

Announcements

- **Take Home Assessment 3: Pokemon** due Thursday, April 30th at 11:59pm!
- **Peer Reviews for THA 2** due Wednesday, April 29th at 11:59pm!
- **Lesson 11 Canvas Quiz** due tonight at 11:59pm!
- **Reading Assignment 3** due Monday, April 27th at 11:59pm!
- **Project Part 1** due Friday May 1st at 11:59pm on Gradescope!
 - Choose either the Project (group of max 3) or Portfolio (individual)
 - You cannot switch once you've selected your option
 - If you are doing the Project, only one person need submit on Gradescope but make sure your partners are added to your submission! Instructions on Project Part 1 assignment!
- **Resub Cycle 2** will be made available today and due on Tuesday, April 28th
 - Eligible assignment(s): **THA 1**

Project / Portfolio Part 1

- Project
 - Original data analysis
 - Go beyond what you have shown in the THAs
 - Groups of max 3 people
 - **Proposal** - what data are you using, what questions are you trying to answer, what are the motivations, and what is your plan?
- Portfolio
 - Polishing and connecting creative component work
 - Portfolio ***does not*** replace original creative component work
 - Individual
 - **Vision Statement** - what are your components, feedback, theme, and plan?

Default Parameters

- Python lets you define default values for parameters in case a client does not pass a value for that argument

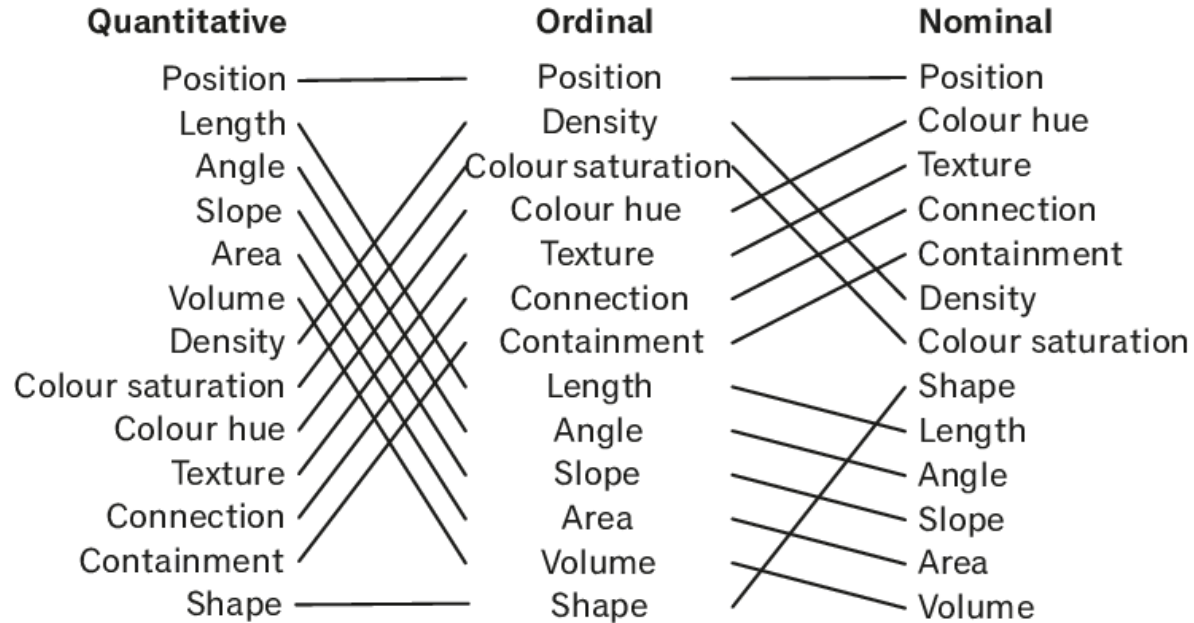
```
def greetings(to: str = 'world'):  
    return 'Hello ' + to + '!'  
  
greetings('Sarah')           # 'Hello Sarah!'  
greetings(to='Antonio')     # 'Hello Antonio!'  
greetings()                  # 'Hello world!'
```

- This allows libraries such as pandas or seaborn to provide default parameters for their functions and initializer functions!

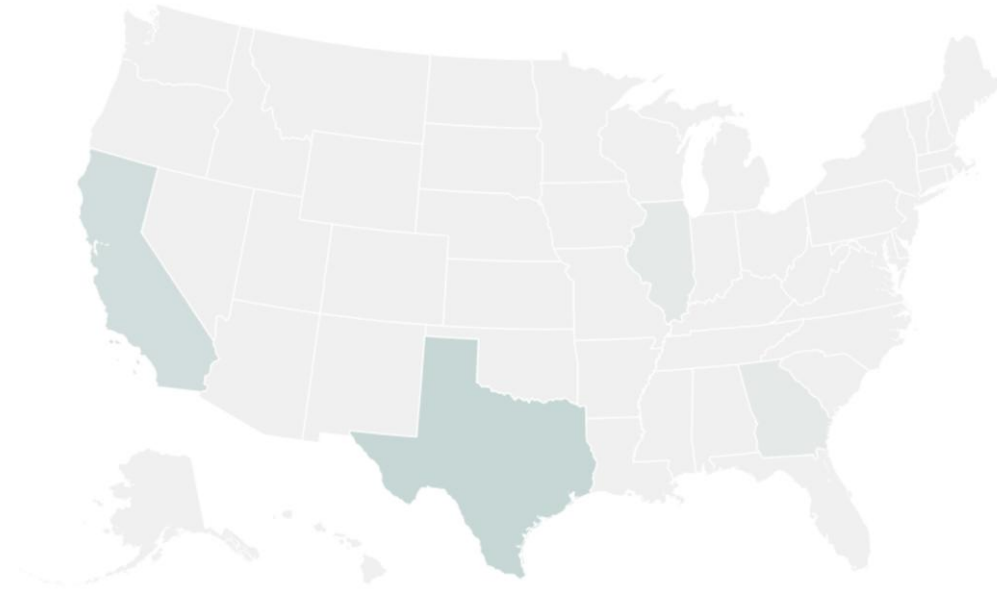
Python Data Visualization

- Many libraries exist for Python that allow you to do graphing and data visualization
- Two of the most popular graphing libraries for Python are
 - `matplotlib`: extremely customizable
 - `seaborn`: great out-of-the-box visualizations
- We will encourage `seaborn` since it is quickly becoming the standard for many data scientists
 - We will still rely on `matplotlib` for some things
 - You are welcome to use it in more depth on your project if you want to make more highly customized charts

Encoding Effectiveness

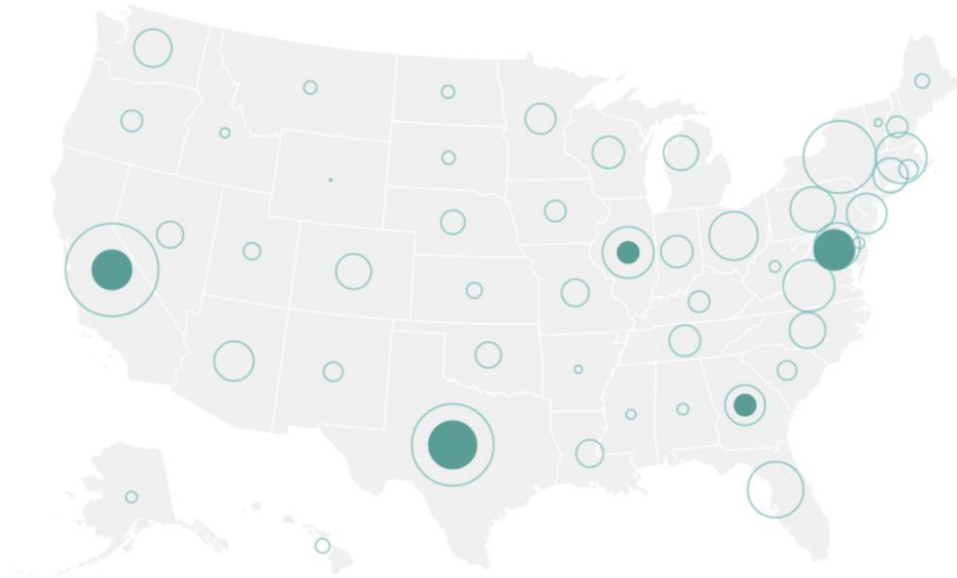


Color Encoding Example



Color Encoding

Area Encoding Example

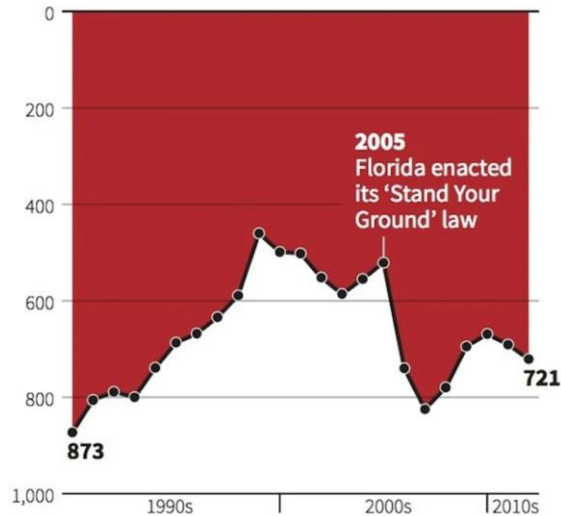


Area Encoding

Readability is Important!

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

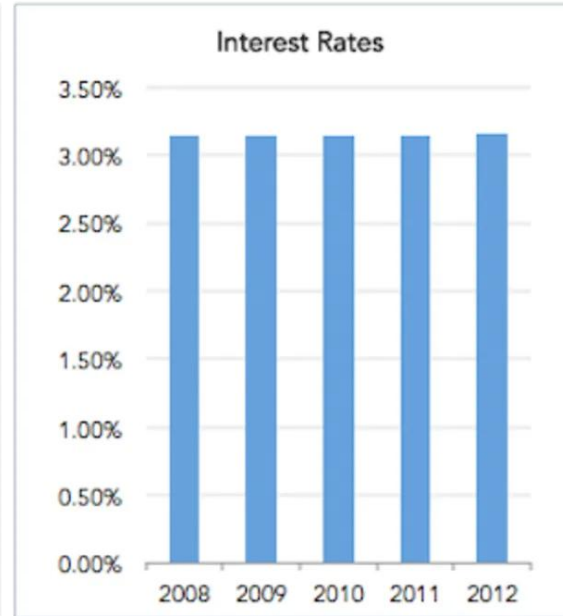
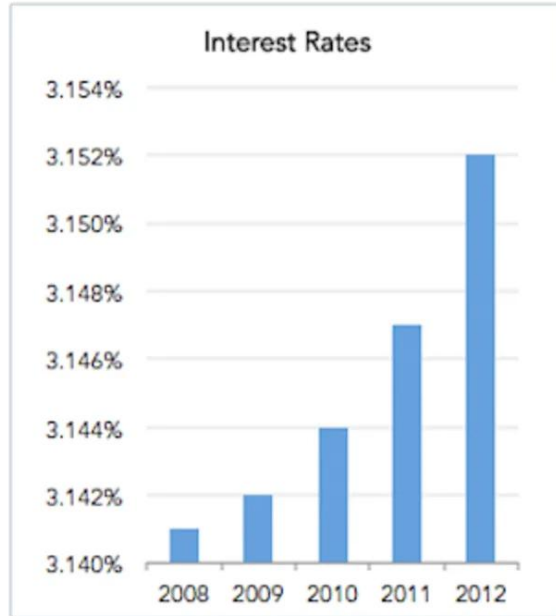
REUTERS

Did the number of deaths go up or down after the law was enacted?

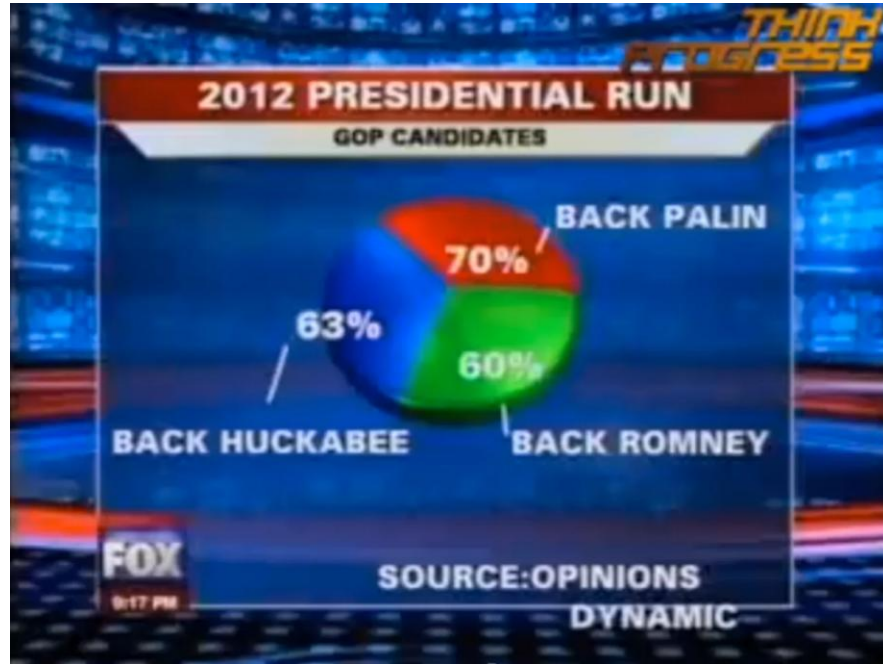
This graphic was posted in an [article by Business Insider](#) (graphic made by [Reuters](#))

How to Lie with Data Visualization

Same Data, Different Y-Axis



How to Lie with Data Visualization



Remember...

- Just because seaborn or matplotlib allow you to do something with your visualizations doesn't mean that's the most effective way to to present results.
- Be mindful of the ethical and conclusion implications that may come from presenting your data in a specific way.
- Data visualizations are meant to give us a relatively quick summary of your results. People should not be trying to decode what your visualization is saying.
- Be mindful of the color, font, and font size you use in your graphics
 - Accessibility matters when it comes to graphics
 - Good starting point for [accessible color palettes](#).