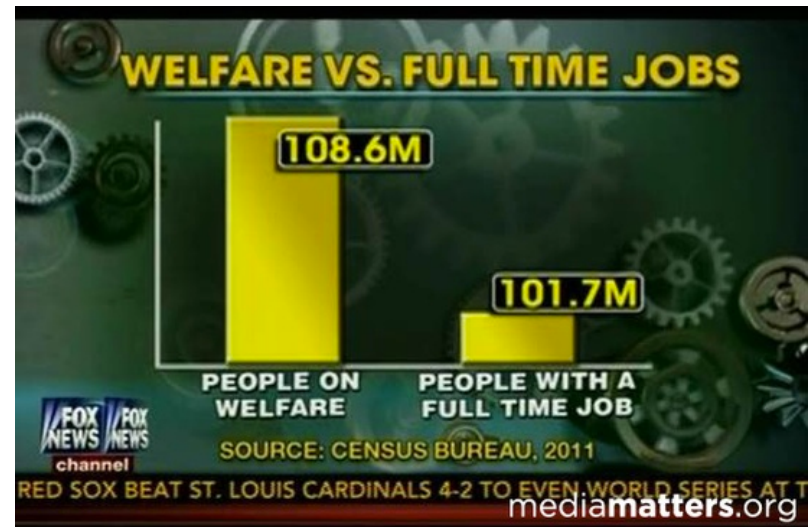


What are some of the most common data misrepresentations?



1. Truncated Y-axis

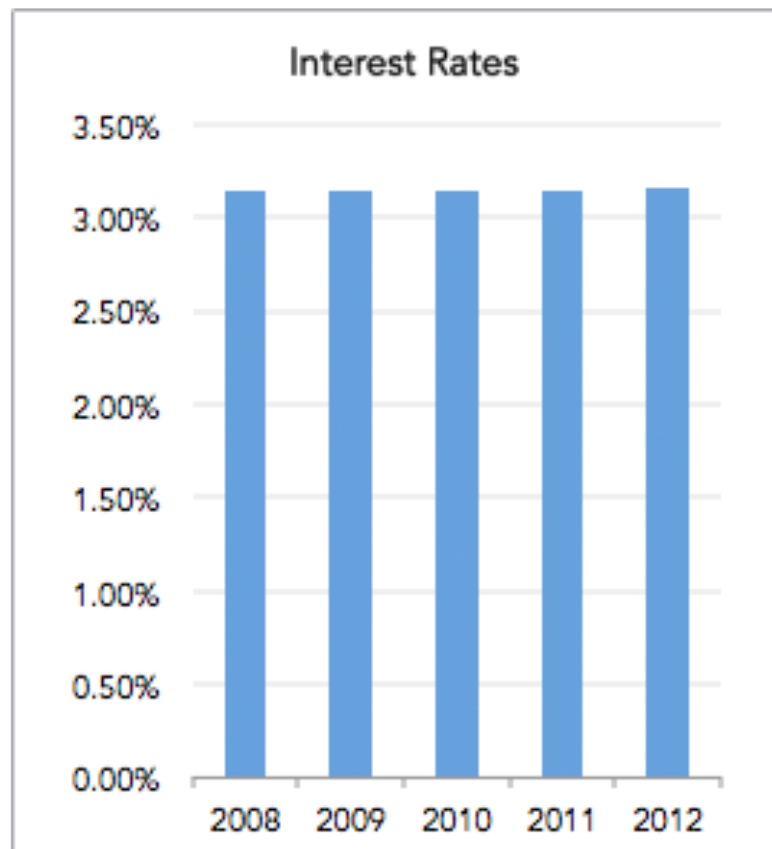
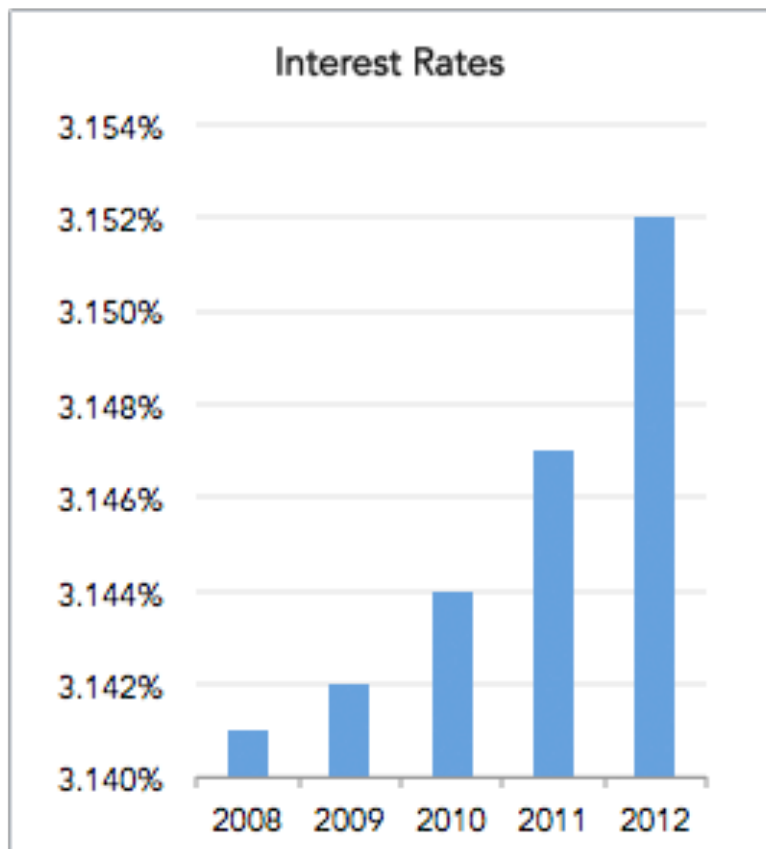
- Infamous for its overuse in politics, the truncated y-axis is a classic way to visually mislead.
- At first glance, the visual dynamics of the graph suggest people on welfare to number four times as many as people with jobs.





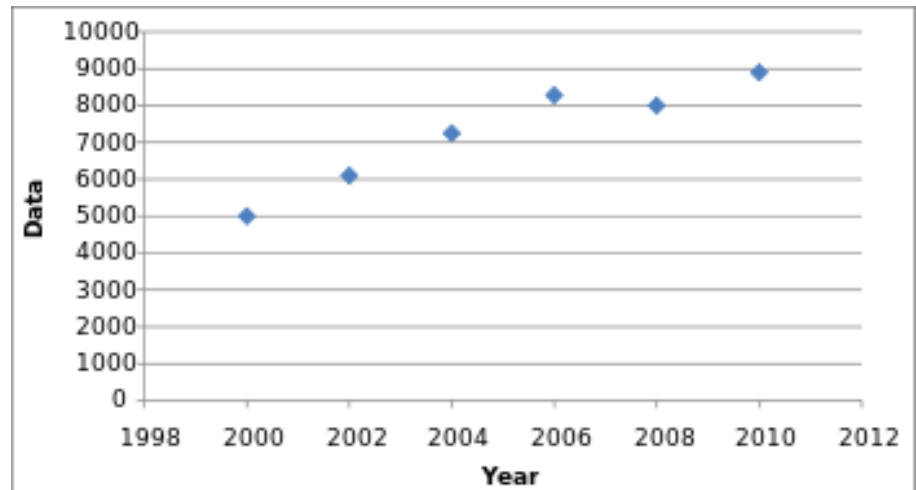
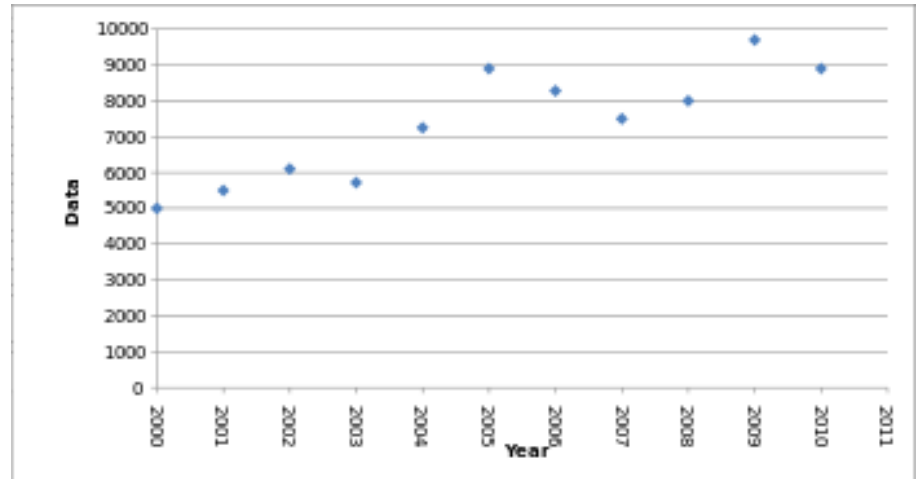
1. Truncated Y-axis

Same Data, Different Y-Axis



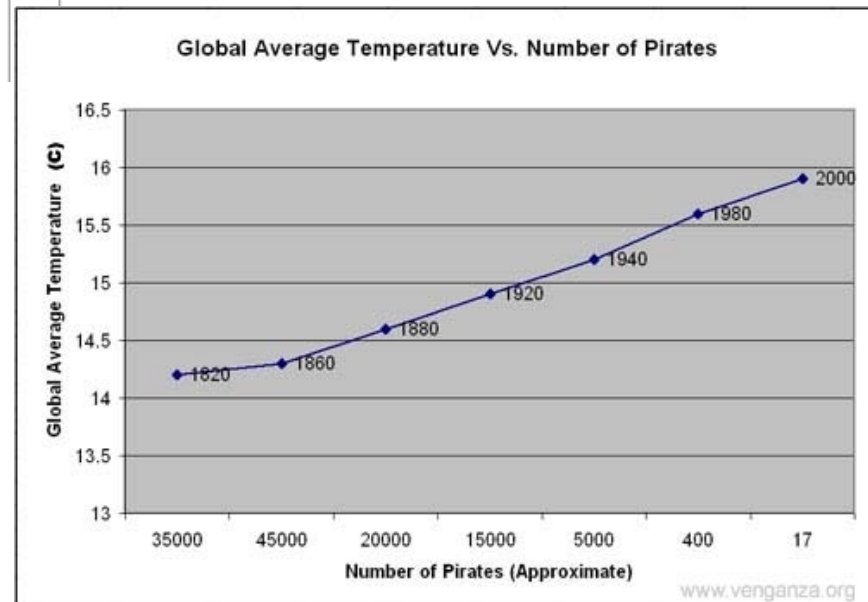
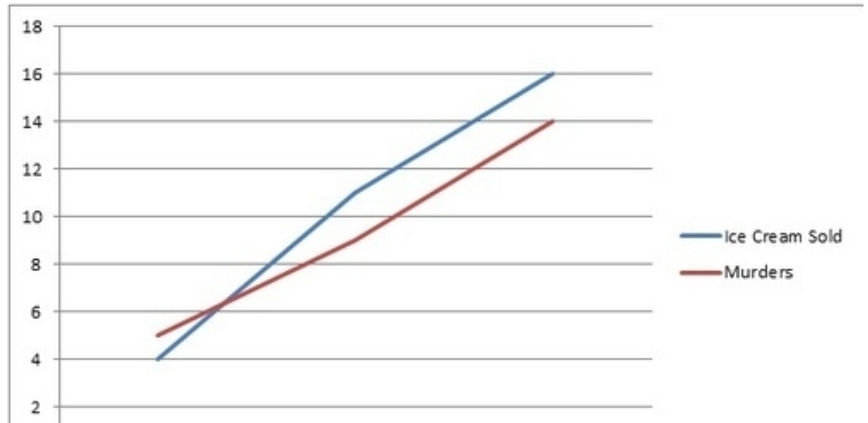
2. Omitting Data

- By omitting certain data points, trends that don't actually exist can easily be created whereas some existing highlights can go unnoticed.
- leaving out some data points, the chart that normally would be filled with dips and spikes, looks much smoother and more stable.



3. Correlating Causation

- Data scientists are finding statistical patterns in data and sometimes care more about -correlation rather than causation.





Conclusion

- Introduction to Data Visualization
- History of Information Visualization
- Common mistakes and misrepresentation of data



JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

© The Johns Hopkins University 2016, All Rights Reserved.