

ISA 401: Business Intelligence & Data Visualization

18: Charts Used for Comparisons, Relationships, Distributions and Correlations

Fadel M. Megahed, PhD

Endres Associate Professor
Farmer School of Business
Miami University

 @FadelMegahed

 fmegahed

 fmegahed@miamioh.edu

 Automated Scheduler for Office Hours

Fall 2023

Learning Objectives for Today's Class

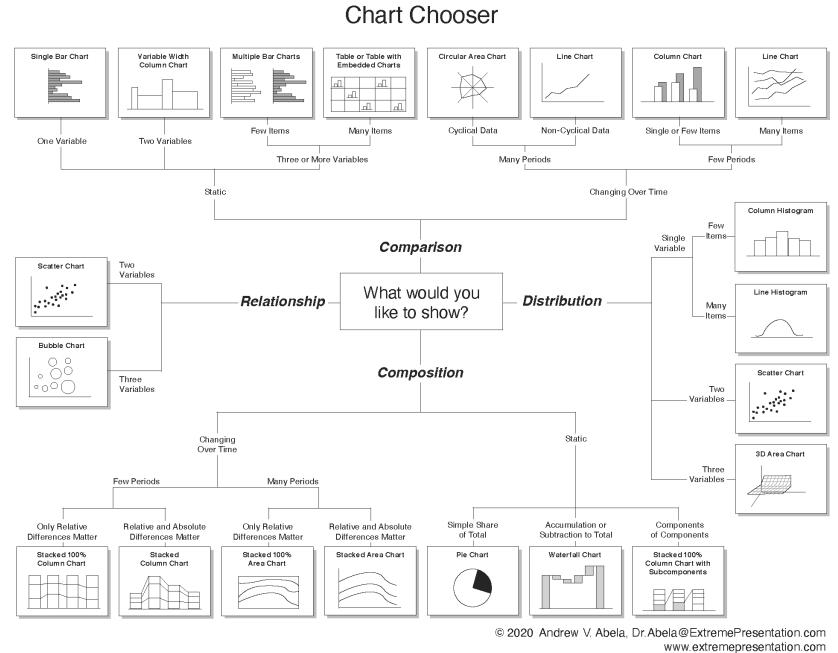
- Identify strengths & weaknesses of basic charts
- Use appropriate charts based on objective
- Avoid using pie charts (never use pie charts)
- Avoid 3D graphs (unless VR changes their utility)

A Catalog of Commonly Used Graph Types

PolicyViz Data Visualization Catalog						
Graph Type	Organization	Author	Small Multiples	Year	URL	Image
Graph Type ▾	Description /Notes	Organization	Author	Small Multiples	Year	URL

Looker Studio

Chart Suggestions



Charts Used for Comparing Data

(Unit of Analysis is Based on a Nominal Categorical Variable)

A Literal Bar Chart

Activity Your Solution



Answer the following questions:

- (1) How many variables do we have in this graph?
- (2) How many observations?
- (3) Please discuss the type of variables in the graph? (i.e. nominal, ordinal, etc.)
- (4) How is the data encoded in the graph?
- (5) Any other comments/observations?

A Literal Bar Chart

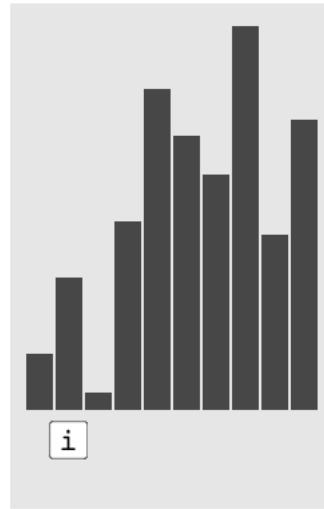
Activity

Your Solution

- Q1:
- Q2:
- Q3:
- Q4:
- Q5:

Using a Bar Chart to Visualize R Code

```
insert_sort <- function(x) {
  i <- 2
  while(i <= length(x)) {
    j <- i
    while(j > 1 && x[j - 1] > x[j]) {
      j <- j - 1
      x[j + 0:1] <- x[j + 1:0]
    }
    i <- i + 1
  }
  x
}
```

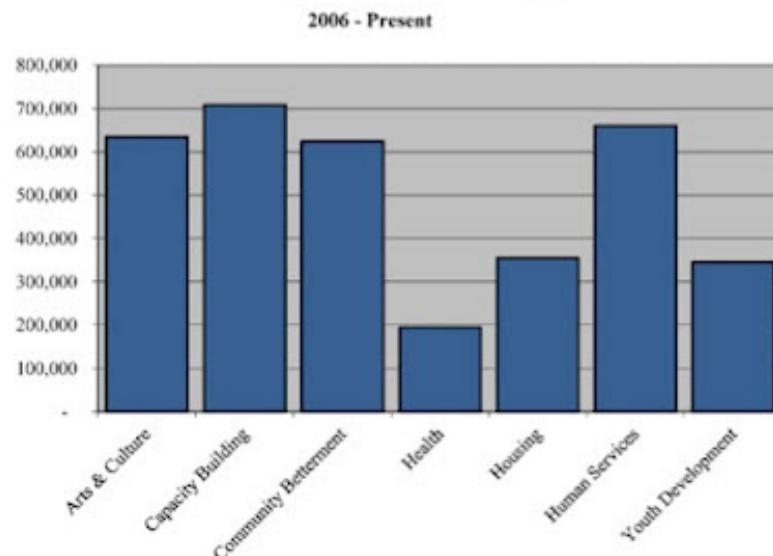


Non-graded activity: Two Bar Charts

Activity Your Solution

Over the next five minutes, identify **3-4 differences that make the graph on the right better**, and suggest **how you can further improve the graph on the right**

Investment by area of impact



We invest primarily in four areas

Since we began investing in 2006, **four areas have received more than \$600K each, accounting for 75% of total grantmaking activity**

Investment by Area of Impact
2006 - Present



04:00

Non-graded activity: Two Bar Charts

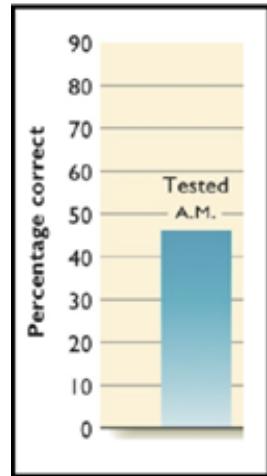
Activity

Your Solution

Insert your differences and suggestions for improvement below.

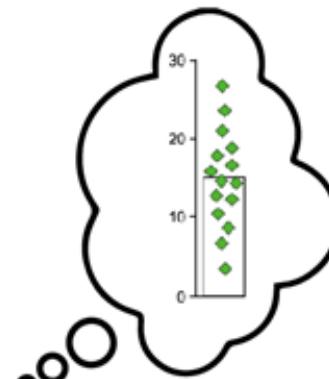
Issues with the Interpretation of Bar Charts

a. Stimulus

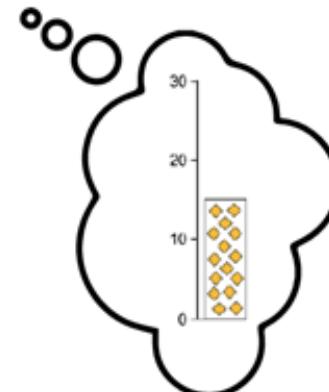


*example bar from
bar graph*

b. Interpretation



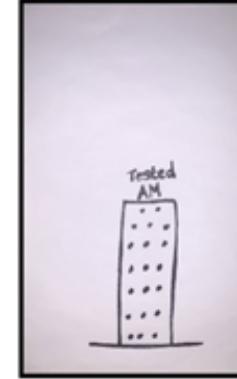
example interpretations



c. Response



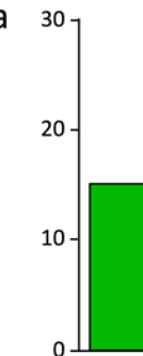
example drawings



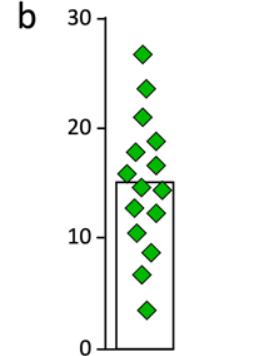
The Draw Datapoints on Graph (DDoG) measure maintains the graph as a consistent reference frame across its three stages.

Issues with the Interpretation of Bar Charts

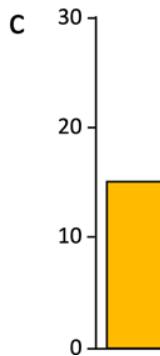
Mean bar graph



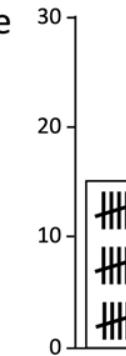
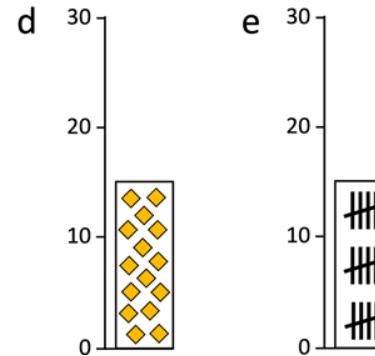
Mean bar graph data



Count bar graph



Count bar graph data



Data distribution differs categorically between mean and count graphs. (a) Mean bar graphs and (c) count bar graphs do not differ in basic appearance, but they do depict categorically different data distributions.

Key Takeaway 1

The typically used **bar** chart should not be to depict means of categorical variables.

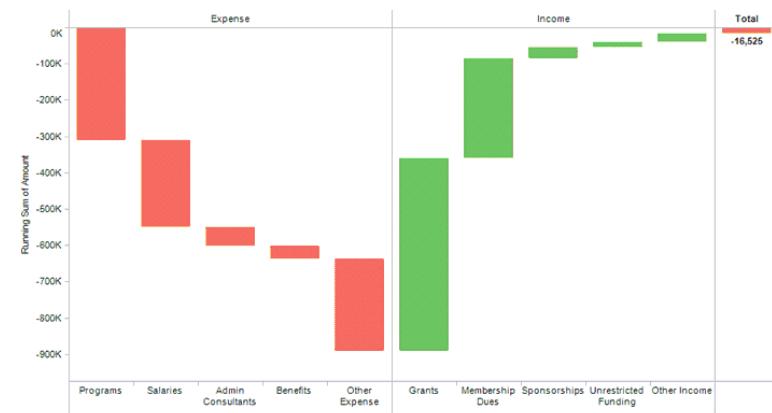
Waterfall Charts

Activity Your Solution

What are the advantages and disadvantages of these two charts? They are using the same exact data. Please try to list 2-4 in each category for each chart.



Whereas **income is primarily from a single channel** (Grants, \$531K in 2010, or 61% of total income), **2010 expenses were spread roughly evenly across programs, salaries, and other expenses**. This means [...]



Waterfall Charts

Activity

Your Solution

Insert your advantages and disadvantages below

3D Bar Charts are Awful

 Adam Grant 
@AdamMGrant · [Follow](#) 

When our status is secure, we don't emphasize it. When it's ambiguous, we do.

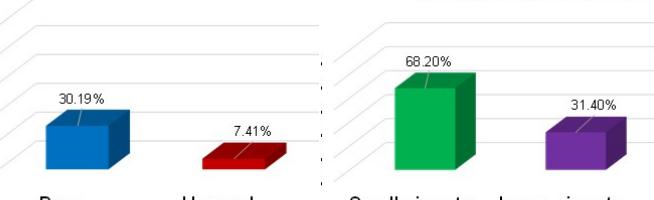
Penn students are more likely than Harvard students to mention that they go to an Ivy League school.

Smaller airports are more likely to refer to themselves as international.

journals.sagepub.com/doi/full/10.11...

Say They Attend an Ivy League Sc

Call Themselves an International Airport

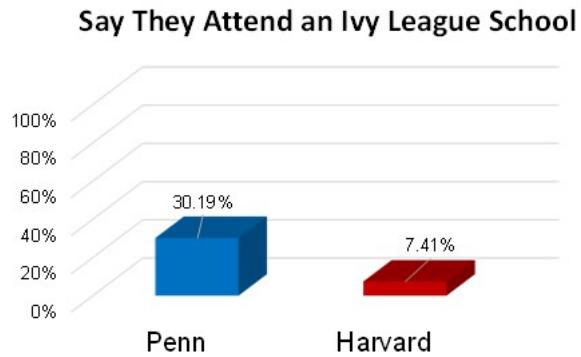


Category	Penn	Harvard	Small airports	Large airports
Say They Attend an Ivy League Sc	30.19%	7.41%		
Call Themselves an International Airport			68.20%	31.40%

8:30 AM · Dec 18, 2020 

 982   [Read 39 replies](#)

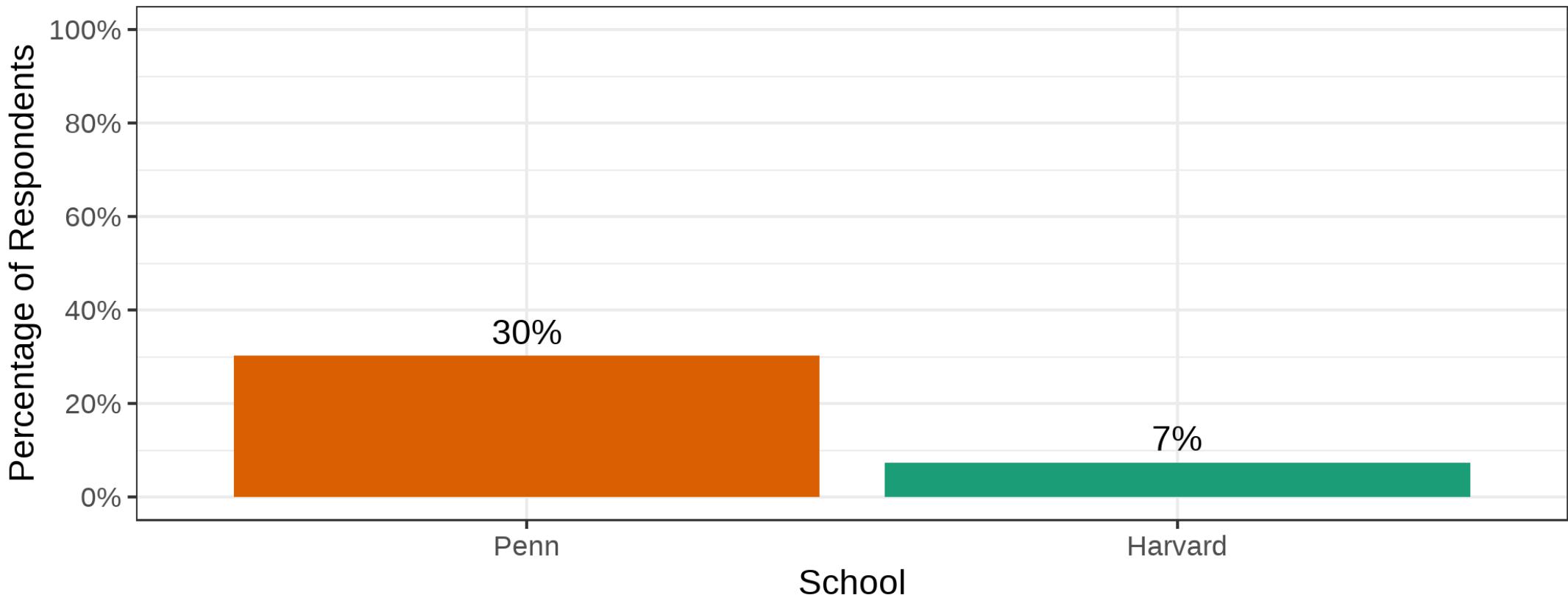
3D Charts are Awful



Adam Grant's Plot of the Penn and Harvard Bar Chart

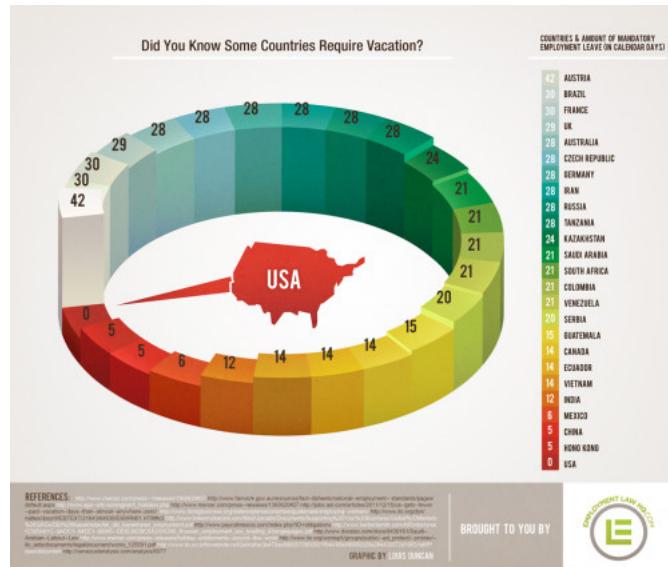
3D Bar Charts are Awful

Say They Attend an Ivy League School



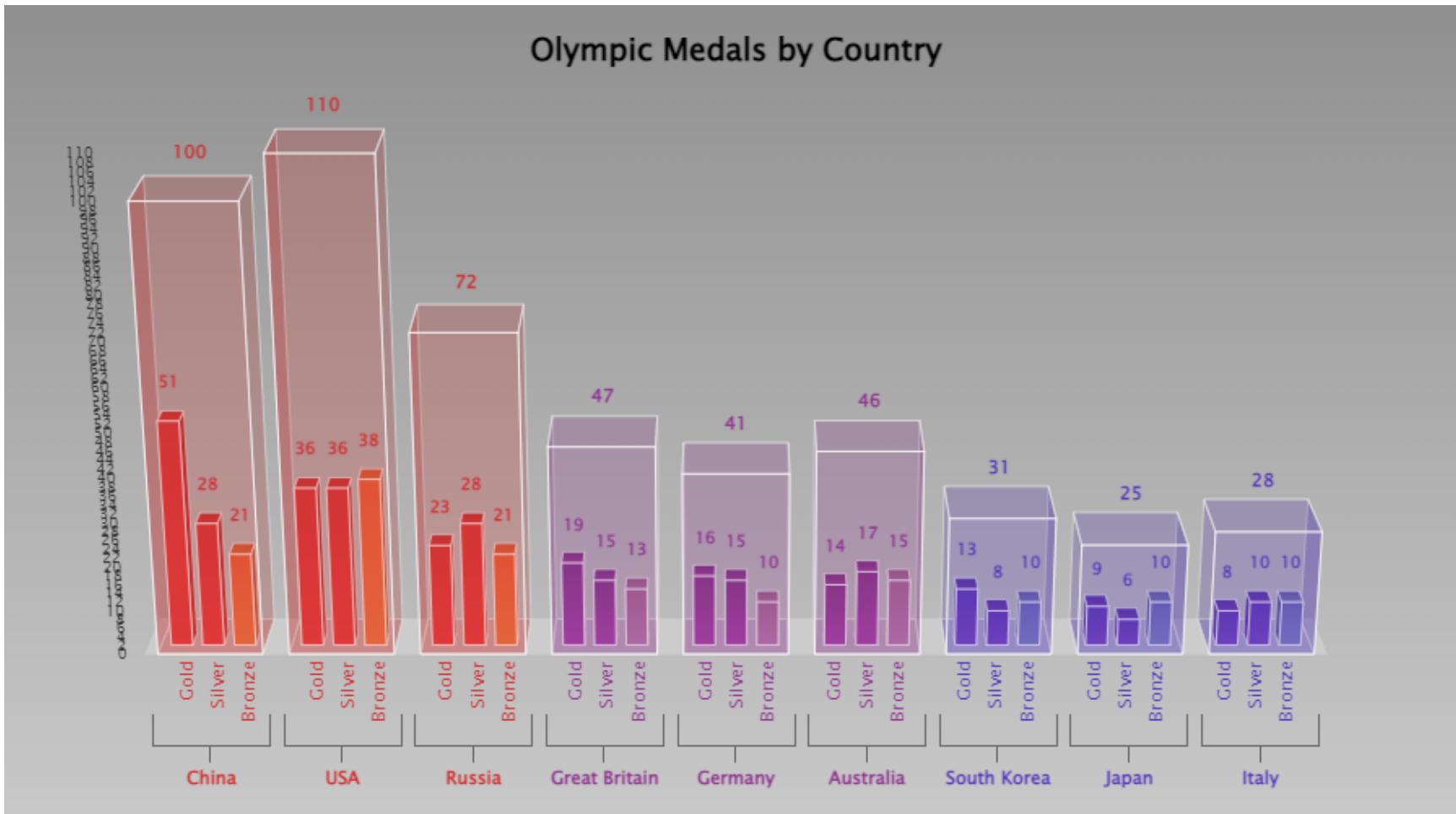
A remake of the plot with colorblind-friendly colors and a 2D bar representation to avoid distorting the data

3D Bar Charts are Awful: Including This

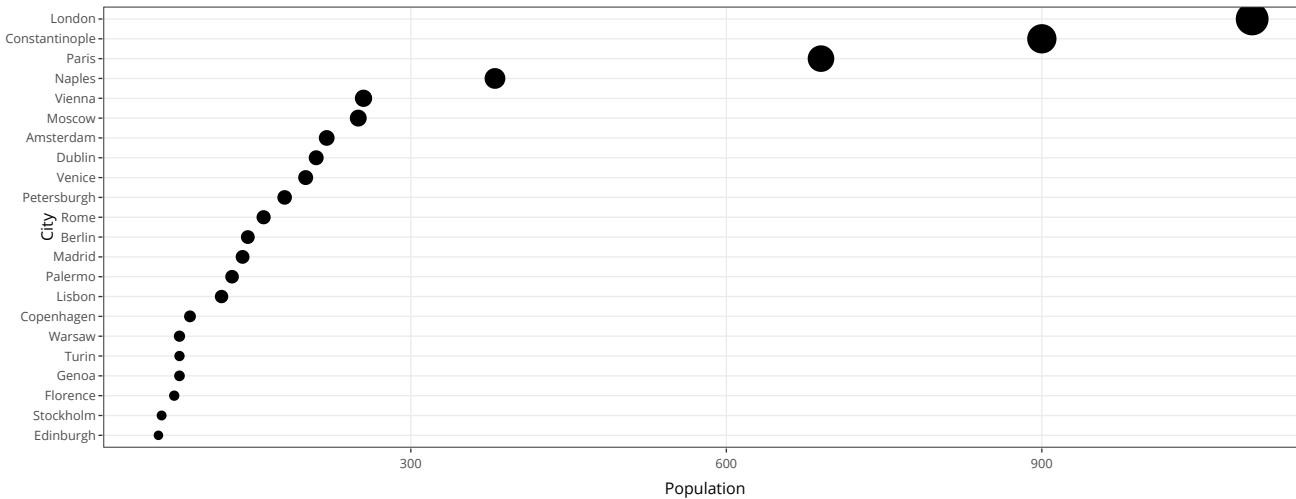


Source: See the [visually](#) post.

3D Charts are Awful: Even This



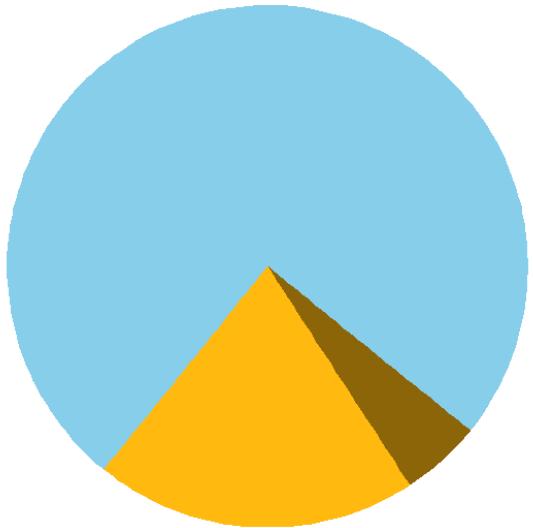
Dot Charts: Recall the Playfair Example



Proportions

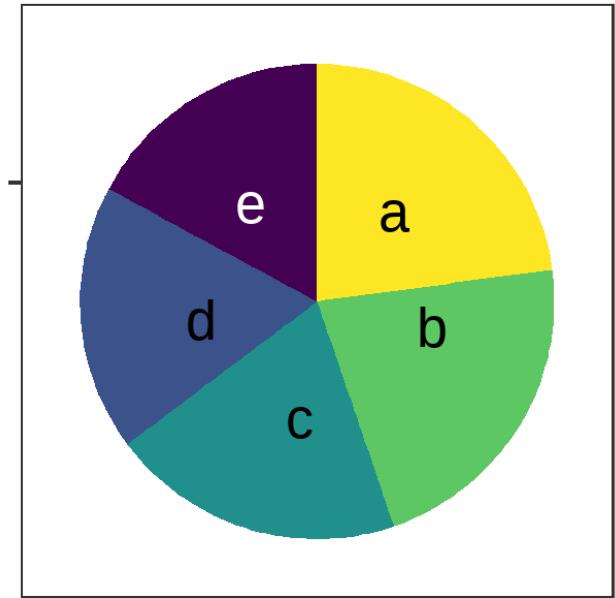
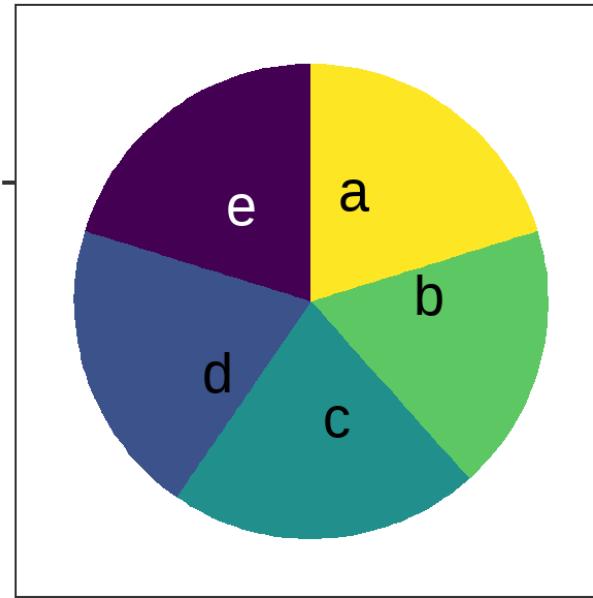
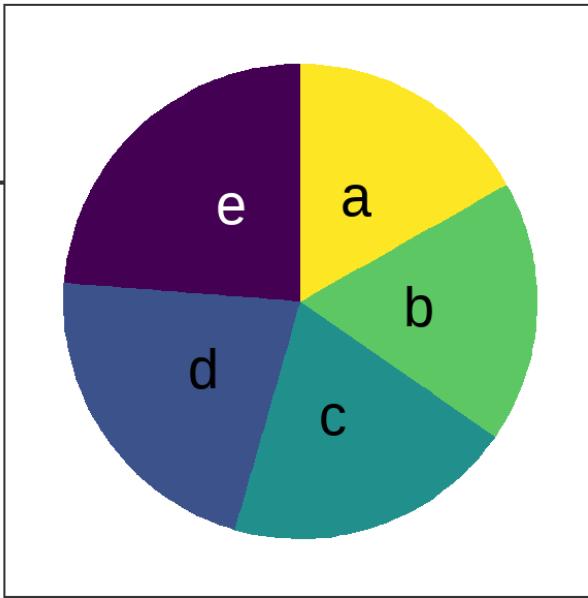
My Favorite Pie Chart

The Egyptian Pie Chart

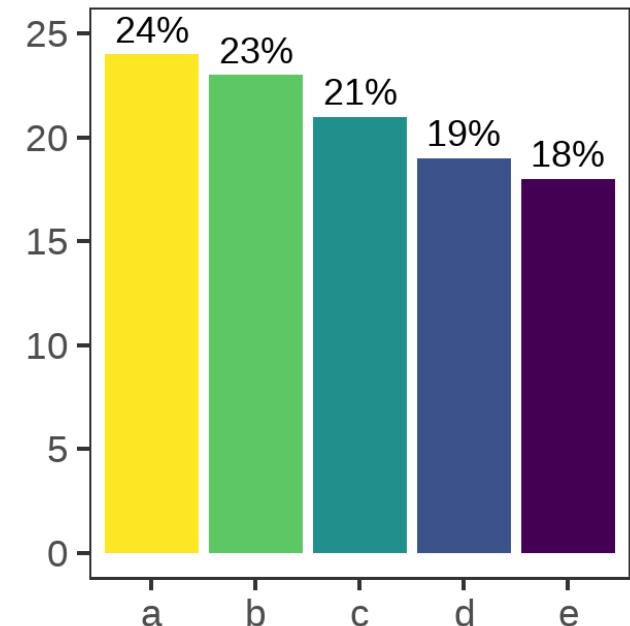
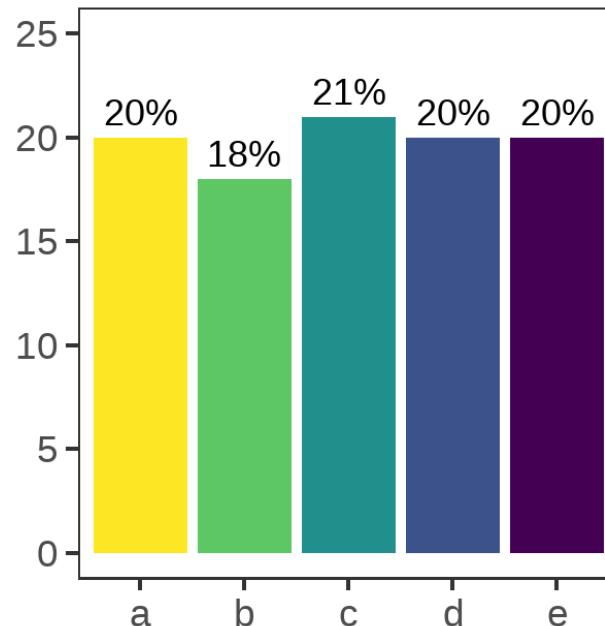
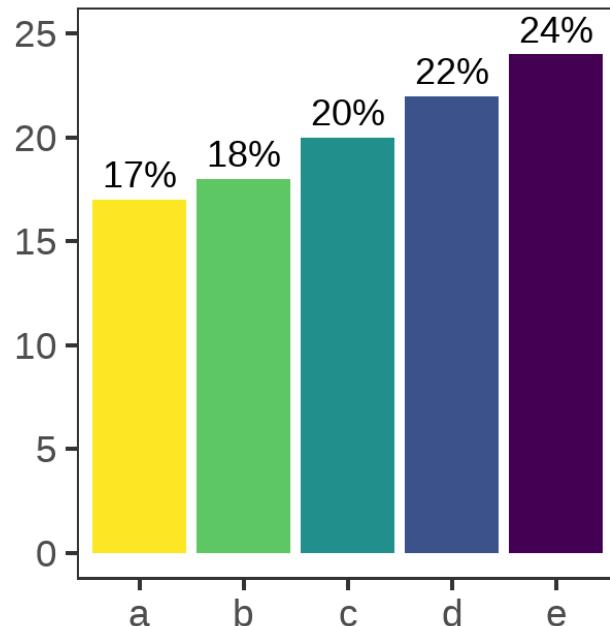


- Sky
- Sunny Side of the Pyramid
- Shady Side of the Pyramid

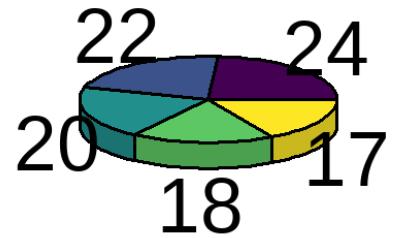
Pie Charts are Awful By Design



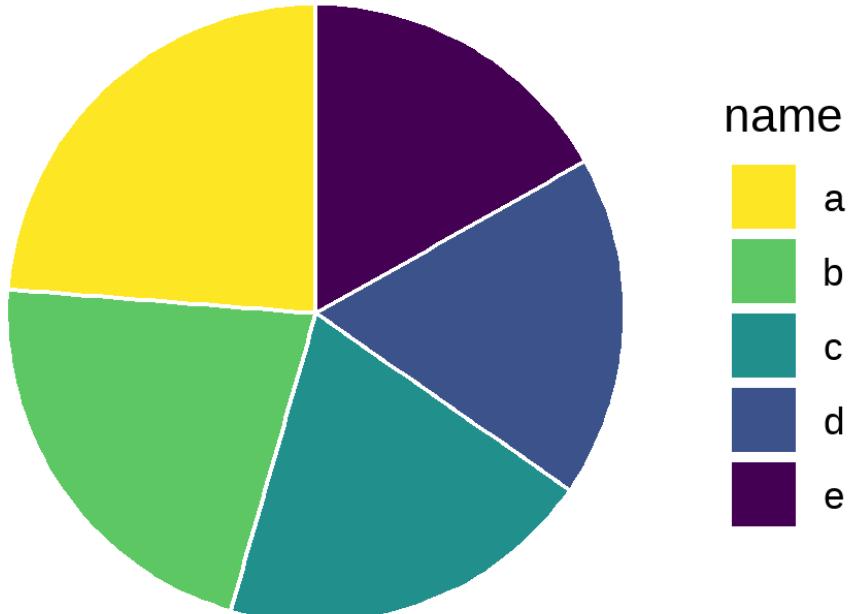
Pie Charts are Awful By Design



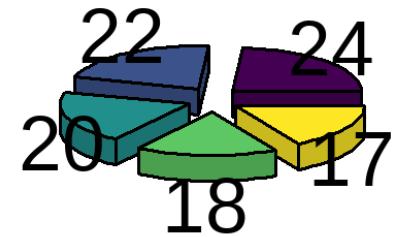
And often made even worse: 3D



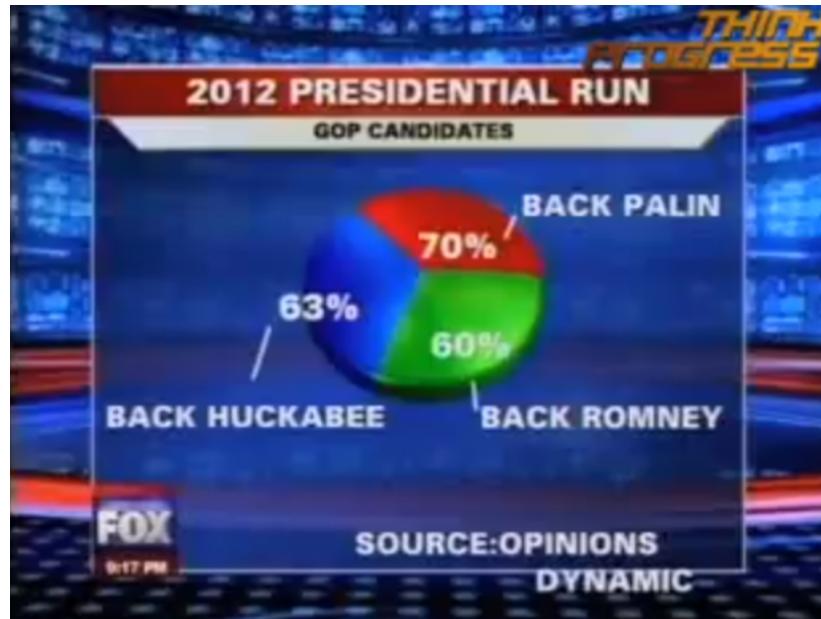
And often made even worse: Side Legend



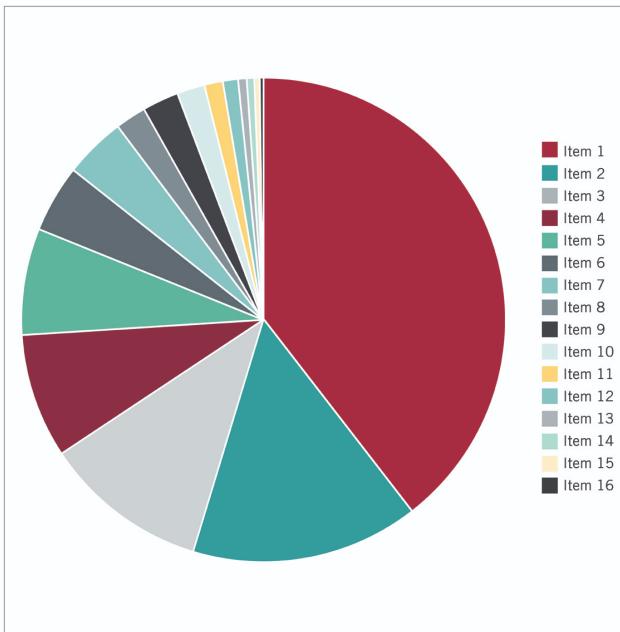
And often made even worse: Exploded Pie



And often made even worse: $\text{SUM}(\%) \neq 100\%$



And often made even worse: Many Levels

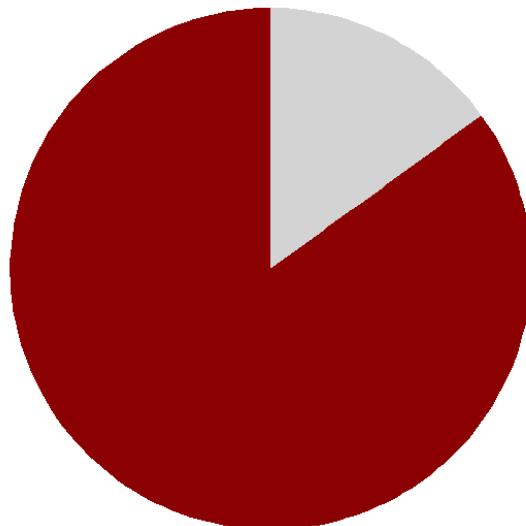


Key Takeaway 2

Please do **NOT** use pie charts.

- If you need any further evidence, please check `?pie()` in R. Even statistical software are recommending against using pie charts!!

Should you use Pie Charts?

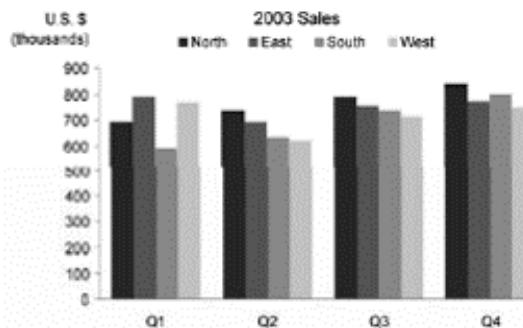
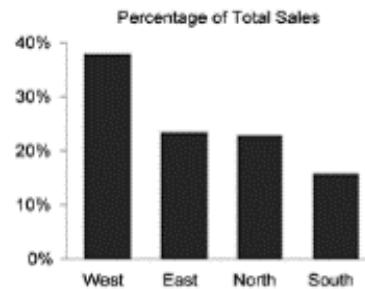
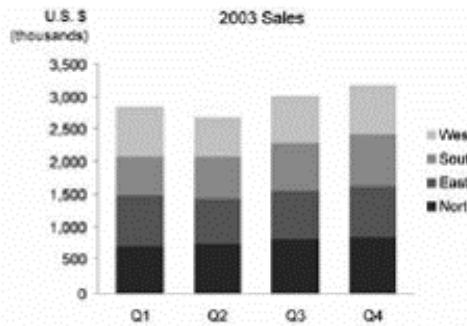


No
No, but in gray

Stacked Bar Charts

Activity Your Solution

When it is best to use the four charts below? They are using the same exact data.



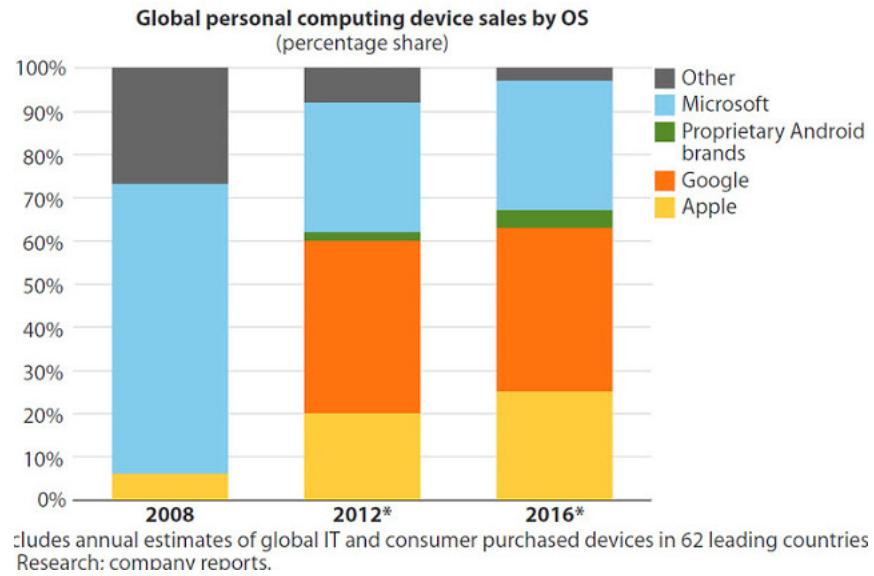
Stacked Bar Charts

Activity

Your Solution

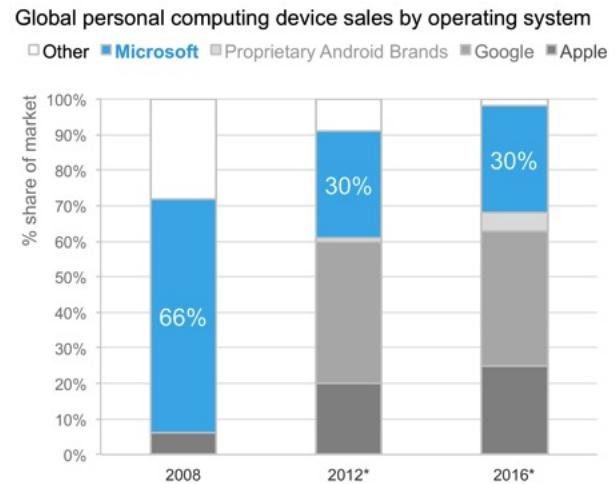
Insert best usage scenario for each chart below

A Note on Stacked Bar Charts



A Note on Stacked Bar Charts

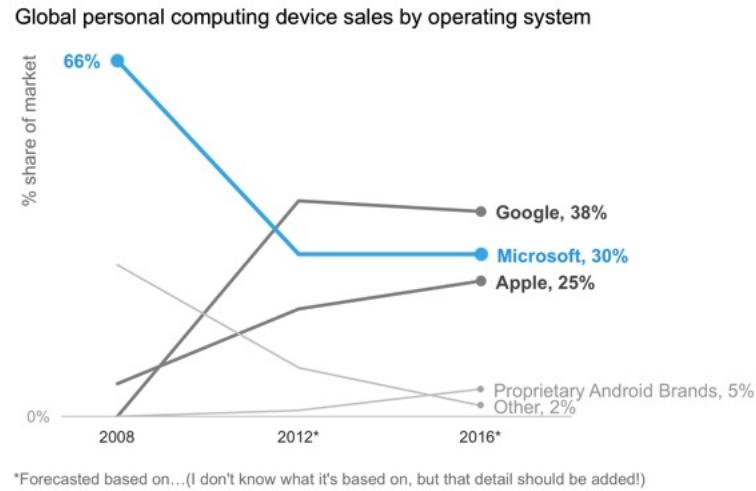
After spending years dominating the operating system market,
Microsoft is destined to become one of three major players



*Forecasted based on...(I don't know what it's based on, but that detail should be added!)

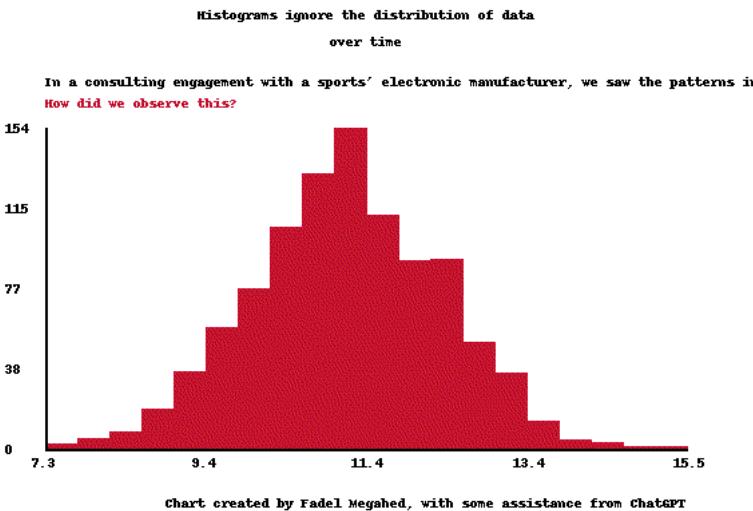
A Note on Stacked Bar Charts

After spending years dominating the operating system market,
Microsoft is destined to become one of three major players

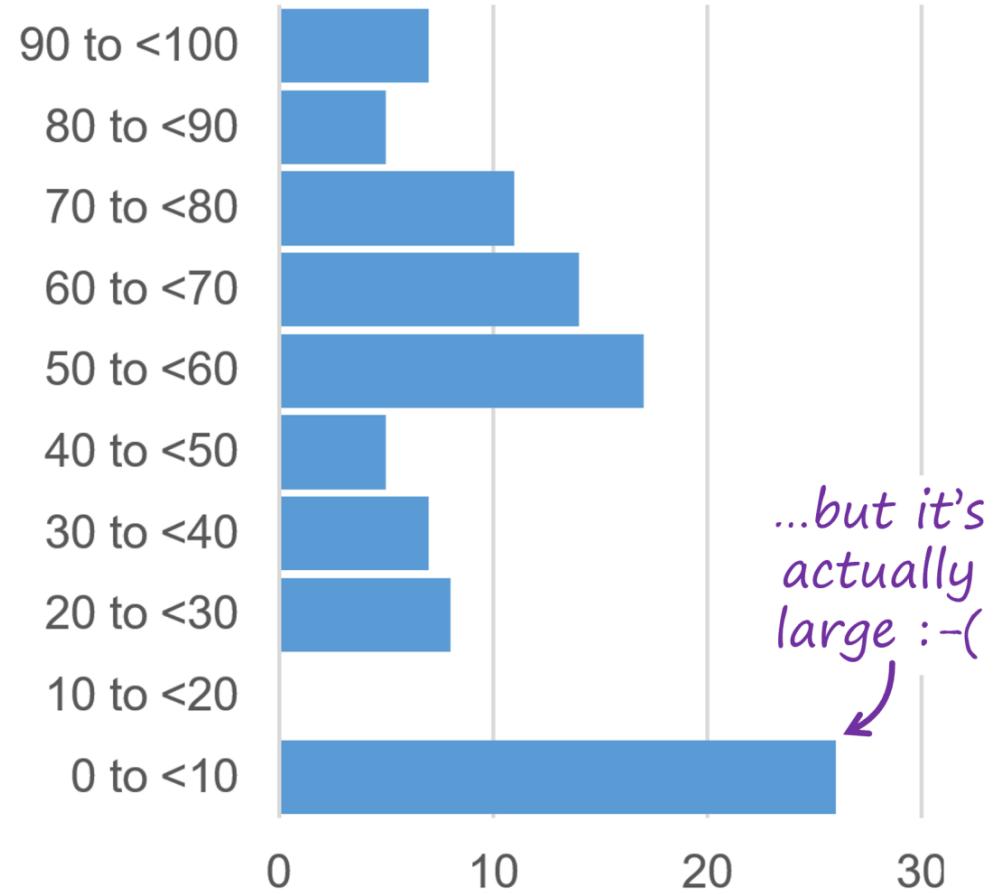
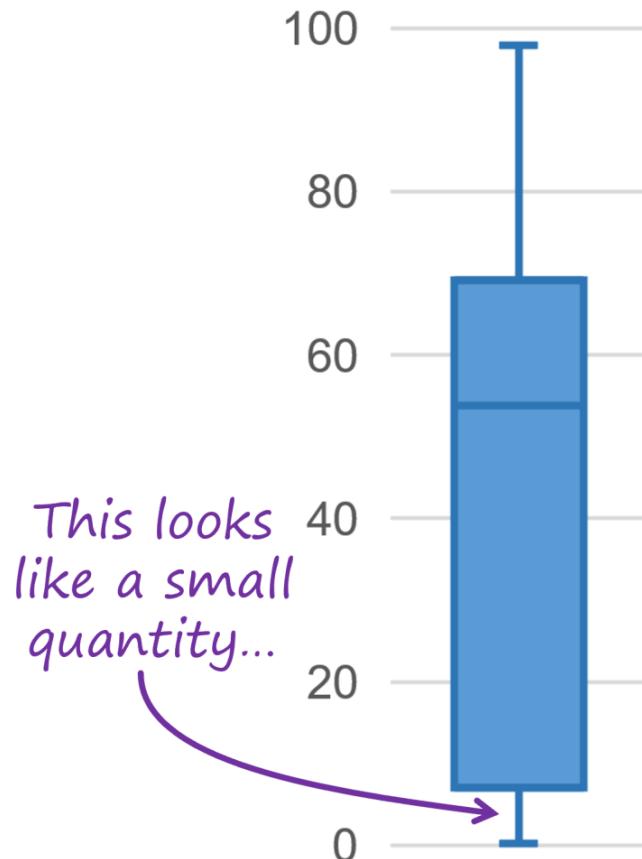


Distributions and Correlations

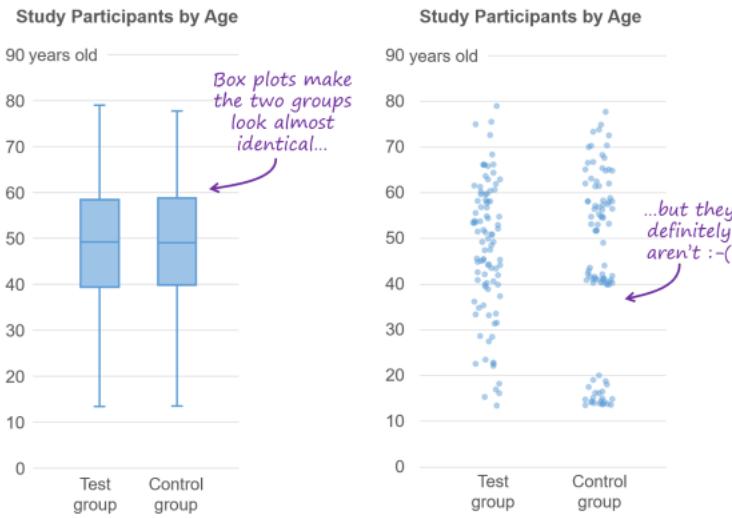
Issues with Histograms



Issues with Box Plots



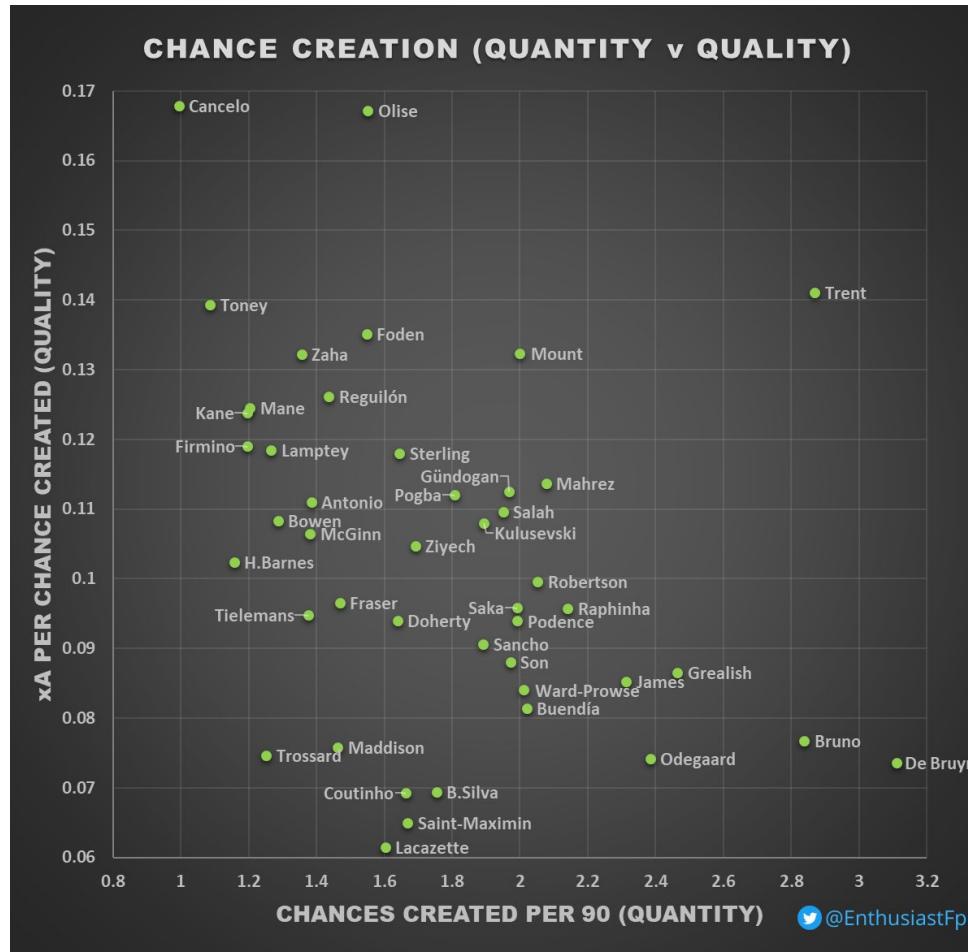
Issues with Box Plots



Key Takeaway 3

To capture the variability in a dataset, the use of box plots **may not be** appropriate!!!!

Scatter Plots



Recap

Summary of Main Points

- Identify strengths & weaknesses of basic charts
- Use appropriate charts based on objective
- Avoid using pie charts (never use pie charts)
- Avoid 3D graphs (unless VR changes their utility)