How To Lie With Charts

by Gerald Everett Jones

Florian Fallenbüchel

Seminar “How do I lie with statistics?”

Supervisor: Prof. Dr. Ullrich Köthe

Heidelberg, 24.10.2019

Outline

- Introduction
- Pie Charts
- XY-Charts
- Trends
- Radar Charts

Introduction

- Modern tools allow easy creation of various charts
  - Design choices made by algorithm
  - Little thought required
  - Charts abstract and distort reality
  - Favor (un)intentional misuse

- Internet grants access to a lot of (un)useful information
  - Everyone can reach many people through social media
  - Everyone can create charts at a click

➢ We need to learn about distortions to recognize them!

Introduction

- The numbers don't lie – do they?
  - Decisions based on solid data *must* be reliable!?
- Remember discussion about color of this dress?
  - Human perception is subjective
  - Data is collected by humans
    - Biased labels
    - Biased numbers

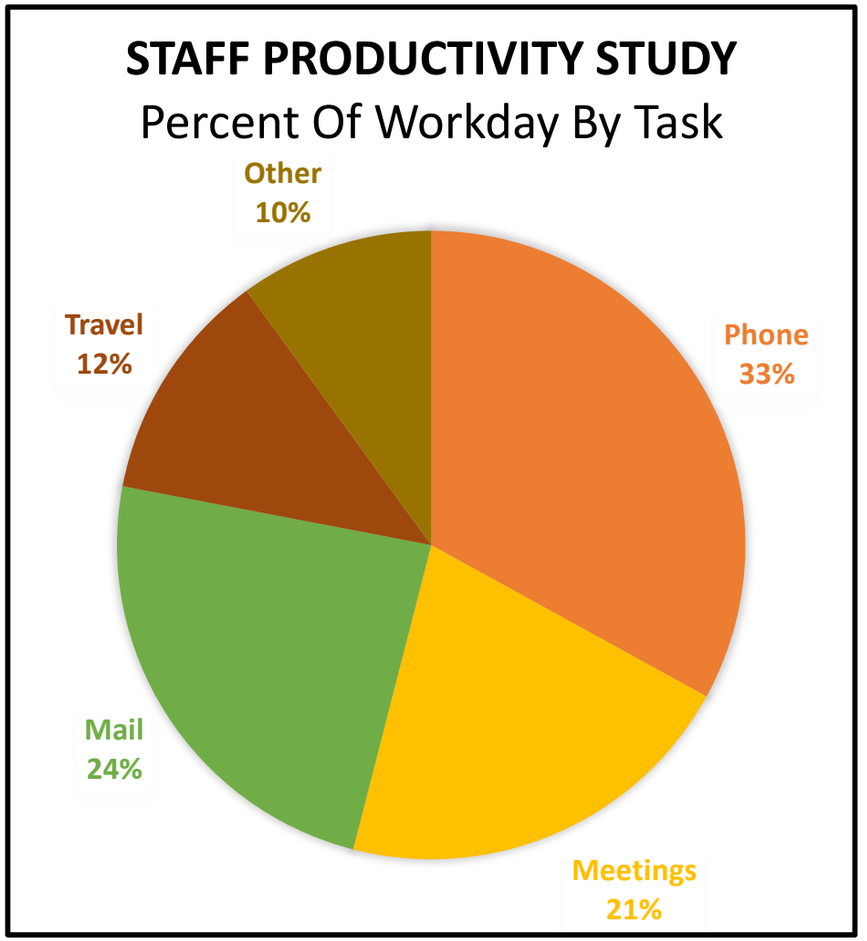
Introduction

- Simply counting things already requires interpretation of reality
  - Example: counting fruits in the next supermarket, resulting number: 42
    - 42 what? Just apples or every fruit?
    - Counted the rotten fruit?
    - Pieces or boxes?
    - Boxes or crates?

➢ You cannot trust a number or chart that you haven't checked yourself!

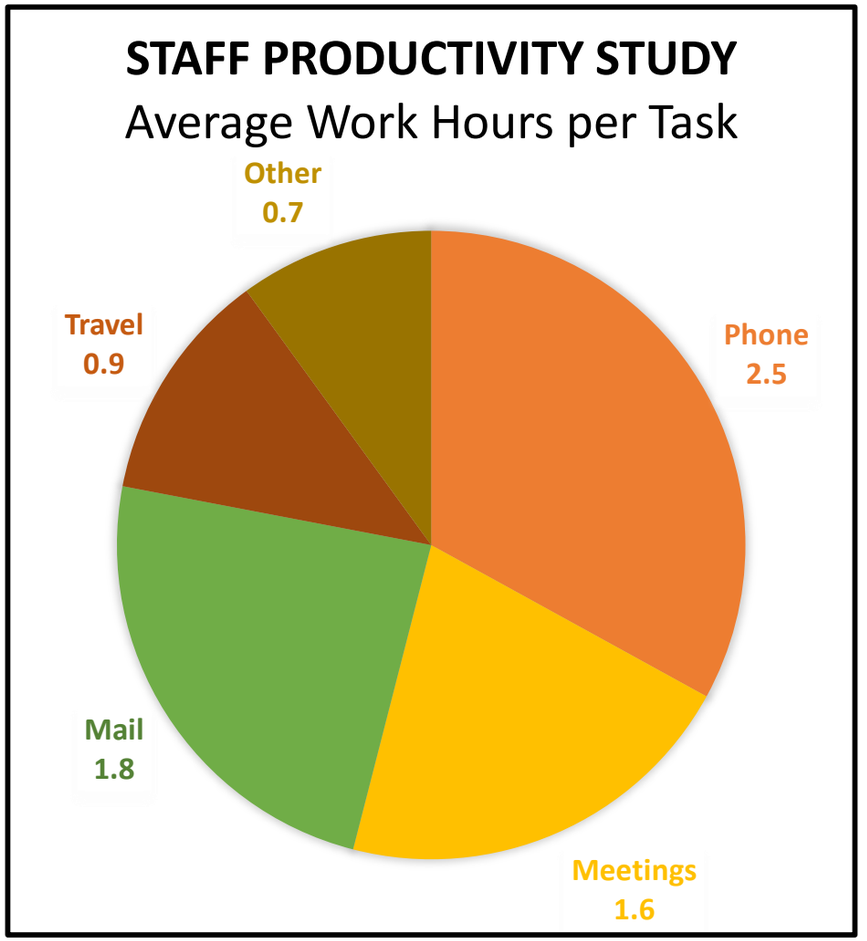
Pie Charts

- Pie charts are only for percentages!
  - Purpose to focus audience on proportions
  - Whole amount should not be important
    - No actual numbers
    - Always labeled with percentage
  - Title should label whole chart



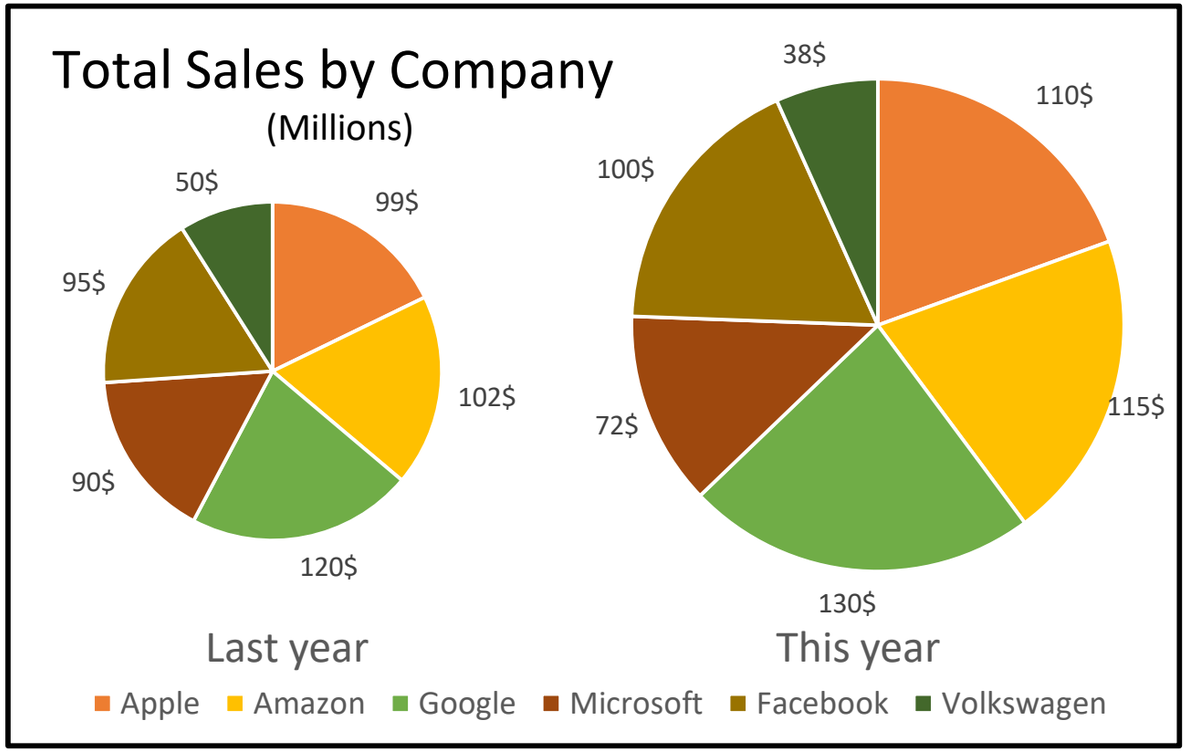
Pie Charts

- Putting actual values on slices distracts from message
  - Audience tempted to sum the values
  - If they don't, they still have a mental impression
- Here: values sum up to 7.5 hours
  - But most staff works 8 hours or more
  - Difference through accounting part time workers
  - ➢ Management might be worried about how staff spends remaining 0.5 hours



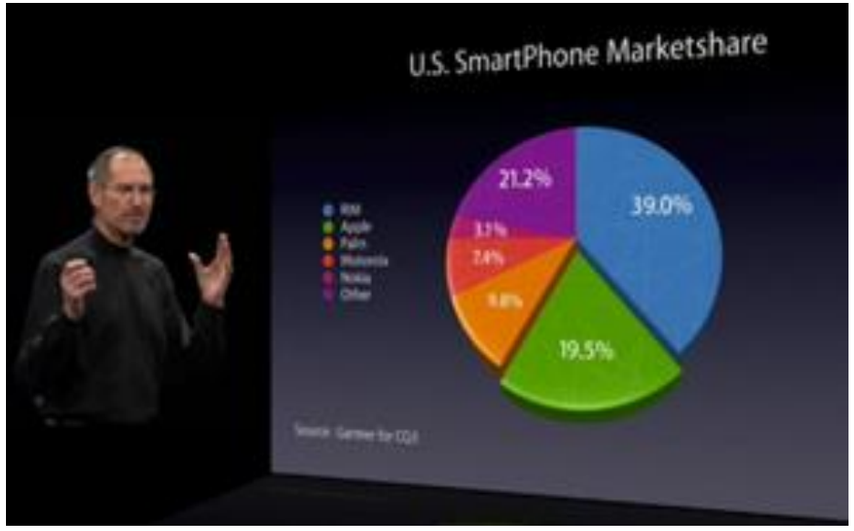
Pie Charts

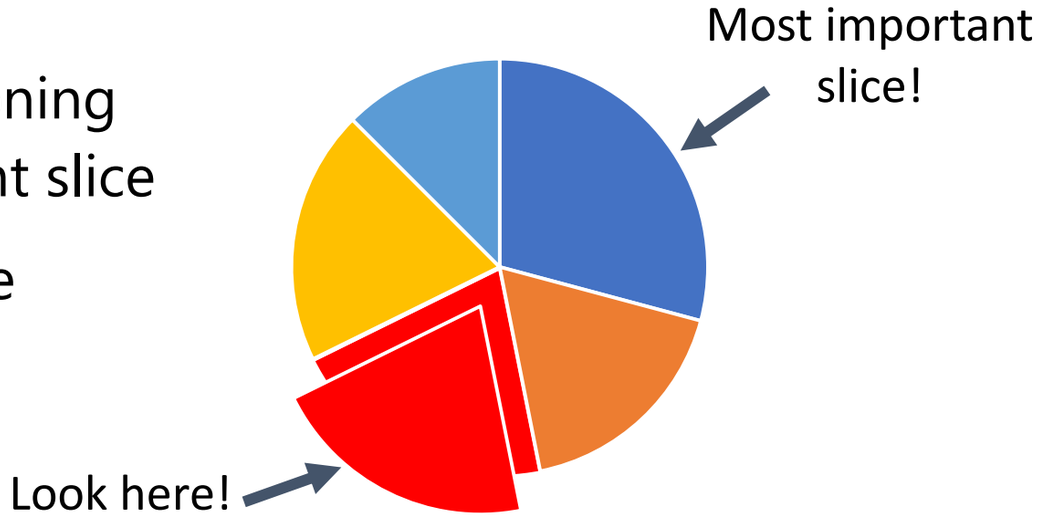
- Messing with size confuses even more
- Example: comparison of last years company market share with current year
  - Size falsely adjusted to reflect total gain
  - Sales dollars instead of market share
  - Top companies prominently placed
  - ➢ Loss of VW and MS vanishes

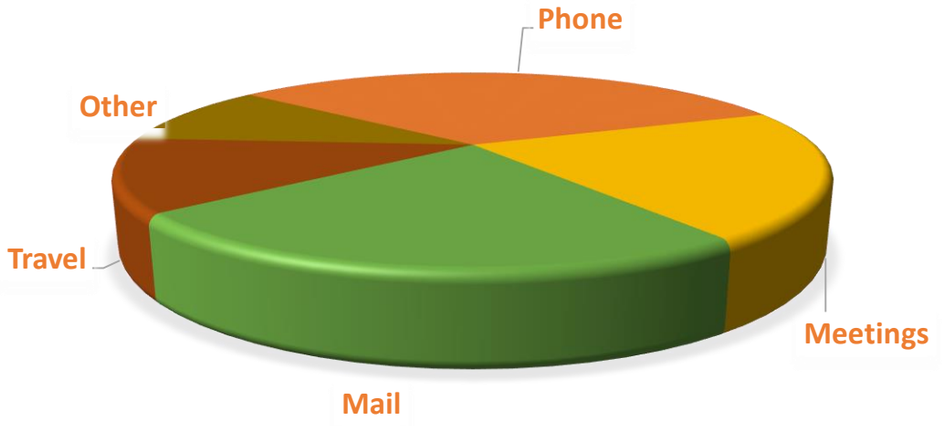


Pie Charts

- Certain slices can be emphasized with strategic positioning
  - People usually give most importance to upper-right slice
  - Slices can be exploded to further stress importance
    - Distract from unfavorable slices
  - 3D pie chart emphasizes bottom slice
    - Distorts apparent size with lower edge

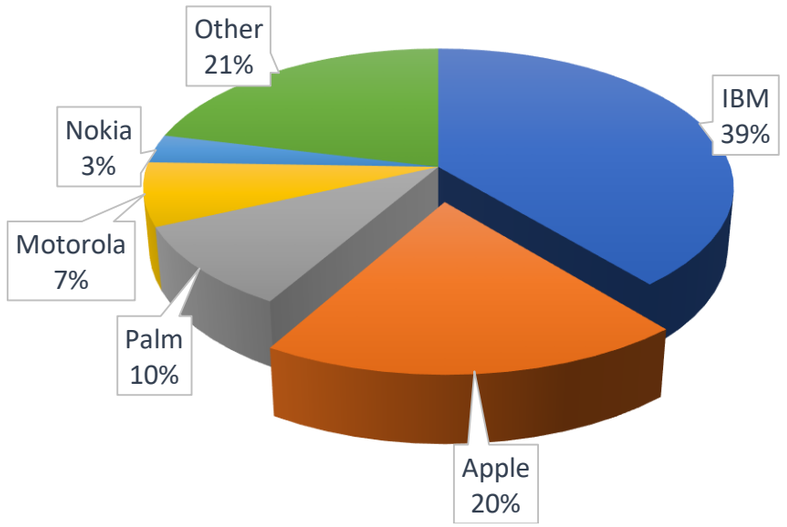


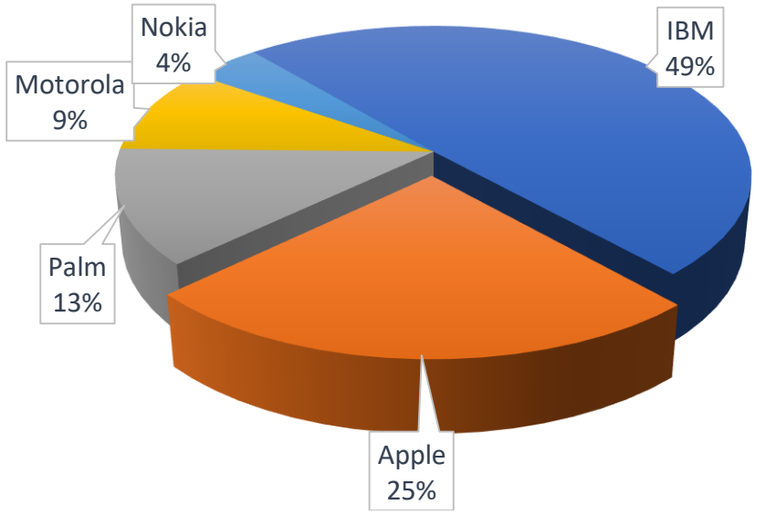




Pie Charts

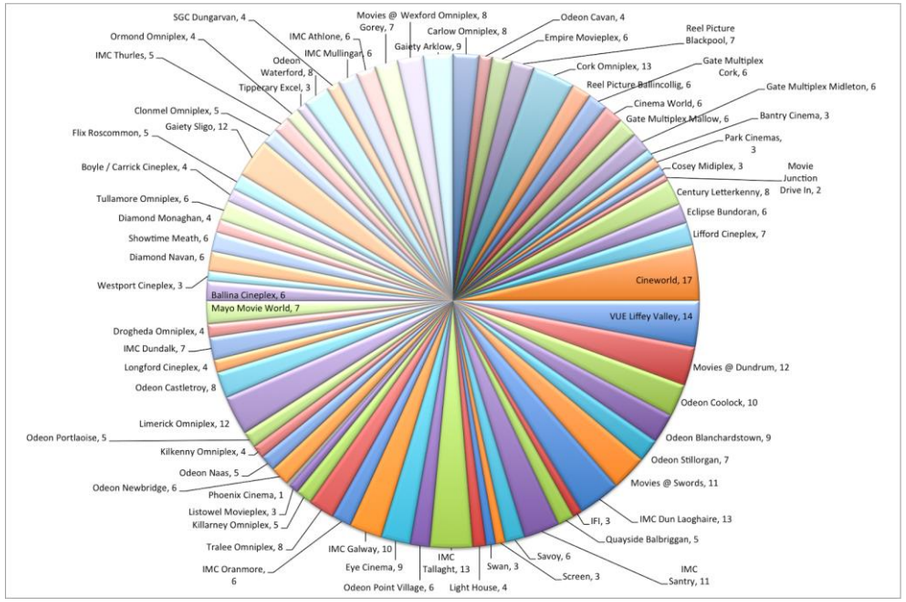
- Abuse the “All Others” category
  - Inconvenient data can be put into mystery slice
    - Might hide valuable data
  - If data complicates story: exclude from pie
    - Remaining slices grow



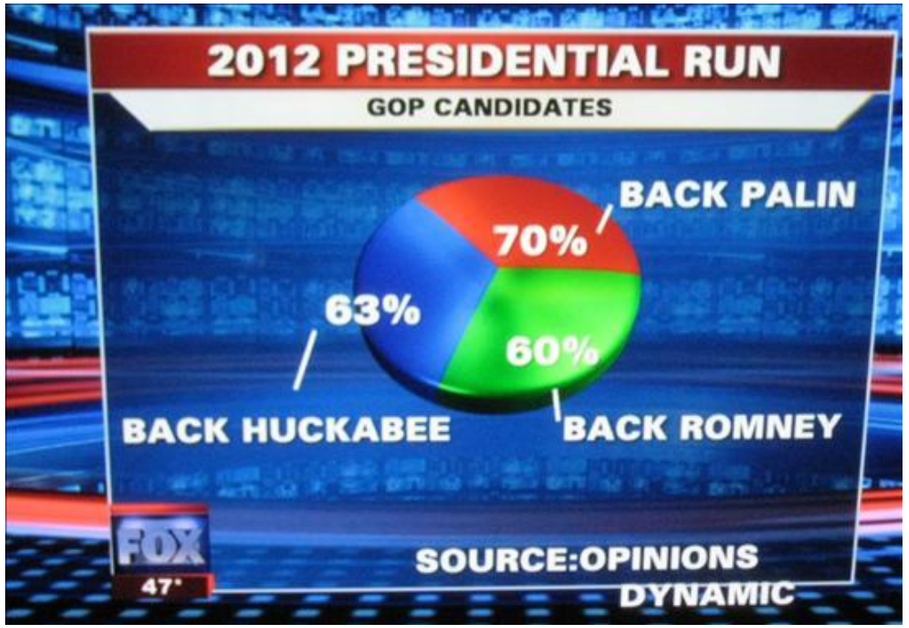


Pie Charts

- Cluttered pie charts hinder proportional comparison
  - Hides inconvenient data
  - Almost no information gain

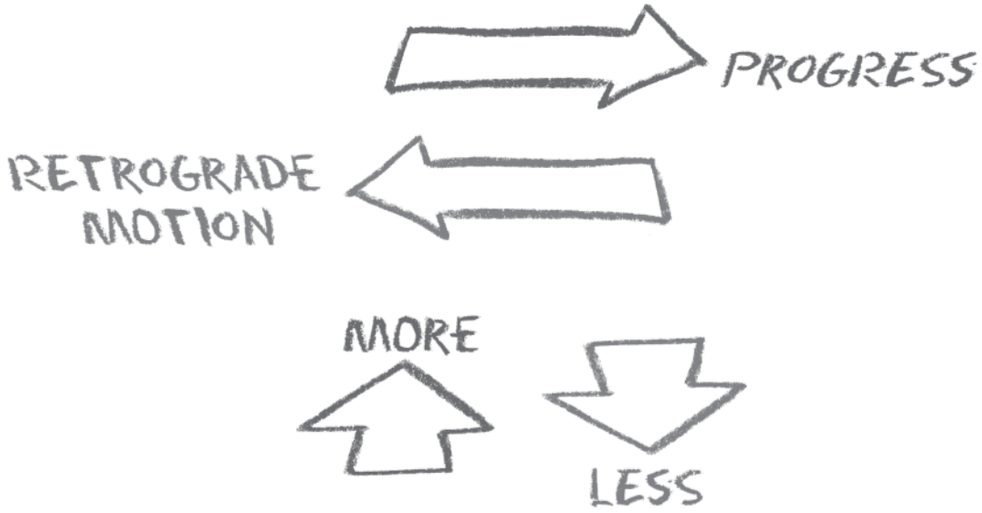


- Percentage labels not summing up to 100%



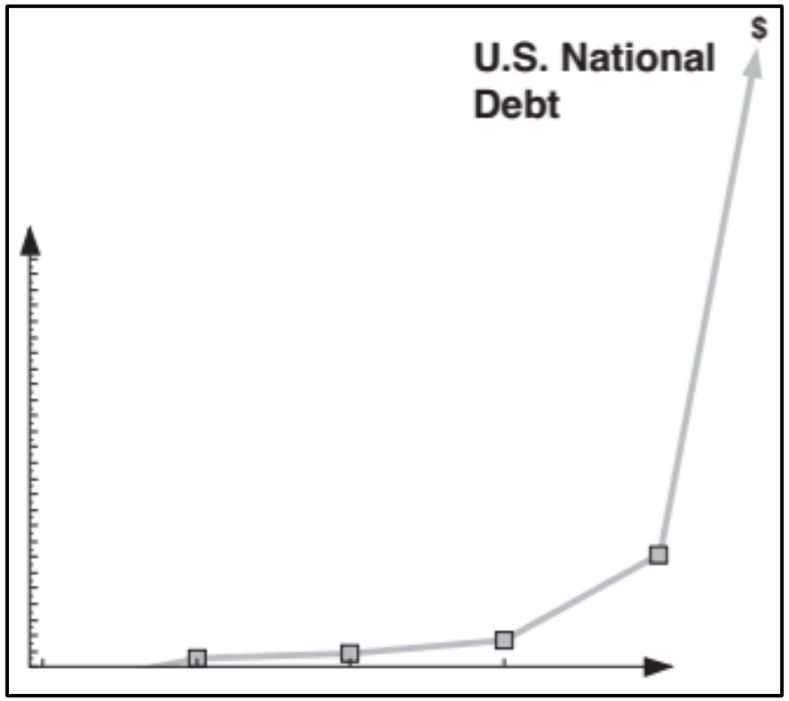
XY-Charts – Orientation

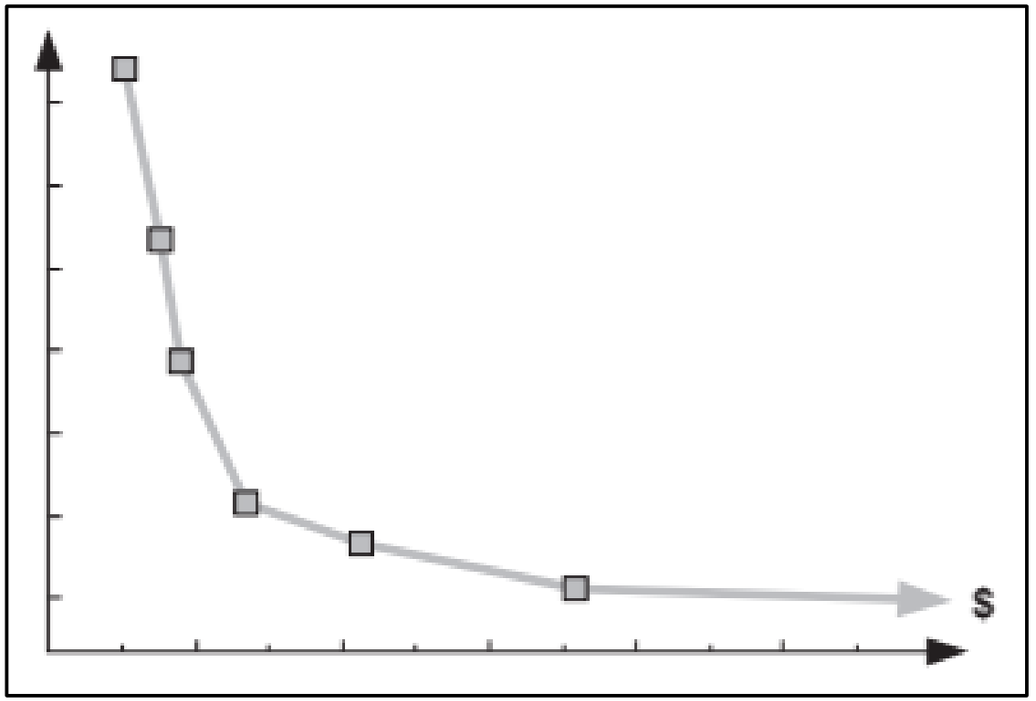
- Every audience has subconscious assumptions about meaning of orientation
  - Most cultures associate upwards movement with gain, downwards with loss
  - Western audiences read from left to right
    - Rightward motion associated with progress / time / positive movement
    - Conversely leftward motion considered backward / bad
    - These effects combine



XY-Charts – Orientation

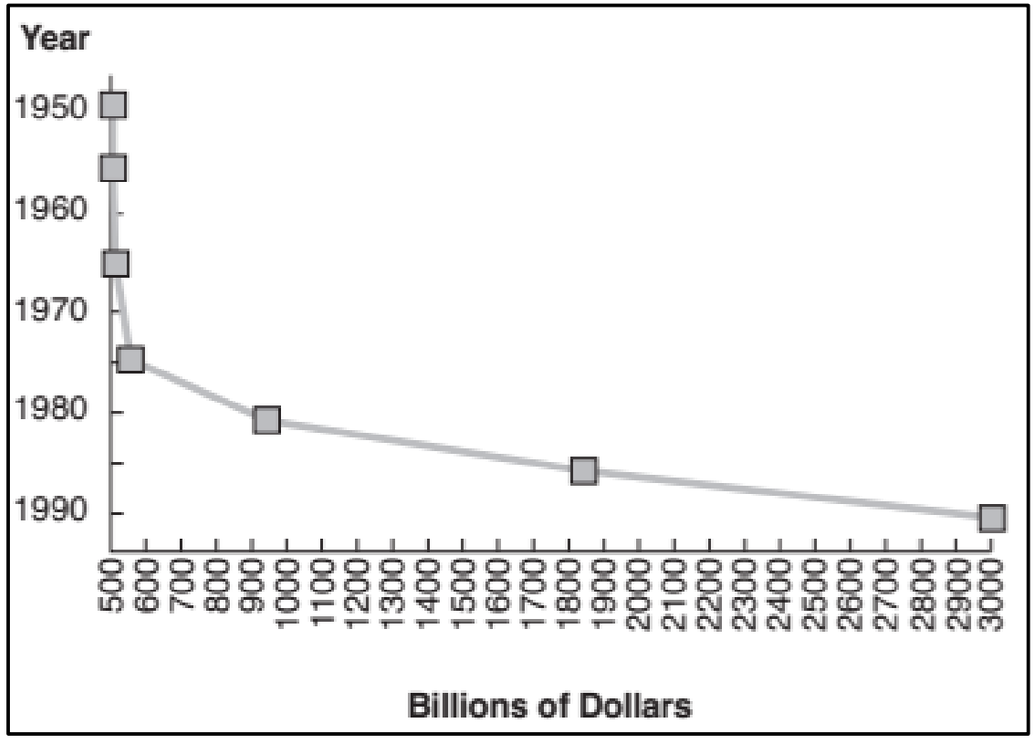
- Orientation greatly influences impression of XY-chart
  - Implies message without much context or labeling





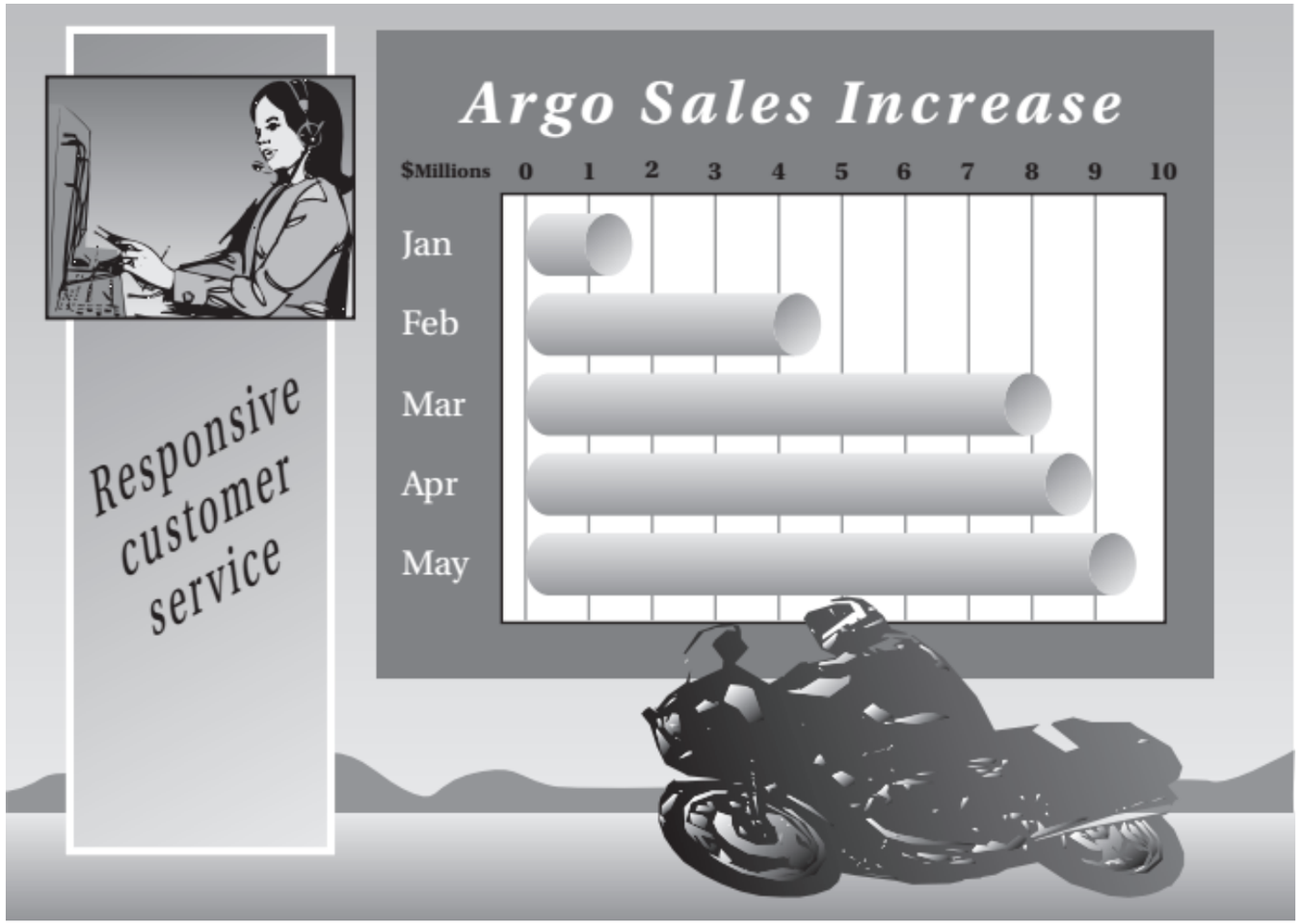
XY-Charts – Orientation

- Original message disguised through unconventional orientation
  - Distracts audience
  - Less concerning despite higher accuracy
  - Perfectly valid math wise



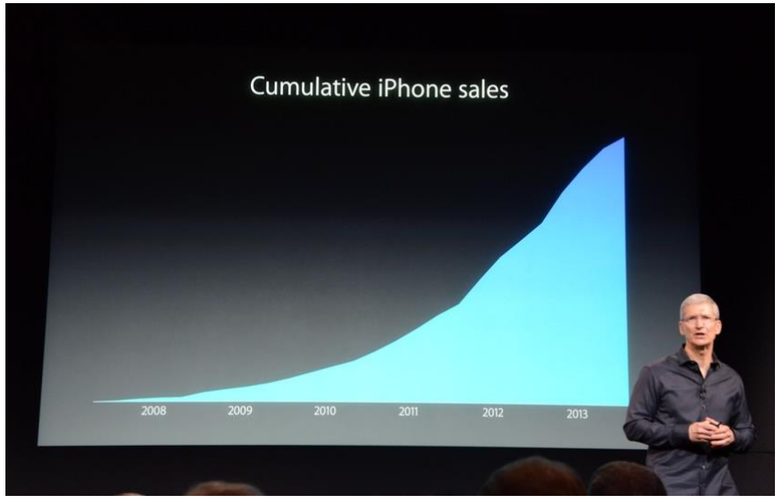
XY-Charts – Orientation

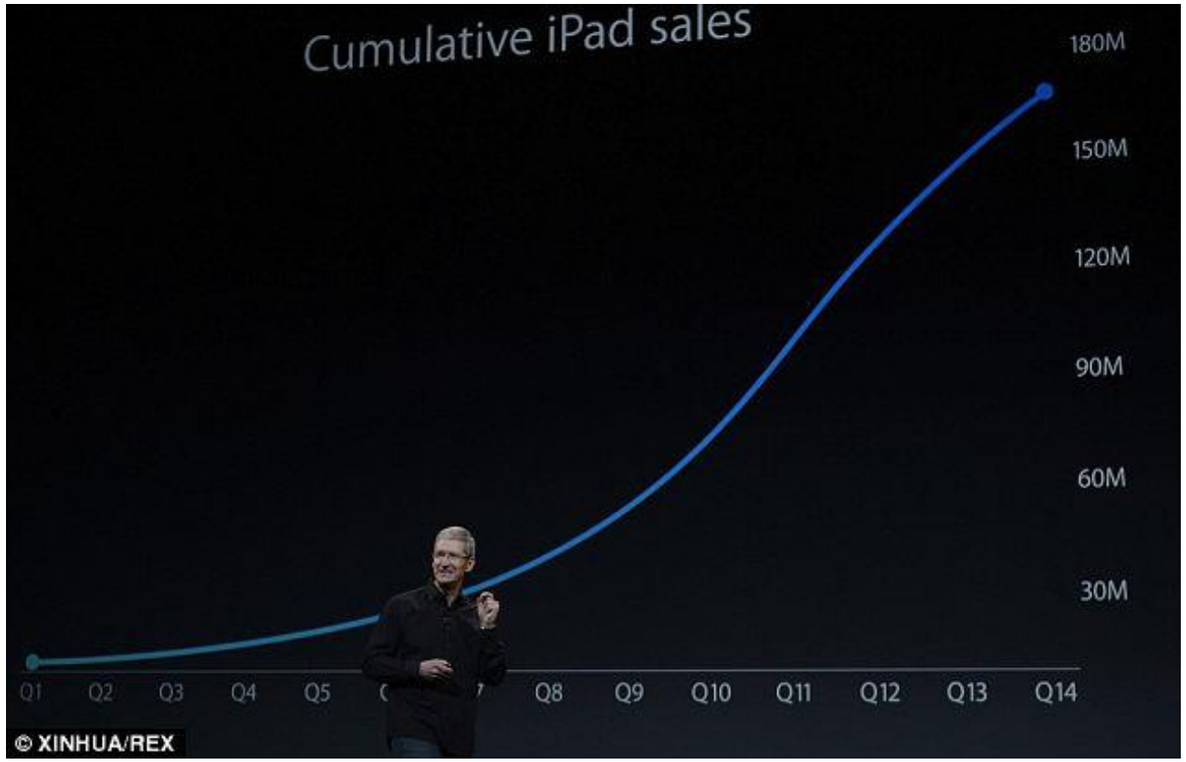
- What is wrong with this chart?



XY-Charts – Cumulative Charts

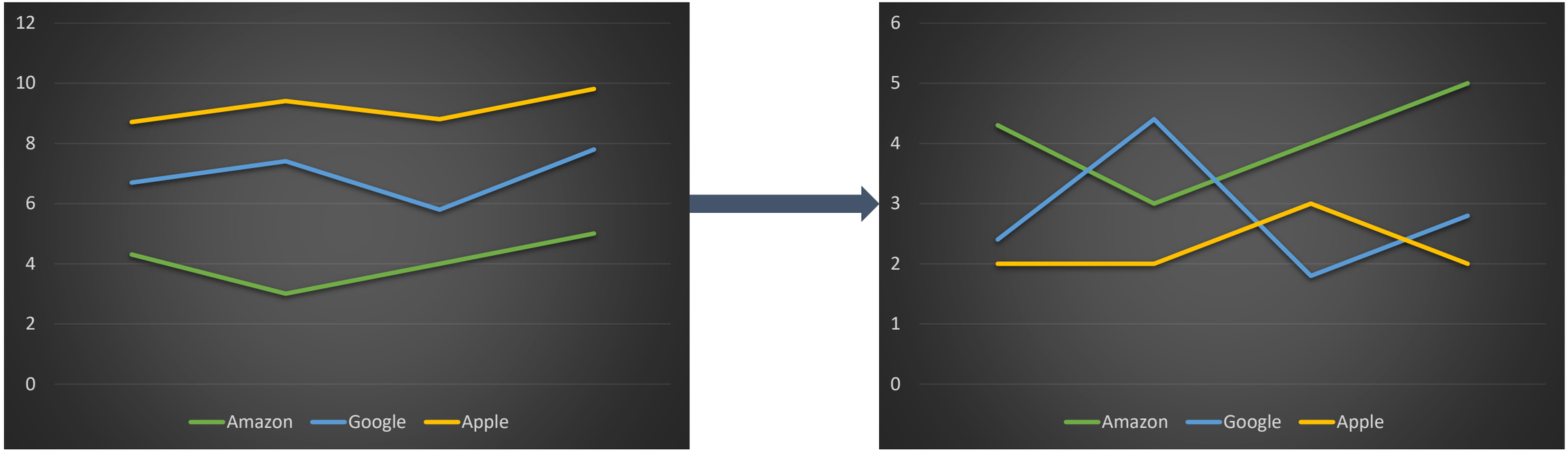
- Cumulative data always looks positive
  - Abuse of upward motion
  - Disguises variation of intermediate data
  - “These sales can only go up!”





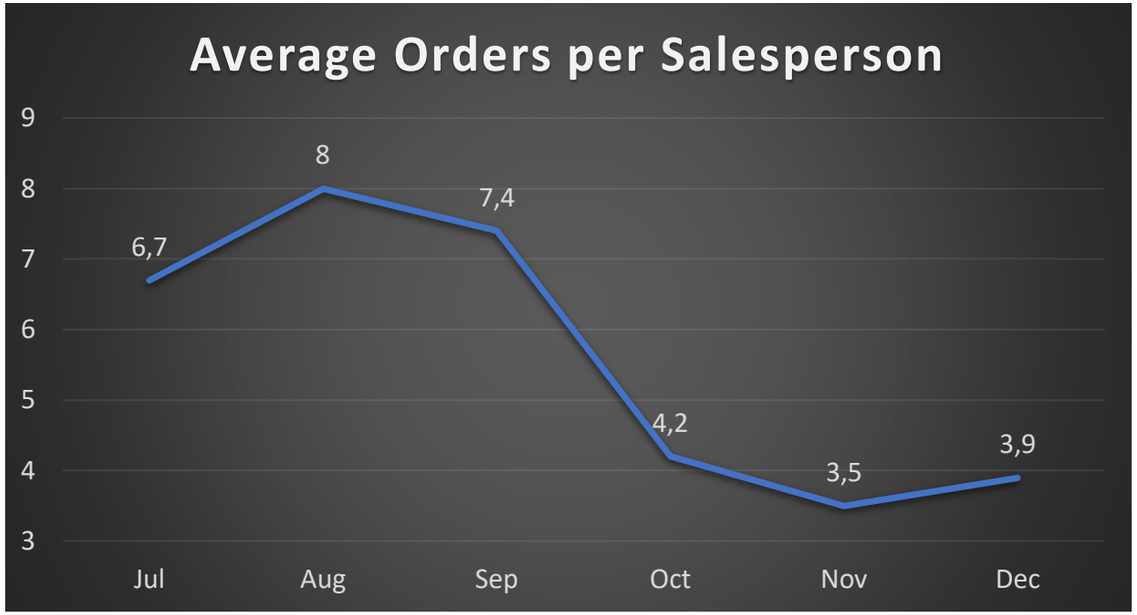
XY-Charts – Stacked Charts

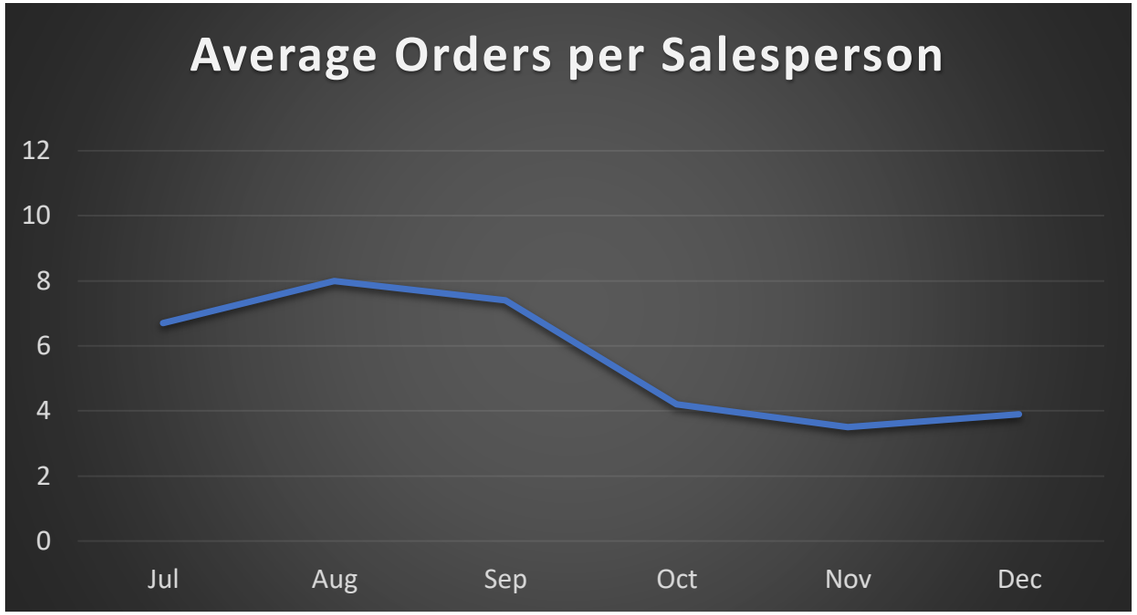
- Stacking line charts is another way to lie about your data
  - Notion of stacking visually not apparent
  - Base line assumed to be x-axis (y=0)



XY-Charts – Messing With Axes

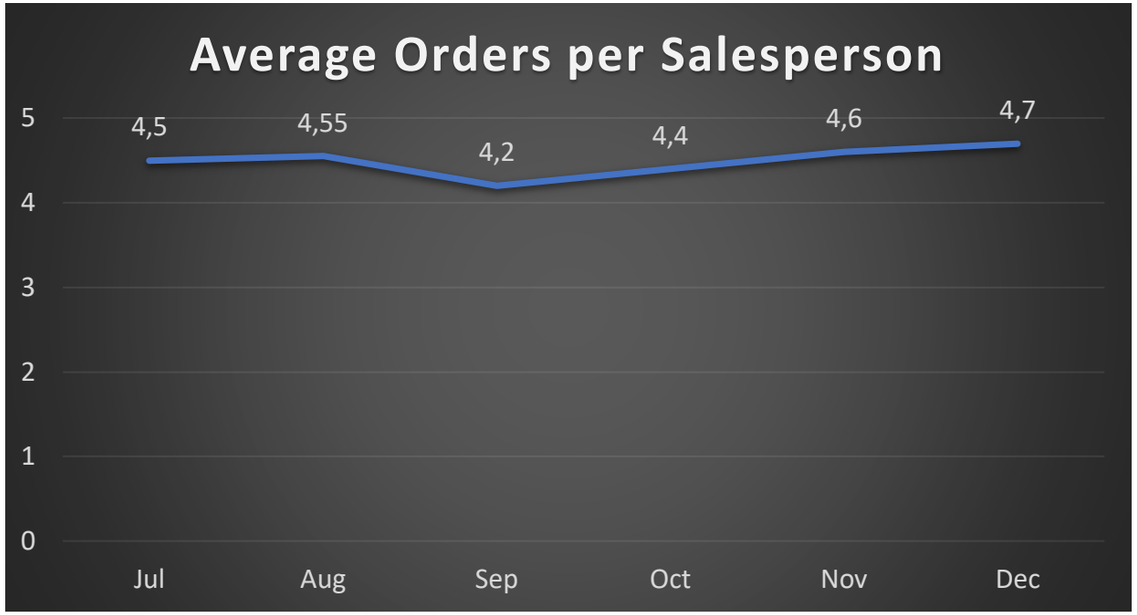
- The displayed range greatly influences appearance of plot
- Increase range to minimize fluctuations
  - Combine with omitting data labels

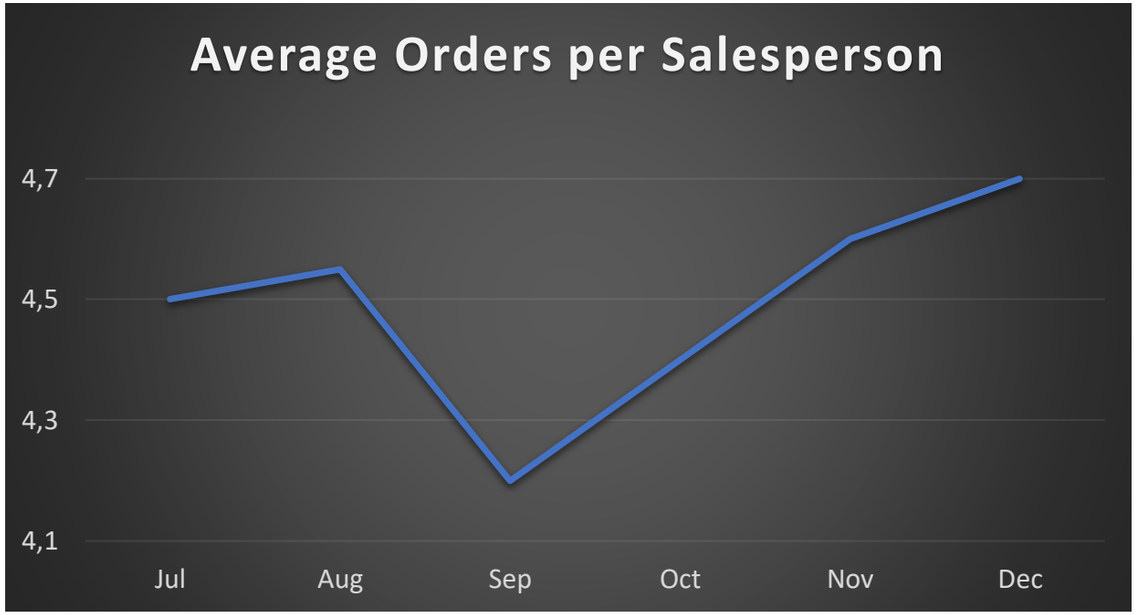




XY-Charts – Messing With Axes

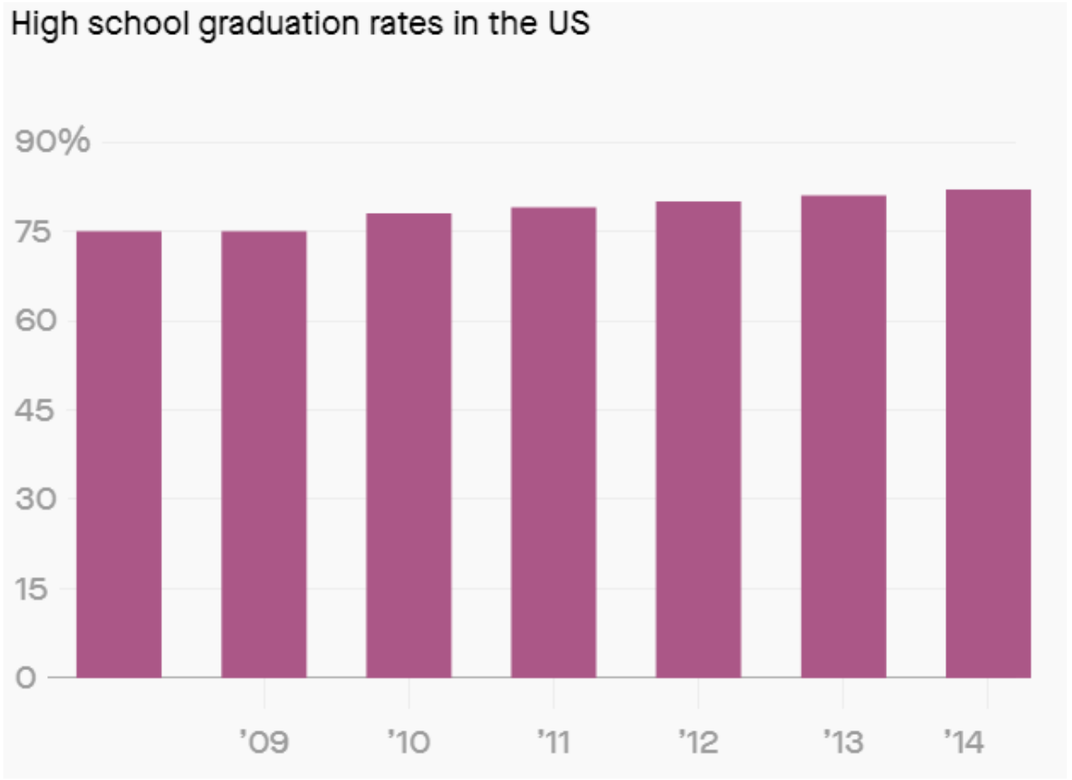
- Reversely, decreasing range maximizes fluctuations
- Widely used trick, even from “reputable” sources





XY-Charts – Messing With Axes

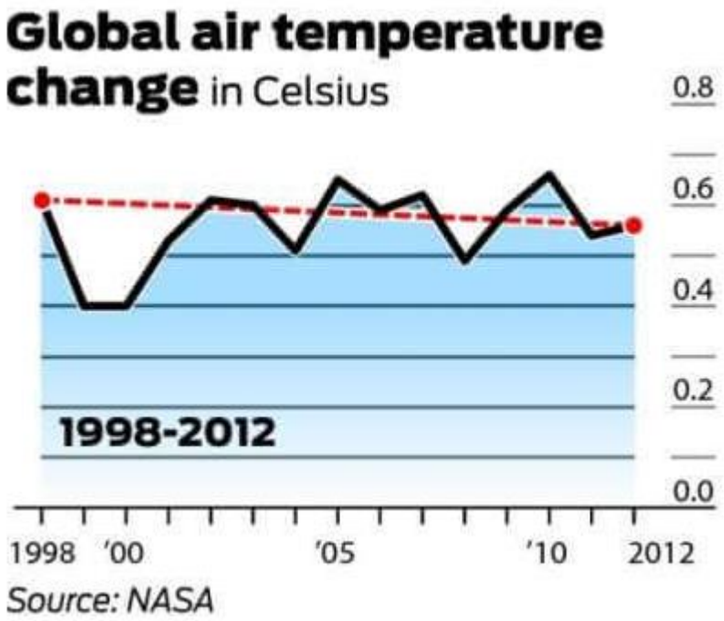
- Tweet from The White House under Obama presidency, 2015[1]

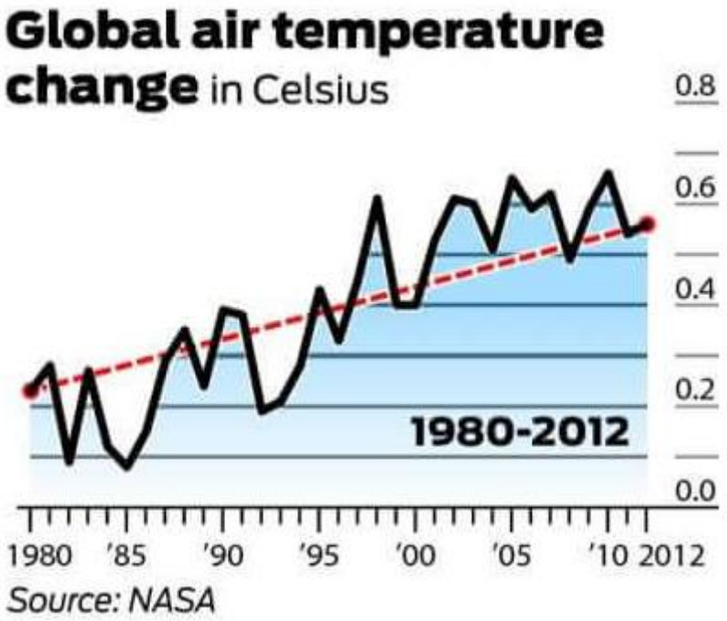


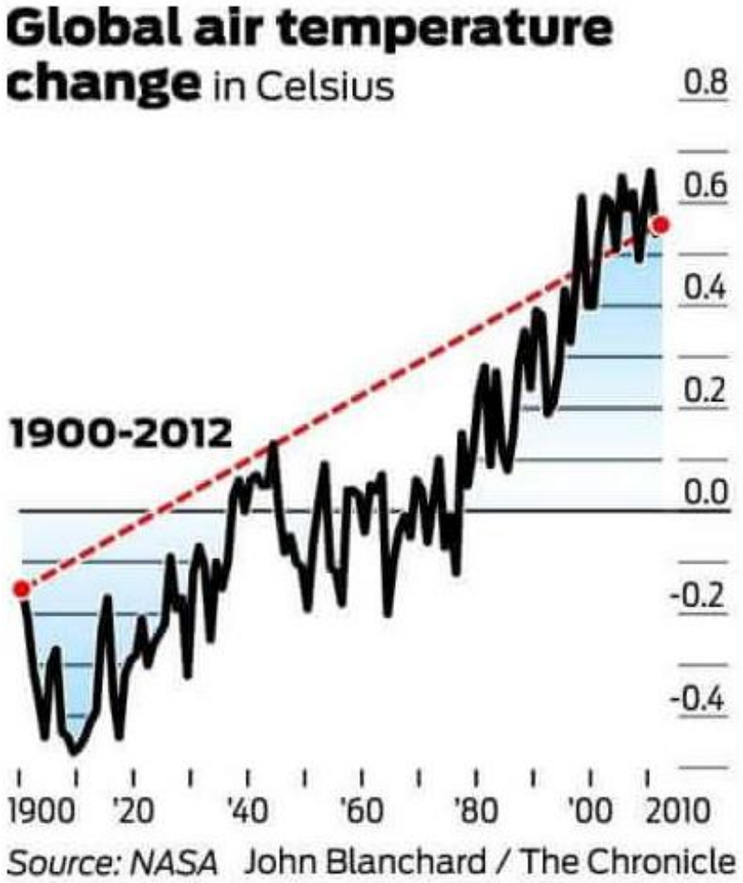
[1] https://twitter.com/ObamaWhiteHouse/status/677189256834609152

XY-Charts – Messing With Axes

- This trick also applies to the time axis

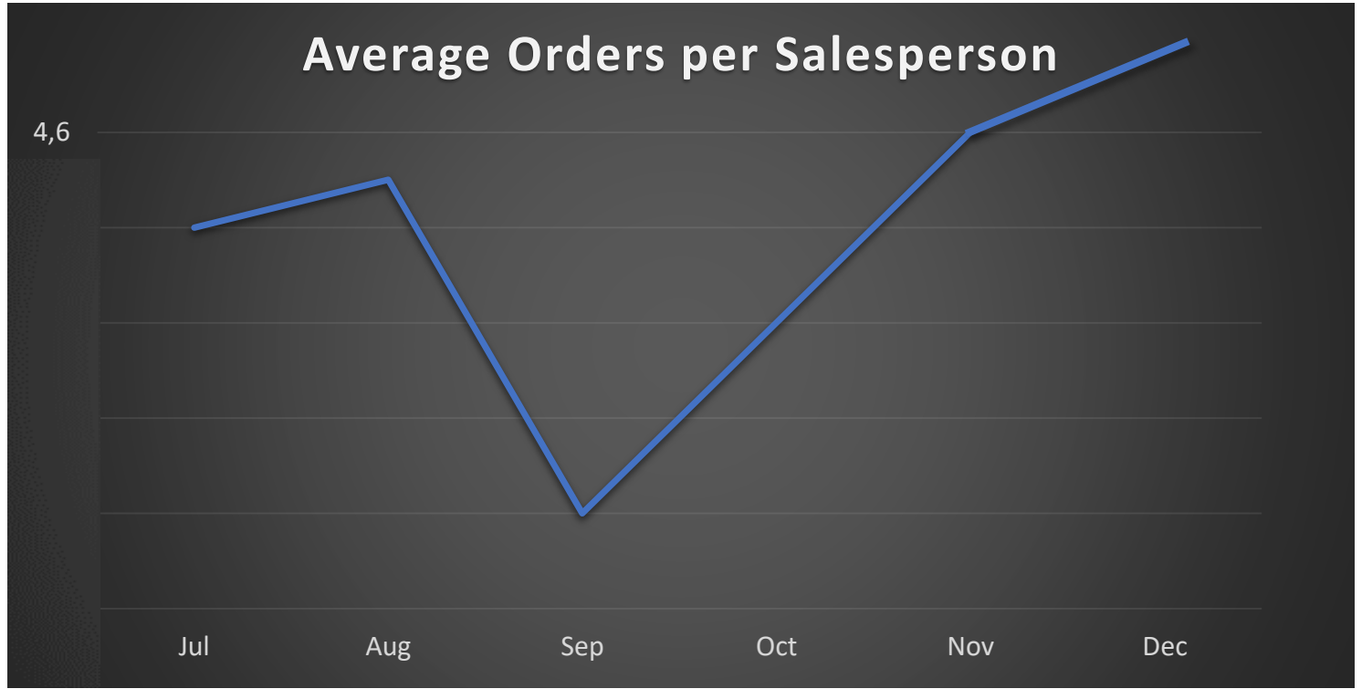






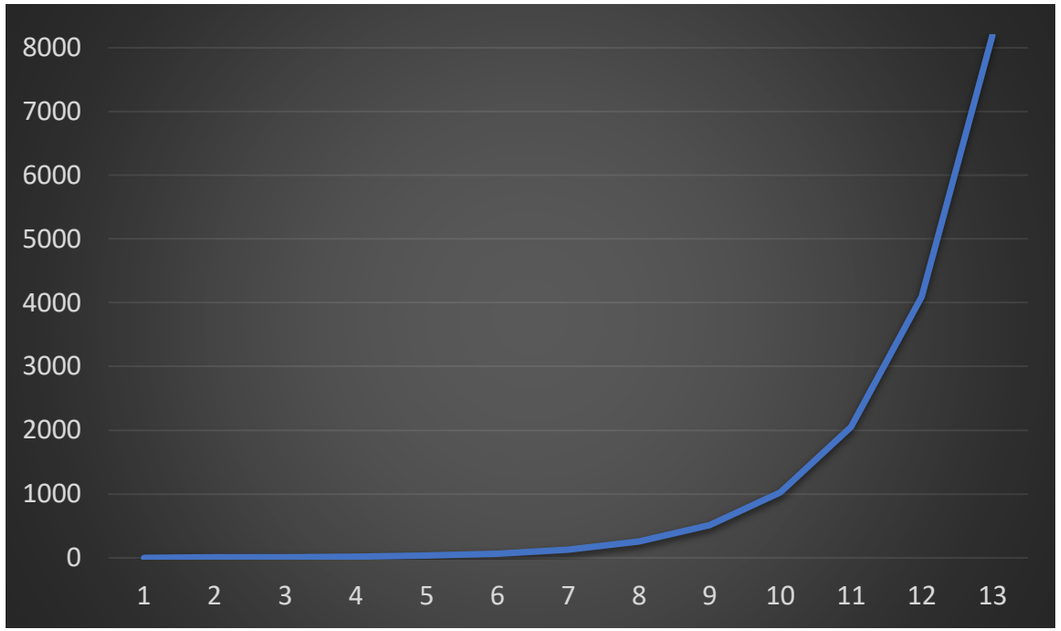
XY-Charts – Messing With Axes

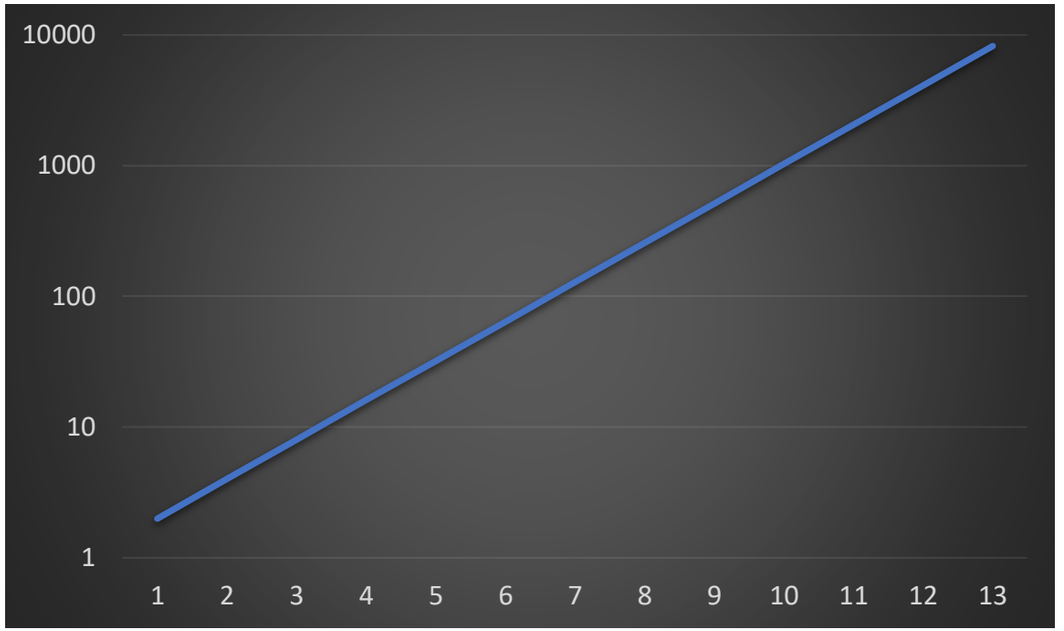
- Want to be even more evil?
  - Adjust range slightly below max data value
  - Omit every but maximum axis label



XY-Charts – Messing With Axes

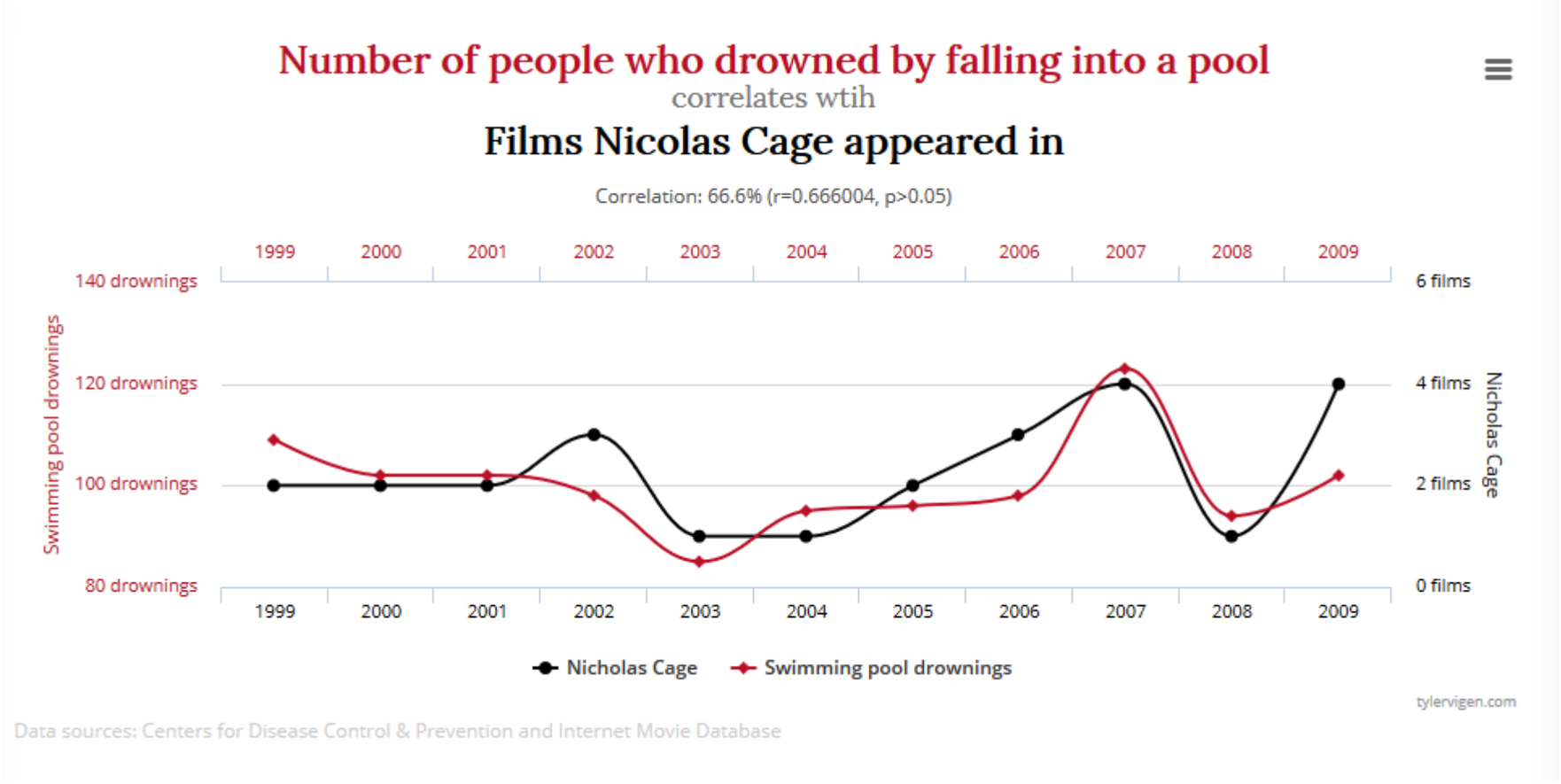
- Logarithmic axes counter impression of exponential growth
  - Possibly unconventional for audience
  - Misleading about growth rate
    - No “spiraling out of control”





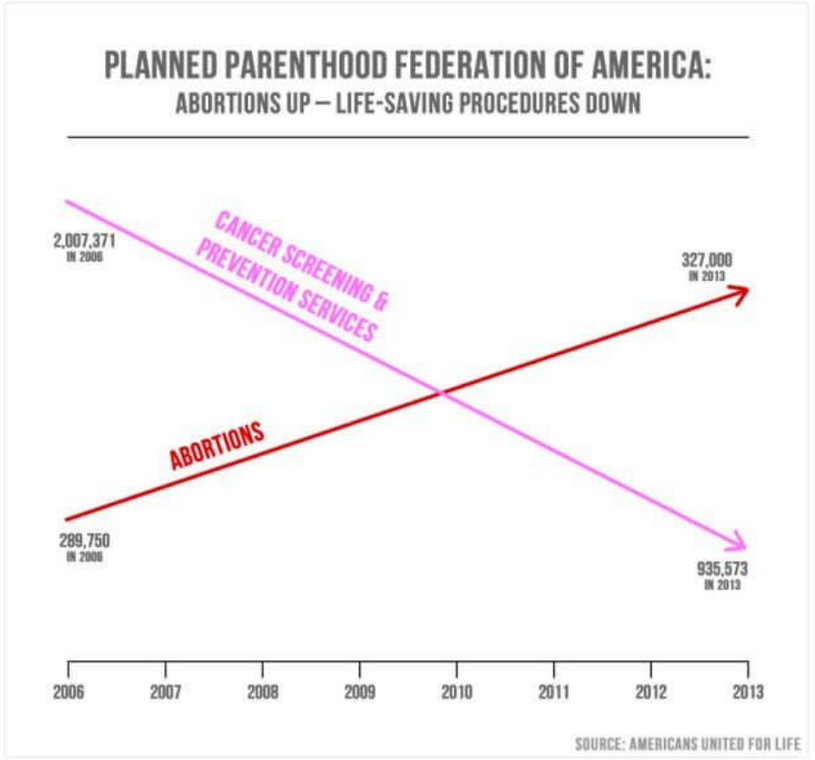
XY-Charts – Messing With Axes

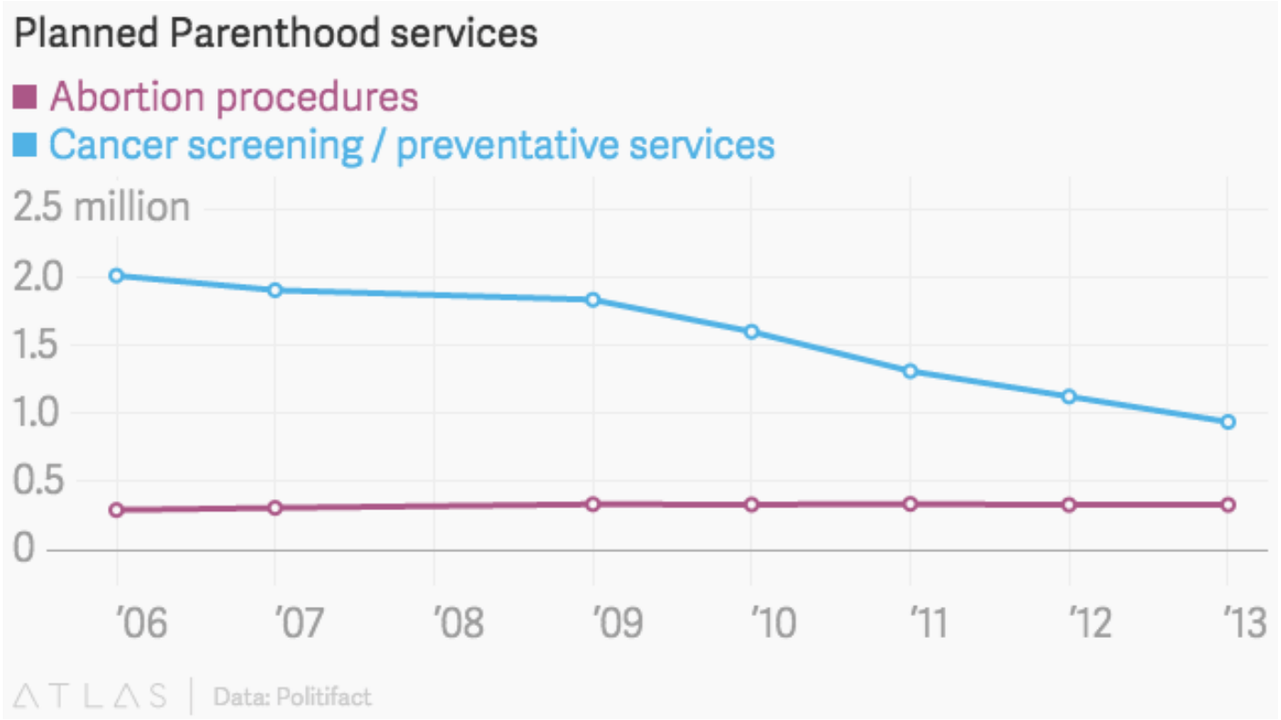
- Multiple axes can be abused to correlate unrelated things
  - Liars would use similar colors for both axes



XY-Charts – Messing With Axes

- With individual scaling of axes all kinds of trends can be implied
- Chart from republican discussion about rising abortion rates, 29.09.2015[1]

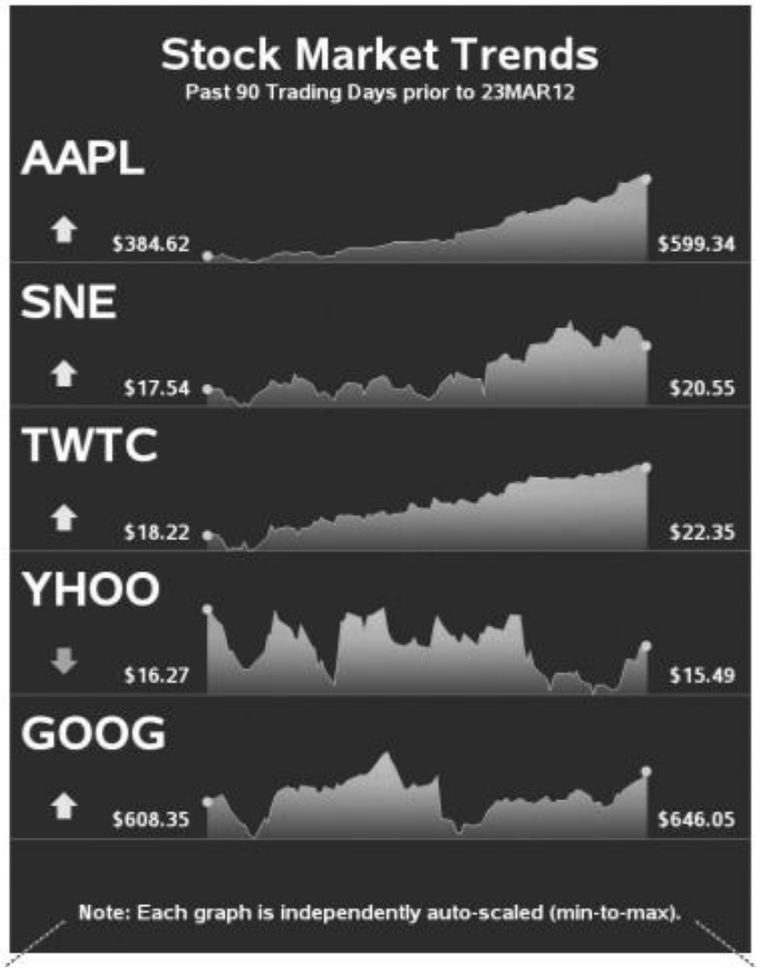




[1] https://www.politifact.com/truth-o-meter/statements/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/

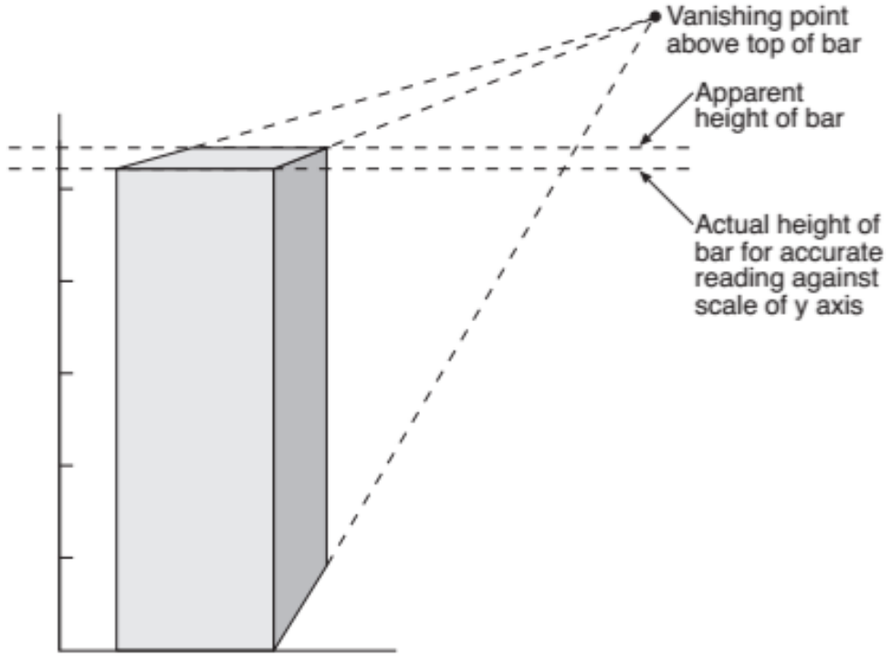
XY-Charts – Viewport

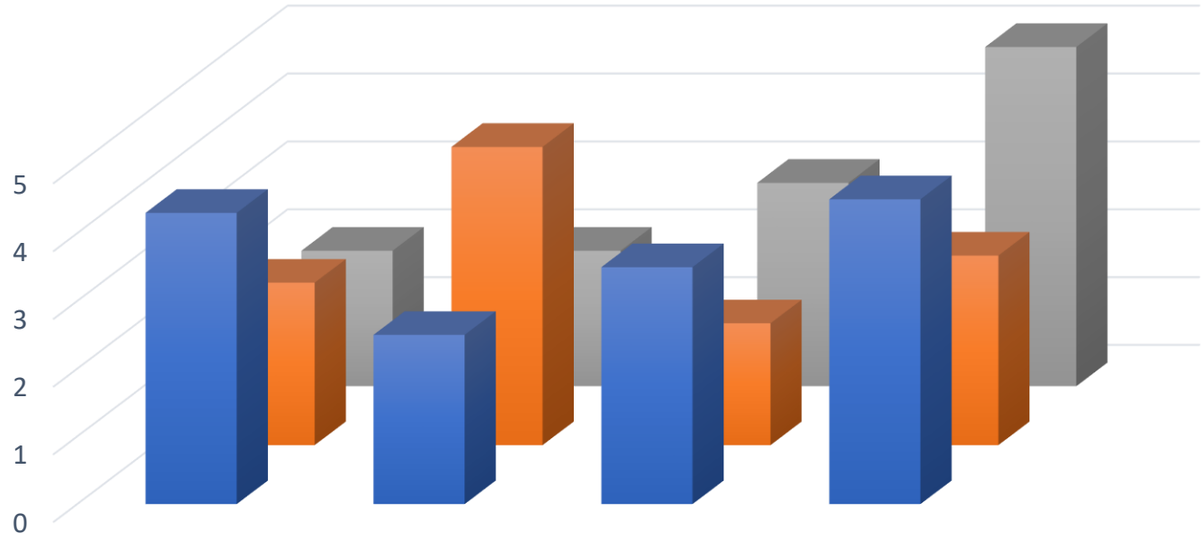
- Sometimes graphics themselves are not misleading, but displaying them side by side
  - Graphics scaled to fit certain viewport
  - Minor change identical to major change
  - Diminishes relation between plots



XY-Charts – 3D Bar Charts

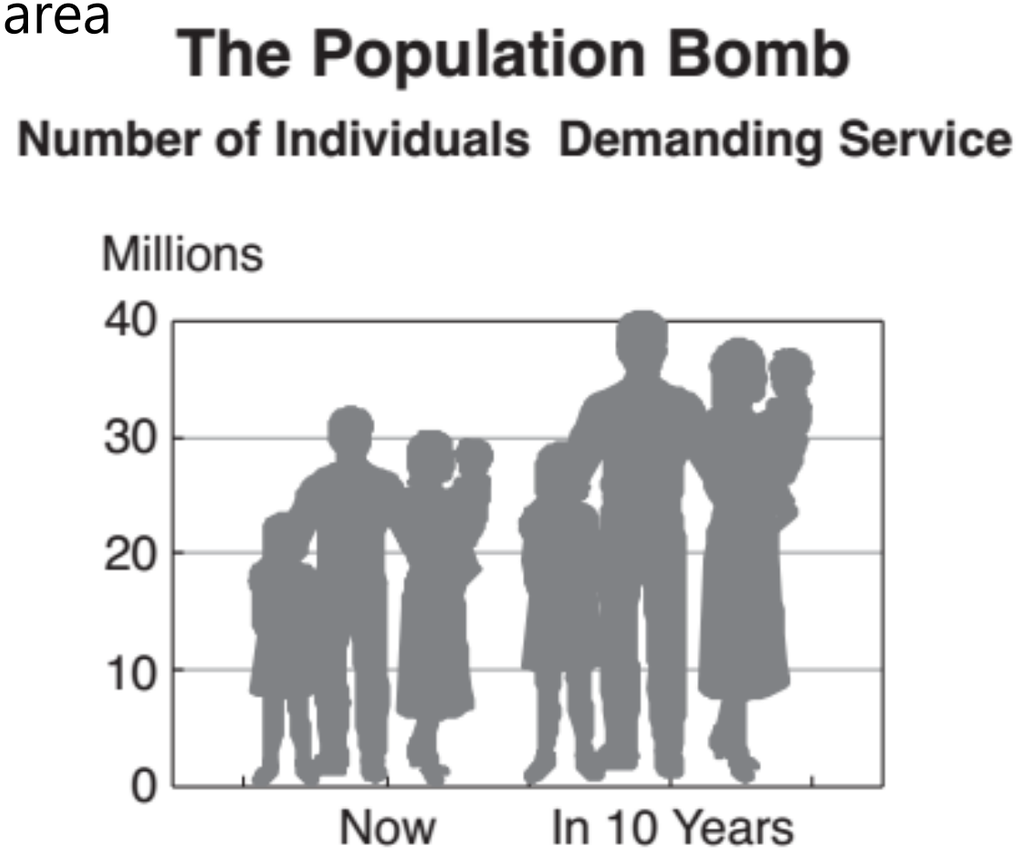
- Vanishing point above top distorts visual height of bar
  - Difficult to estimate actual value
  - Combine multiple 3D plots for maximum confusion





XY-Charts – Bar Charts

- Another common trick is to replace bars with symbols
  - Disproportionate gain of area: 20% height – 200% area
  - More on that in the upcoming presentation



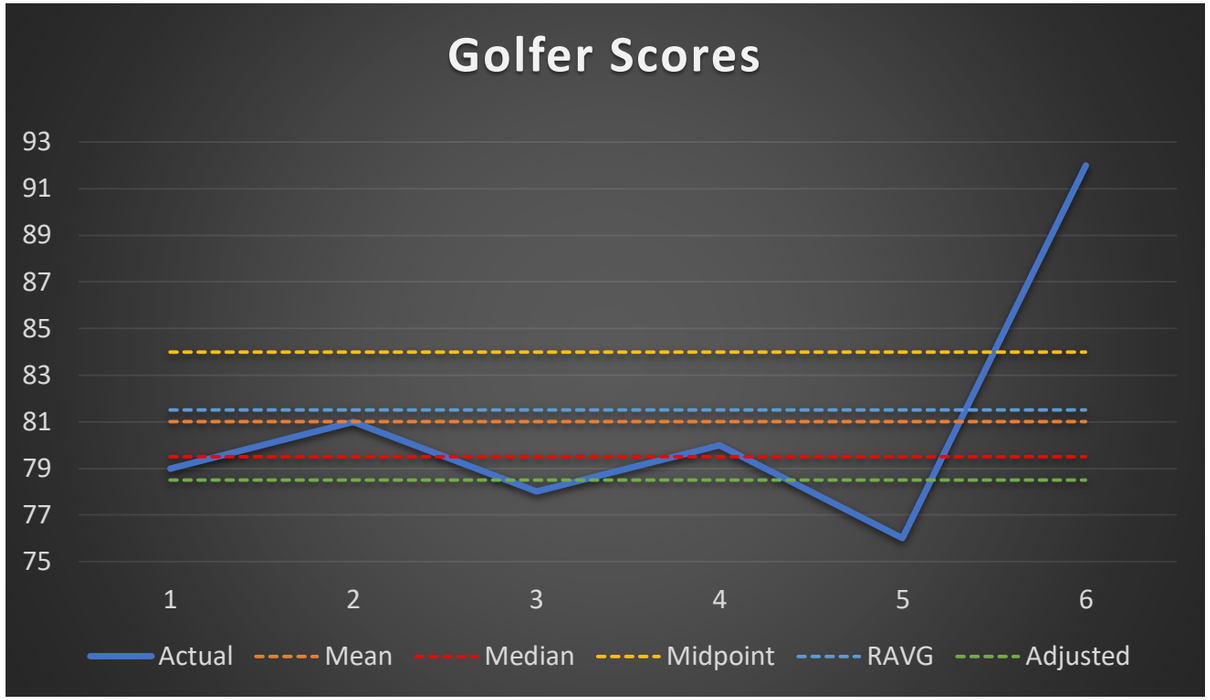
Trends

- Foreseeing the future has always been an inexact science
  - "I believe in the horse. The automobile is a temporary appearance!"
    Wilhelm the Second, 1916
  - "There is no reason anyone would want a computer in their home!"
    ~ Ken Olsen, 1977
  - "There is no chance that the iPhone is going to get any significant market share!"
    ~ Steve Ballmer, 2007

- Most predictions abstract reality and are based on subjective assumptions

Trends

- Many people would argue that the average is a good future prediction
  - But which one? Remember last weeks presentation
- Example: golfer scored 79, 81, 78, 80, 76, 92 in six games
  - Mean = $\frac{79+81+78+80+76+92}{6} = 81$
  - Median = $\frac{79+80}{2} = 79{,}5$
  - Midpoint = $\frac{76+92}{2} = 84$
  - 4 Game Running Average = 81,5
  - Cleared of “outlier” = 78,5

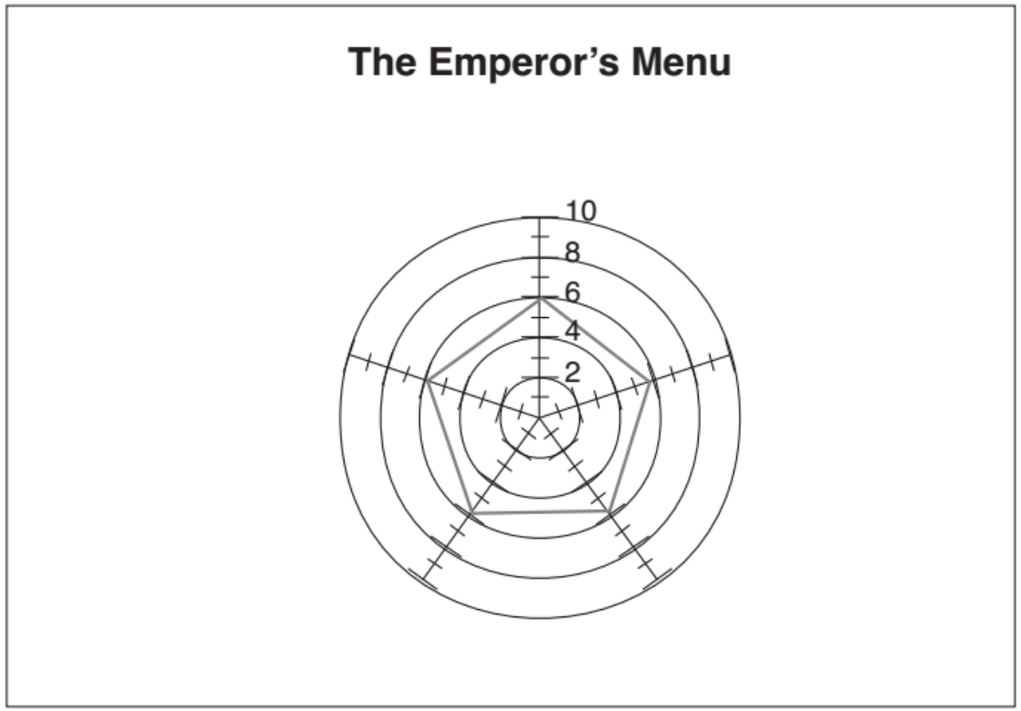


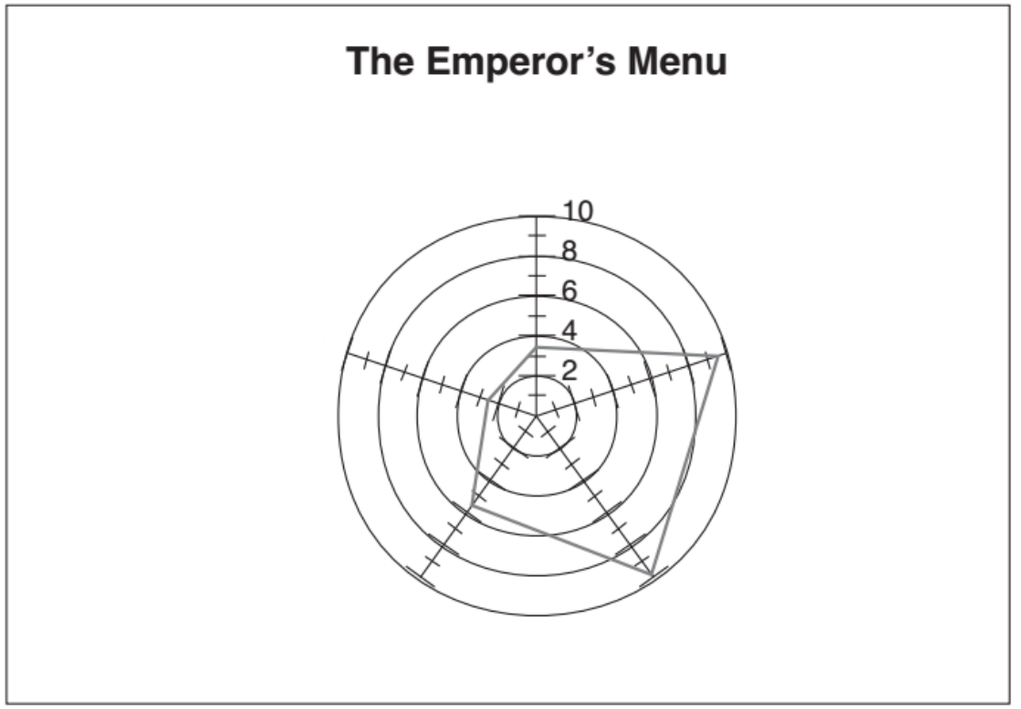
Trends – Regression

- Regression used to fit trend lines on data
  - Linear, exponential, logarithmic
- Impression that best fit estimates best
  - Oversimplification of reality
  - Omit outliers to reinforce hypothesis

Radar Charts

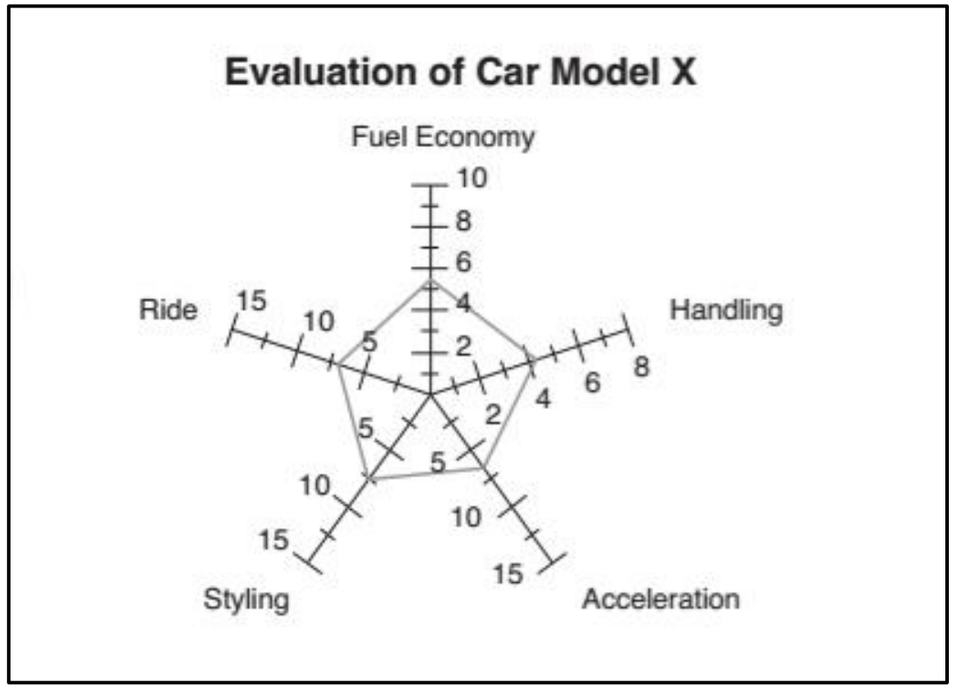
- Radar charts are an exotic example with different biases
  - Biased towards symmetry
  - Audience gives importance to area

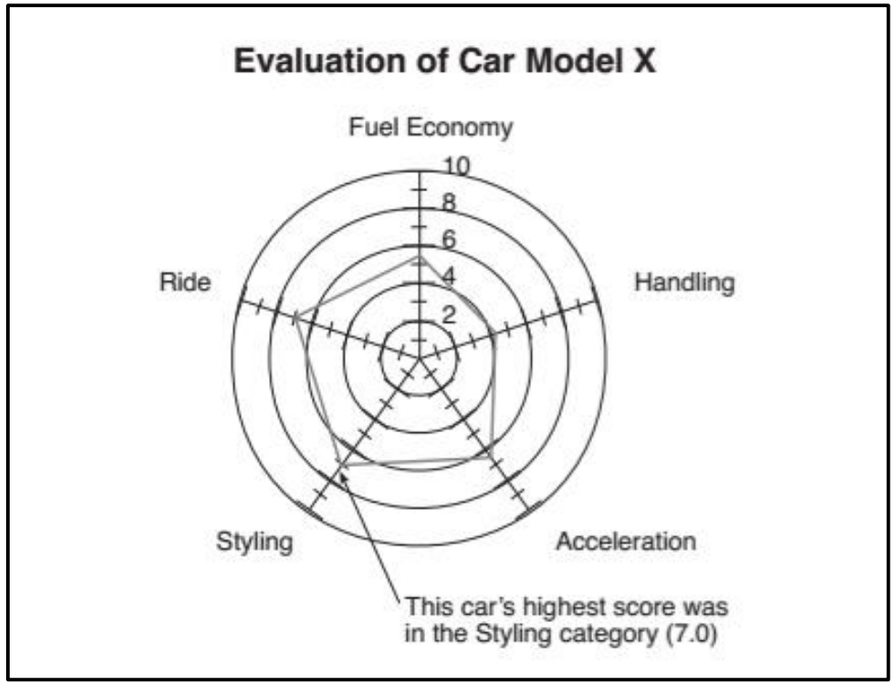




Radar Charts

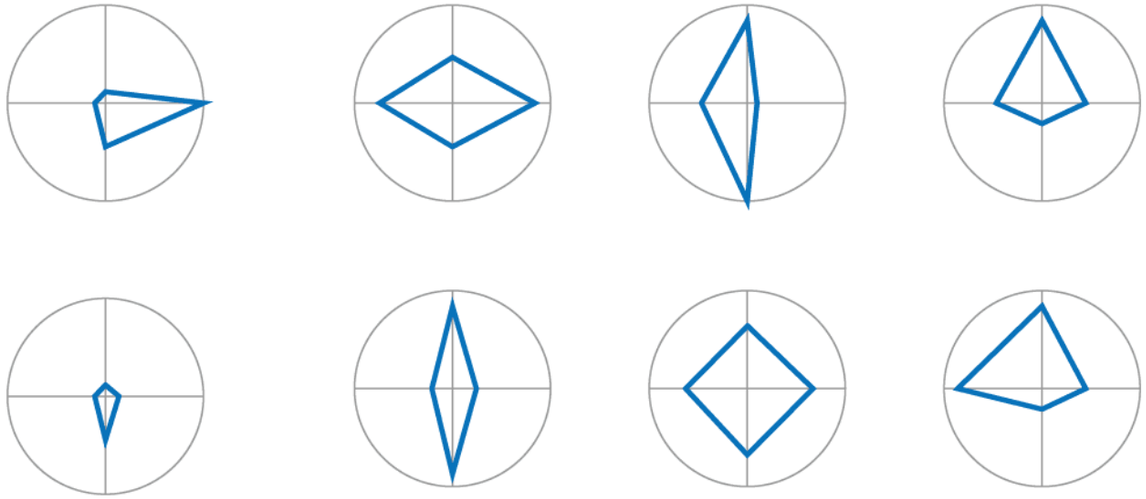
- Radar charts are only reliable with consistent scaling of axes
  - Liars can manipulate scaling to force symmetry
  - Justify modified axes with weighting of categories

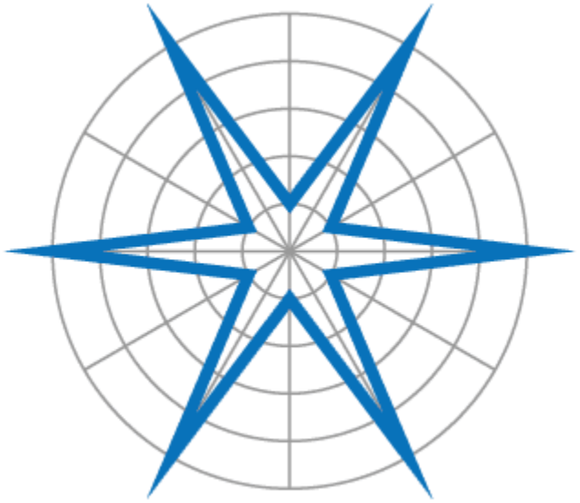


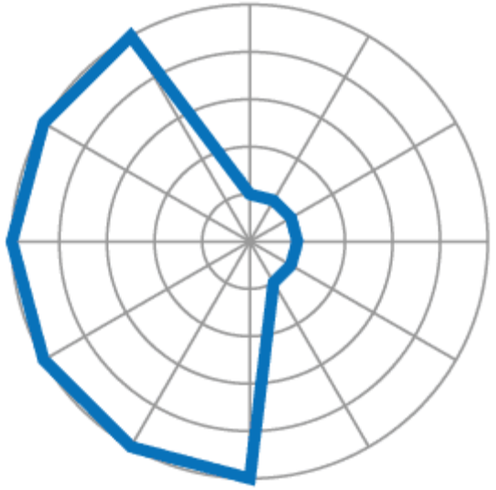


Radar Charts

- Positioning of categories also influences perception
  - Can be abused to imply various things
    - Well rounded or one sided?
  - Misleading through cultural bias of direction







Thank you for your attention!