ISA 401: Business Intelligence & Data Visualization

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

Fadel M. Megahed, PhD

Raymond E. Glos Professor in Business Farmer School of Business Miami University

- @FadelMegahed
- fmegahed
- ✓ fmegahed@miamioh.edu
- ? Automated Scheduler for Office Hours

Fall 2025

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

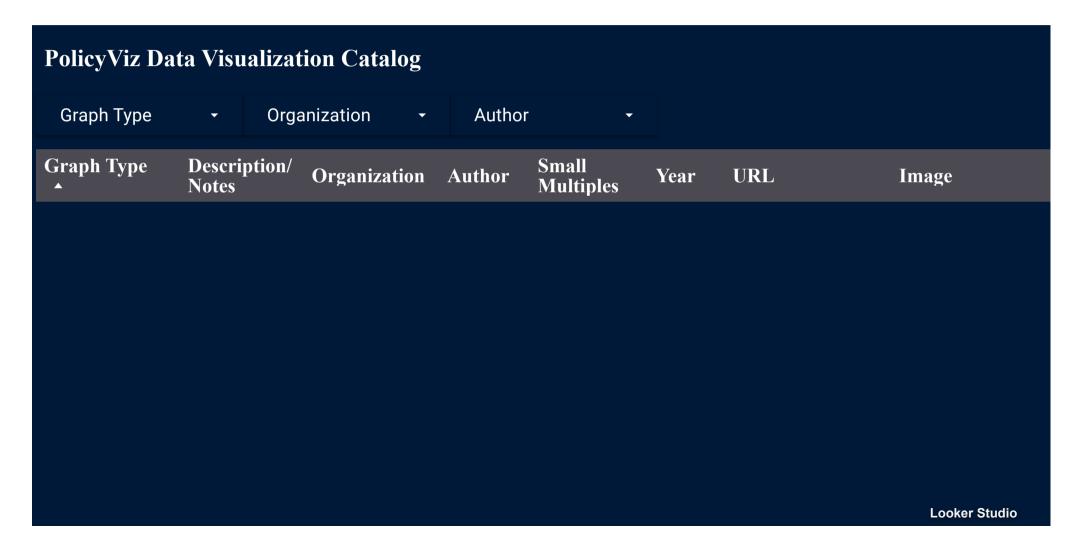


Chart Suggestions

Chart Chooser Single Bar Chart Multiple Bar Charts Variable Width Table or Table with Circular Area Chart Line Chart Column Chart Embedded Charts Cyclical Data Many Items Non-Cyclical Data Single or Few Items Many Items One Variable Two Variables Three or More Variables Many Periods Few Periods Static Changing Over Time Column Histogram Few Comparison Variable Scatter Chart Variables What would you Line Histogram Distribution Relationship Many like to show? Bubble Chart Composition ...O Variables Two - Variables Static Changing Over Time 3D Area Chart Three Variables Few Periods Many Periods Relative and Absolute Simple Share Accumulation or Components Only Relative Relative and Absolute Only Relative of Total Subtraction to Total of Components Differences Matter Differences Matter Differences Matter Differences Matter Stacked 100% Stacked Area Chart Waterfall Chart Stacked 100% Column Chart Column Chart with Subcomponents

© 2020 Andrew V. Abela, Dr. Abela@ExtremePresentation.com www.extremepresentation.com

Charts Used for Comparing Data (Unit of Analysis is Based on a Nominal Categorical Variable)

A Literal Bar Chart

04:00

Activity

Your Solution

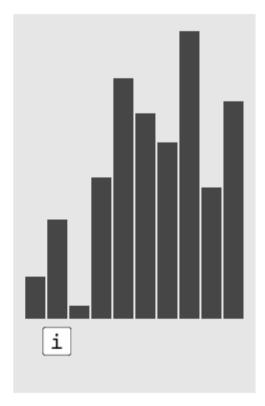


Answer the following questions:

- (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?

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
}</pre>
```



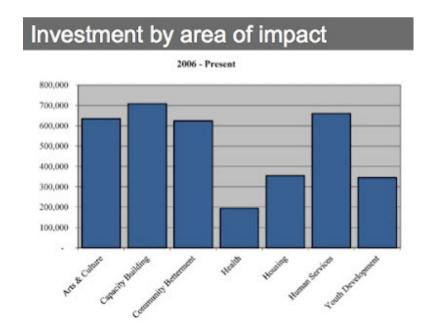
04:00

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



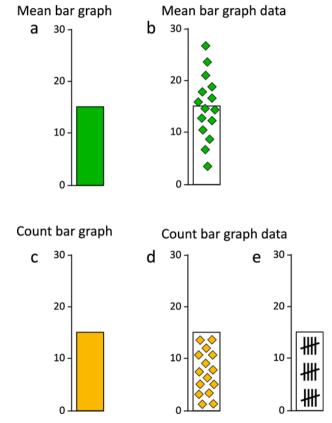
We invest primarily in four areas

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



Issues with the Interpretation of Bar Charts



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

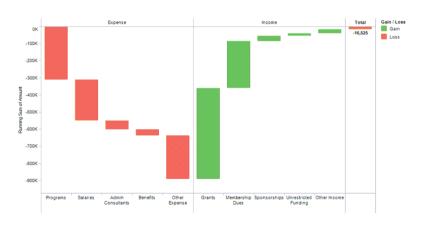
What are the advantages and disadvantages of these two charts? They are using the same exact data. Please try to list 2-4 in each category for each chart.

Income closely matched expenses in 2010

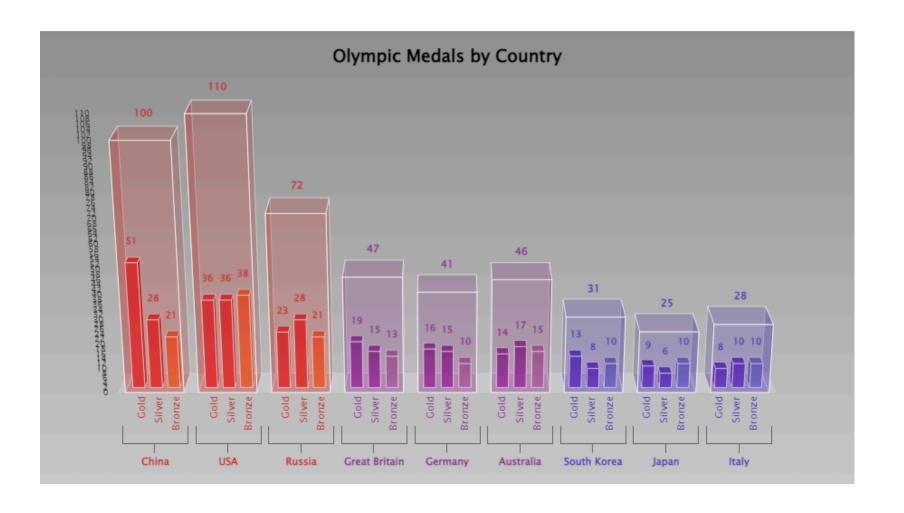
INCOME & EXPENSES: 2010 OVERVIEW EXPENSES INCOME



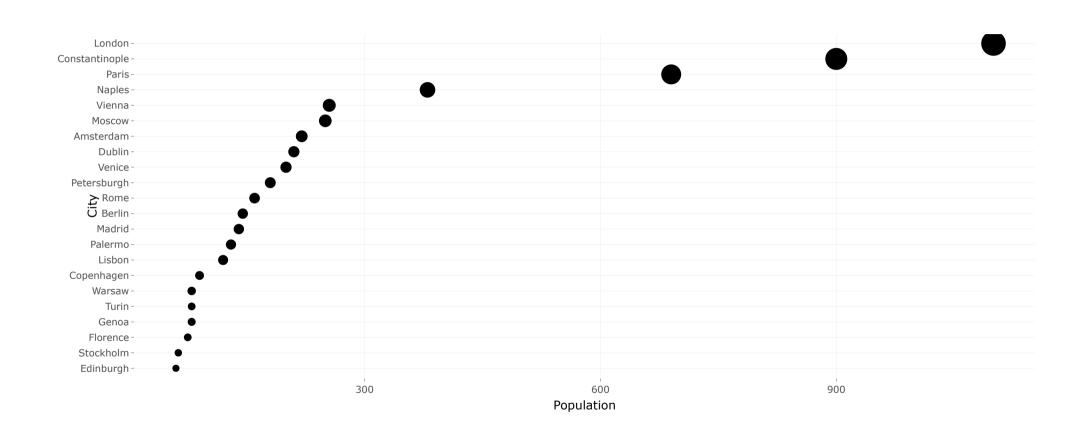
Whereas income is primarily from a single channel (Grants, \$531K in 2010, or 61% of total income), 2010 expenses were spread roughly evenly across programs, salaries, and other expenses. This means [...]



3D Charts are Awful: Even This



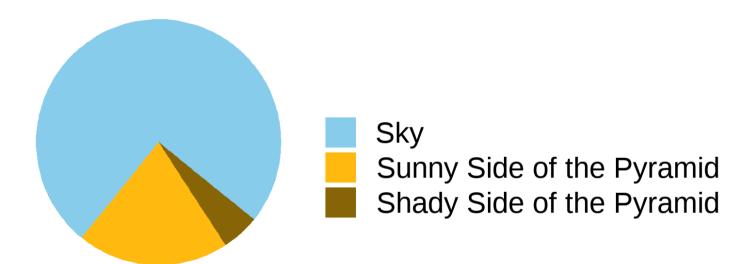
Dot Charts: Recall the Playfair Example



Proportions

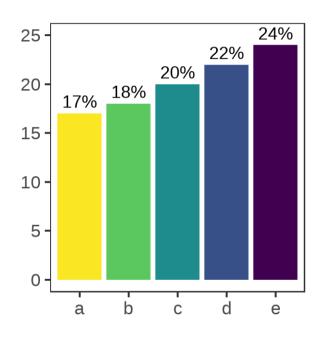
My Favorite Pie Chart

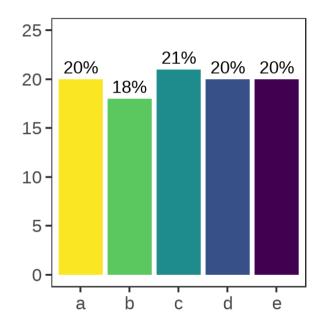
The Egyptian Pie Chart

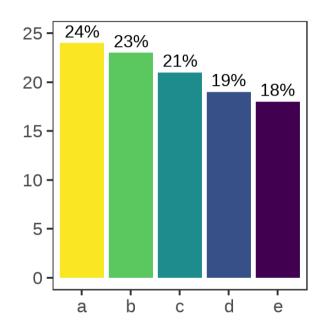


15/32

Pie Charts are Awful By Design

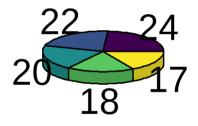




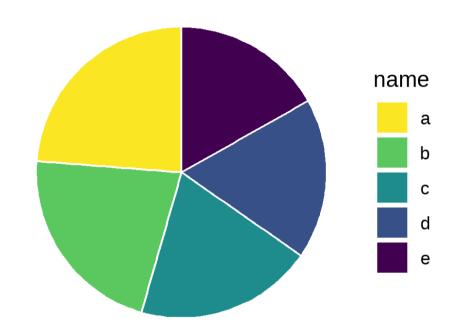


16/32

And often made even worse: 3D



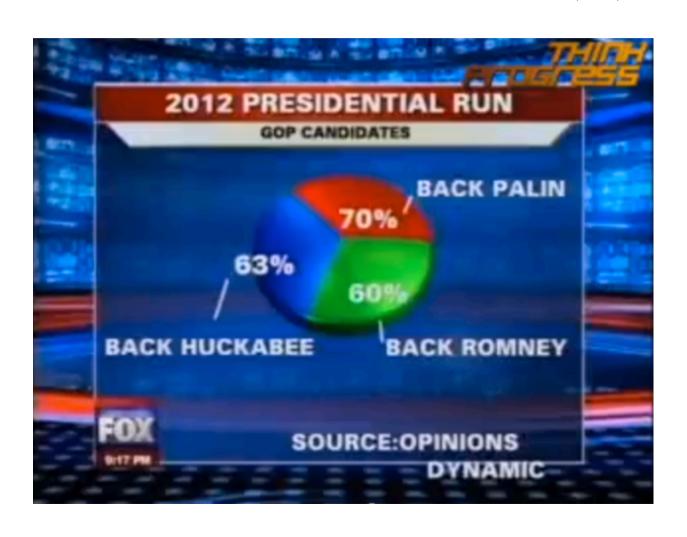
And often made even worse: Side Legend



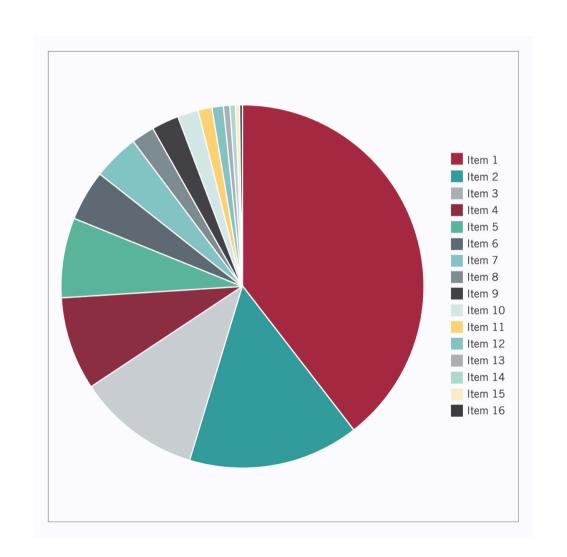
And often made even worse: Exploded Pie



And often made even worse: SUM(%) != 100%



And often made even worse: Many Levels

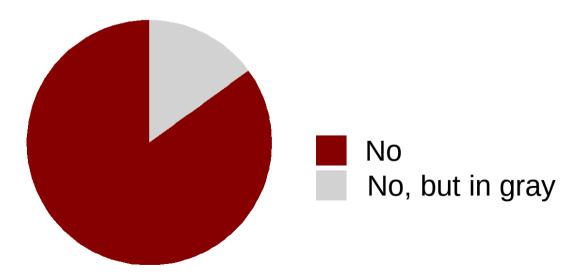


Key Takeaway 2

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?



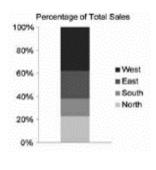
Stacked Bar Charts

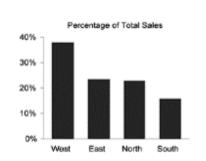


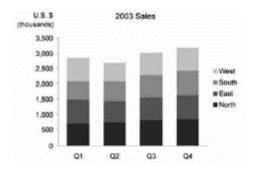
Activity

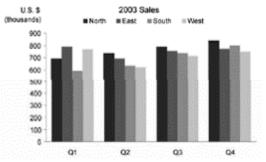
Your Solution

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





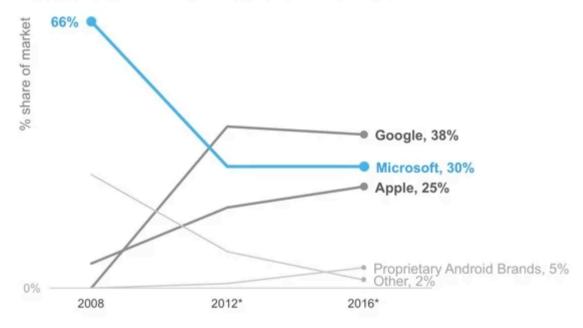




A Note on Stacked Bar Charts

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

Global personal computing device sales by operating system



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

Distributions and Correlations

Issues with Histograms

Histograms ignore the distribution of data over time

In a consulting engagement with a sports' electronic manufacturer, we saw the patterns in How did we observe this?

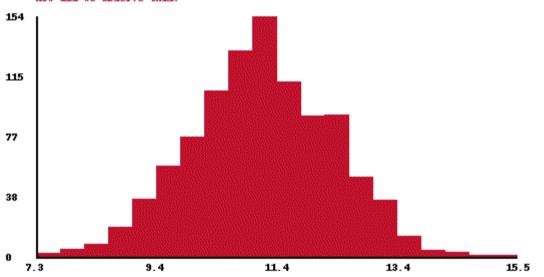
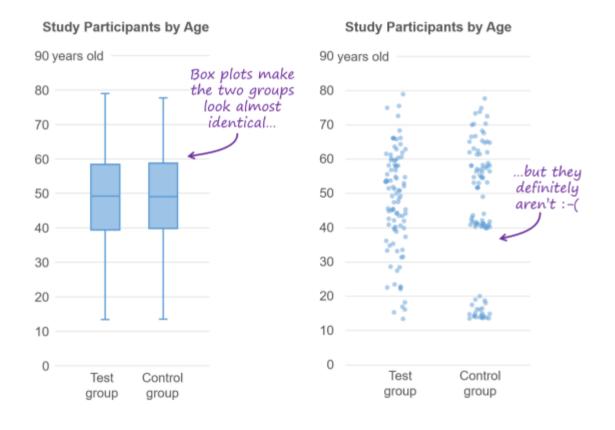


Chart created by Fadel Megahed, with some assistance from ChatGPT

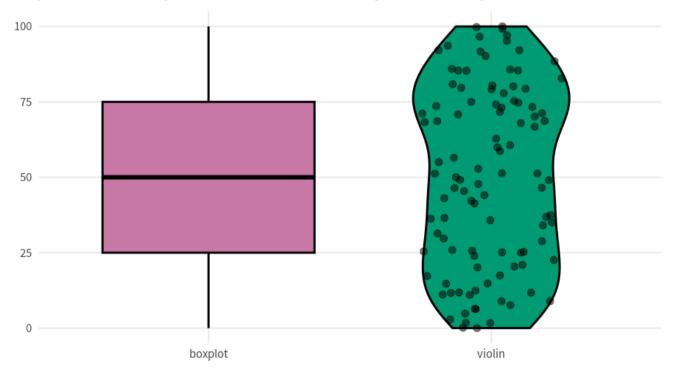
Issues with Box Plots



Additional Issues with Box Plots

Different Data, Same Boxplot

Boxplots can look exactly the same (even when the underlying data is wildly different.)



Source: Albert Rapp (2024) 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)