

Read Chapter 5 (Data Visualization)  
of the Textbook

Pictures worth a thousand word

Human brain is good at understanding information conveyed through visualizations than just text and numbers

## Principle of good Data Visualizations

---

- Serve a reasonably one single clear purpose
- don't over-complicated your graphs
- Encourage viewer to compare different pieces of data
- Induce viewer to think of the substance of data
- Avoid distorting what the data have to say
- Show the data with easy to understand scale
- Reveal the data at several levels of detail

# Data Visualizations Toolbox

---

## Libraries

- Matplot lib
- Pandas plot
- **Seaborn**
- **Plotly**
- Geographic plots

## Type of Plots

- Pie Chart / Bar Chart
- Scatter Plots
- Distribution Plots
- Box Plots
- **FacetGrid**
- Heatmap

## More Advanced visualizations

- Geographical Maps (KML)
- Interactive Plots
- Dash
- D3
- 3D plots
- **Folium**

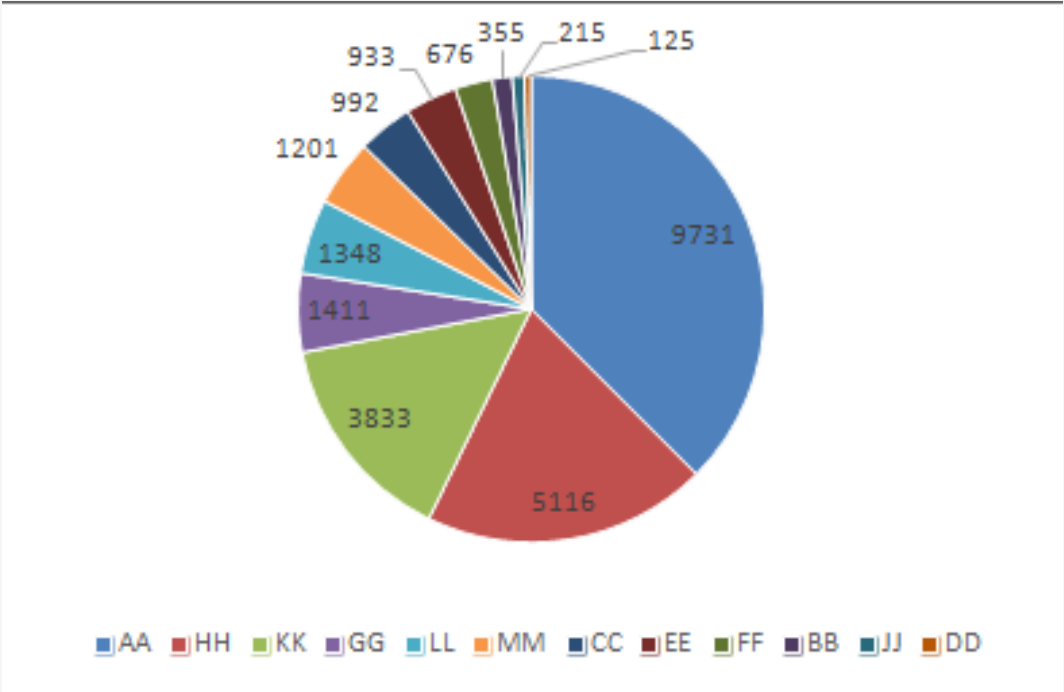
**Box Plot (done in 242)**

**FacetGrid - Puts data in buckets**

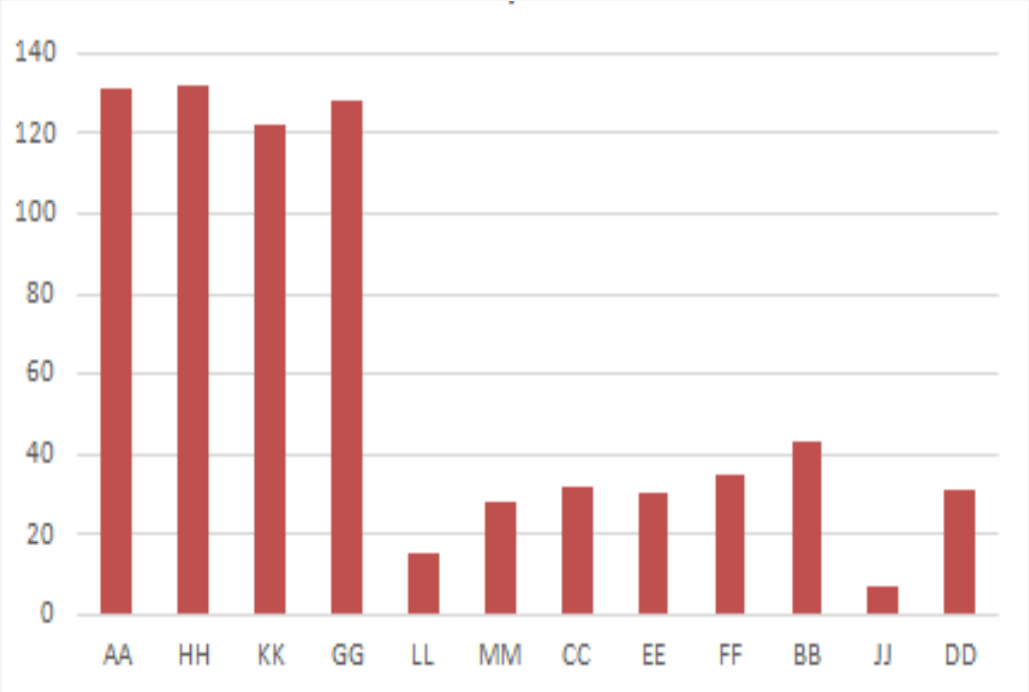
**Heatmap - Like covid vaccines for the hot/high density areas**

# Basic Business Data Visualization

Revenue by Products

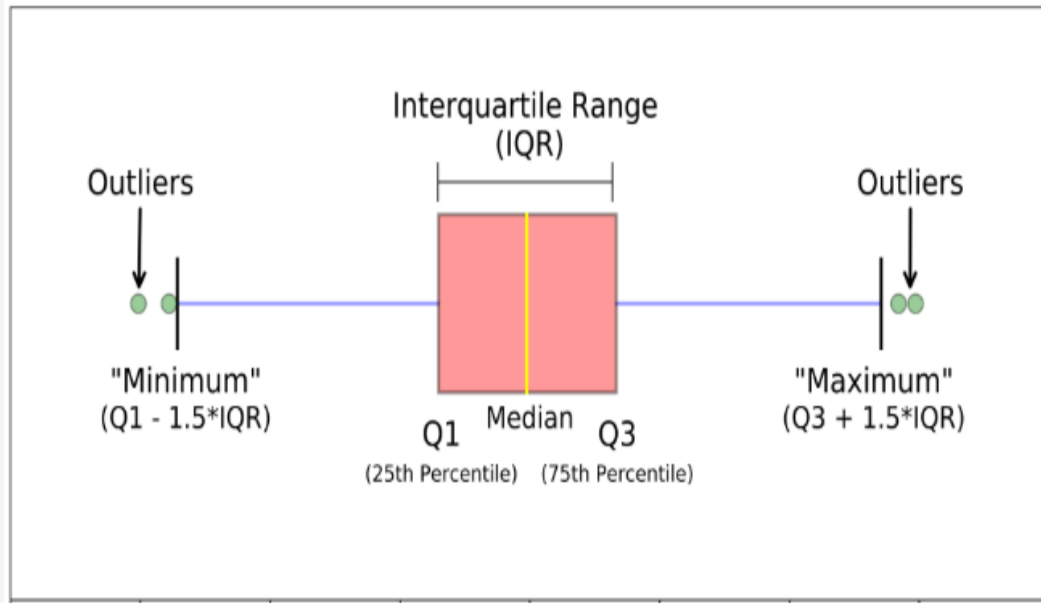


Orders by Products

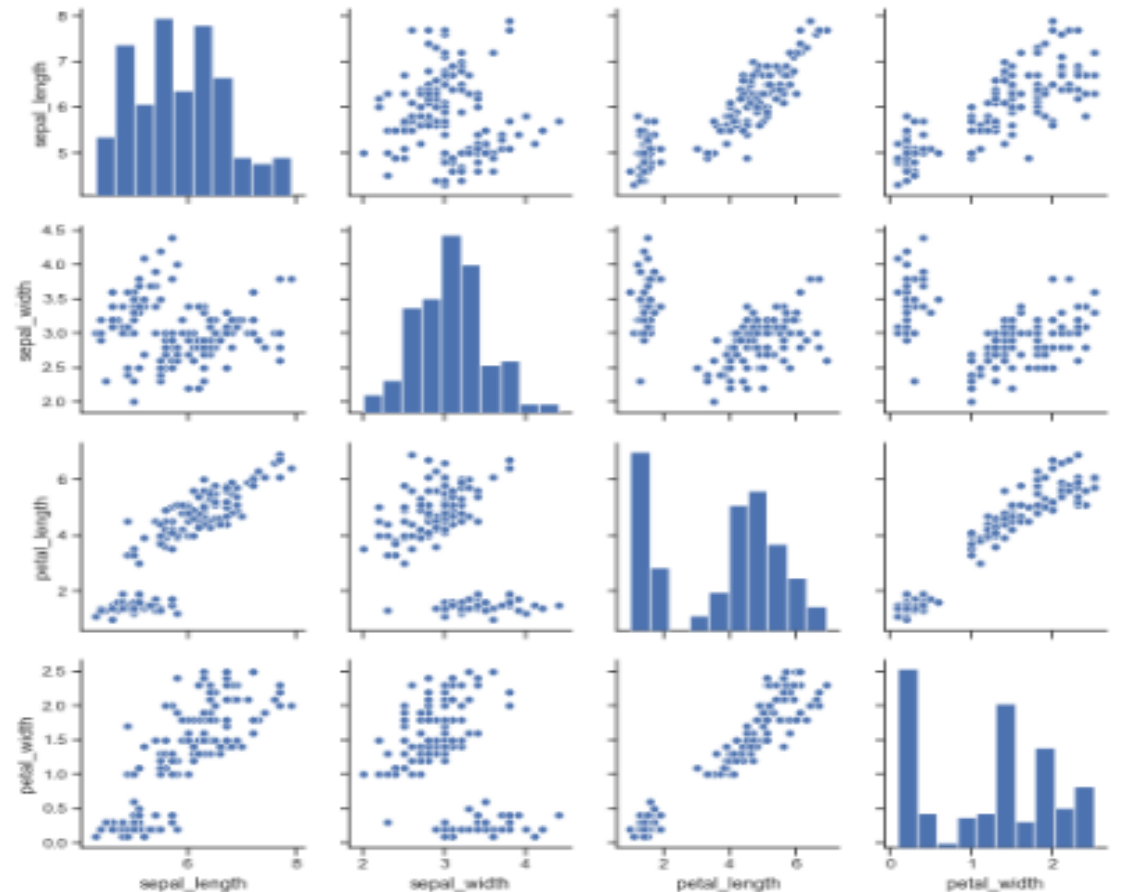


# Common Plot type

## Box Plot



## Pairs Plot



# Common Plot type

---

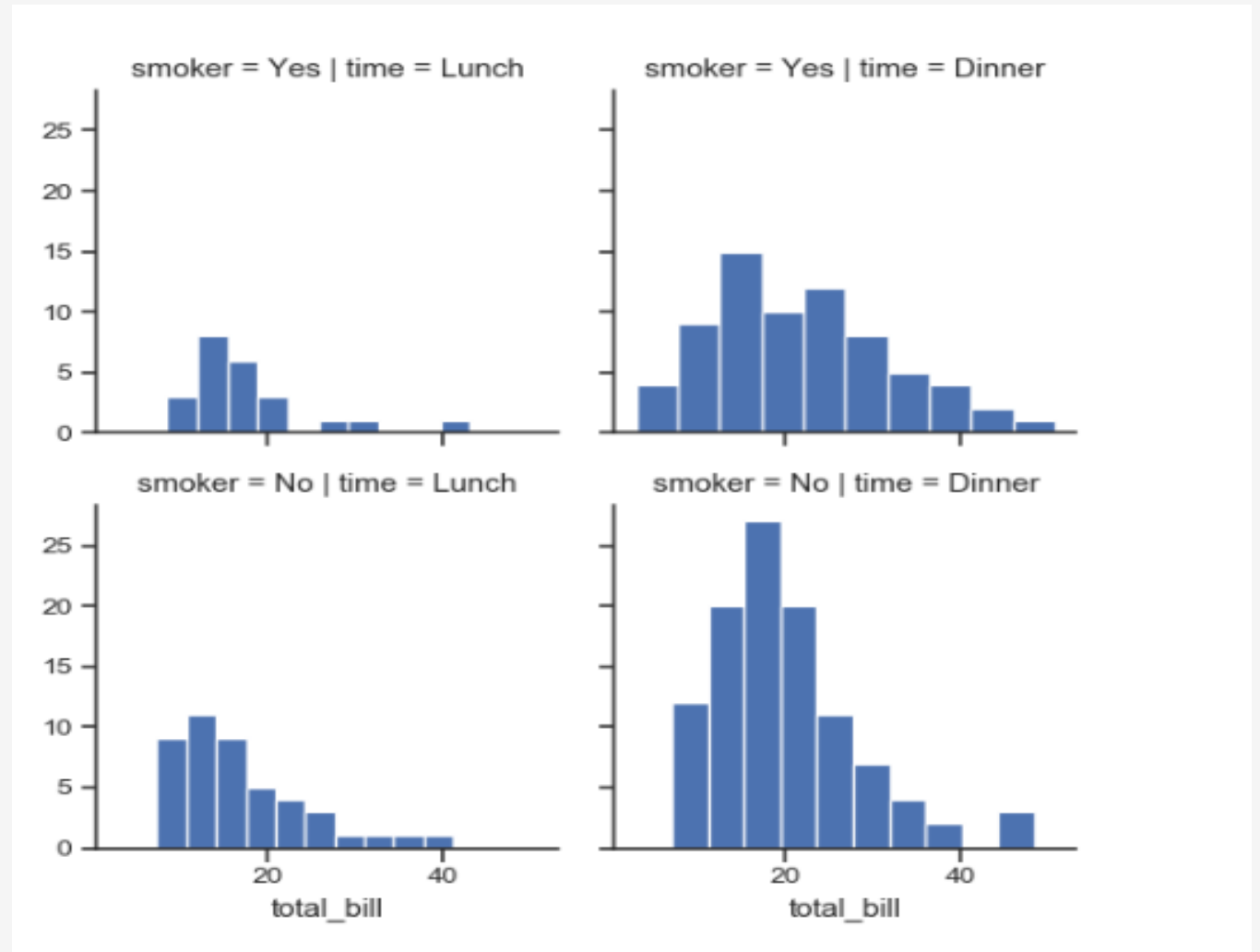
## FacetGrid

Tips data set

Total\_bill

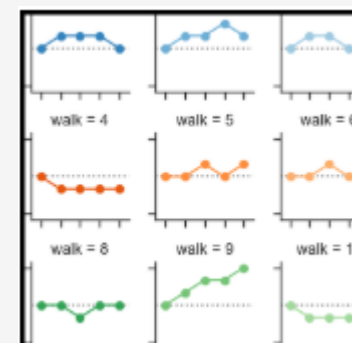
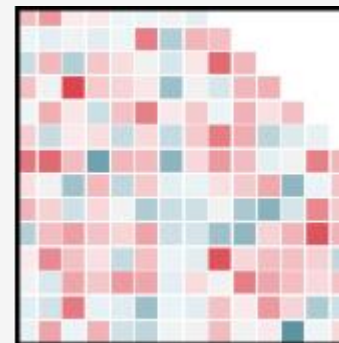
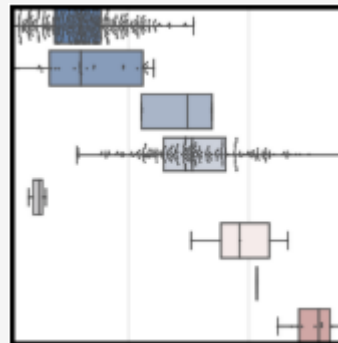
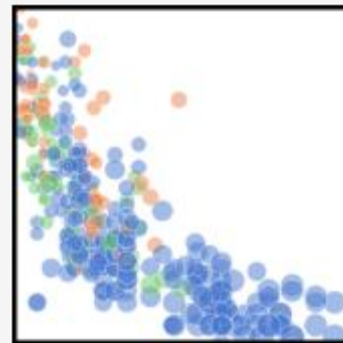
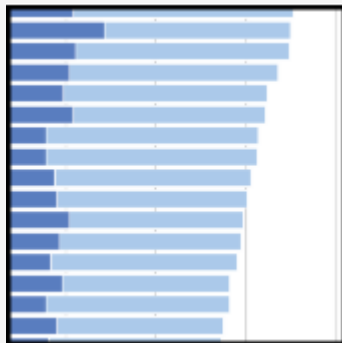
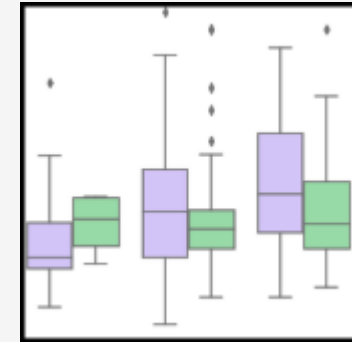
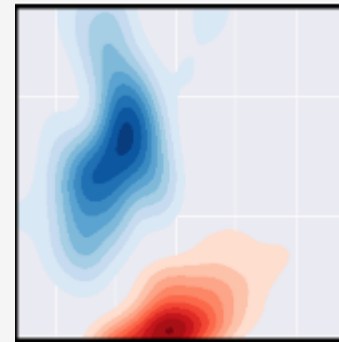
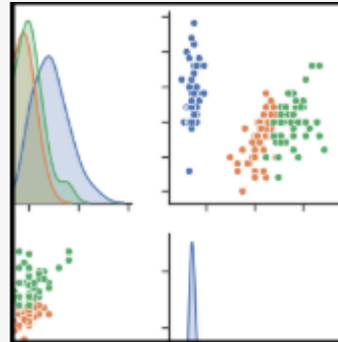
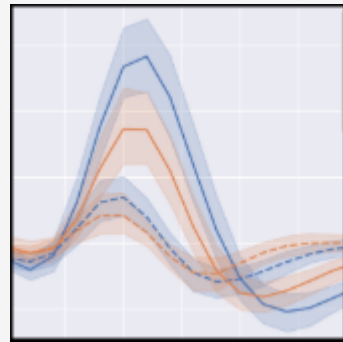
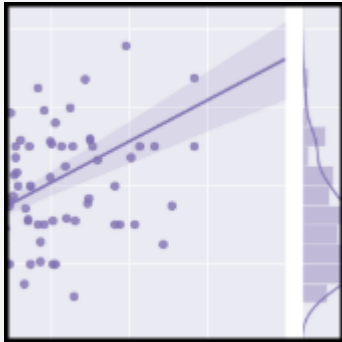
Smoker = Yes or No

Time = Dinner or Lunch



# More advanced Types of Visualizations

---





# More advanced Types of Visualizations

## Geographical Data Visualizations

- Google Map
- Folium Python library (pip install folium)

<https://python-visualization.github.io/folium/quickstart.html>

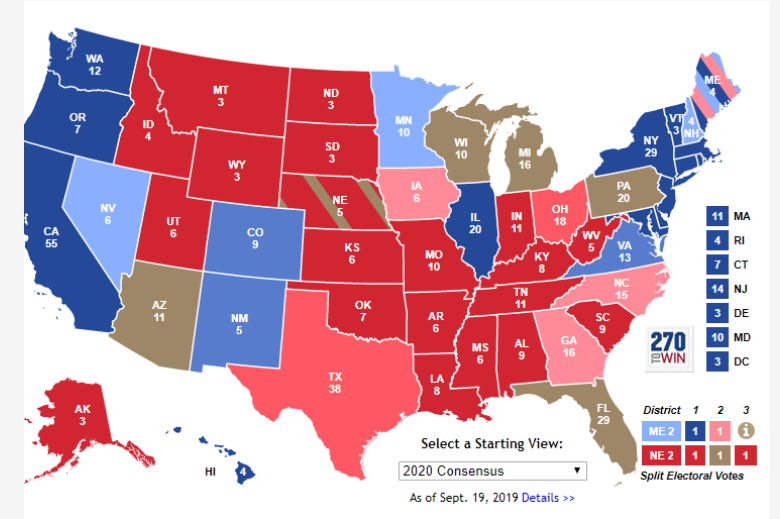
- KML (pip install simplekml)

<https://pypi.org/project/simplekml/>

- Google Earth

<https://www.google.com/earth/versions/#earth-pro>

**Purpose is to show the hot areas or certain heat signatures based on the data**

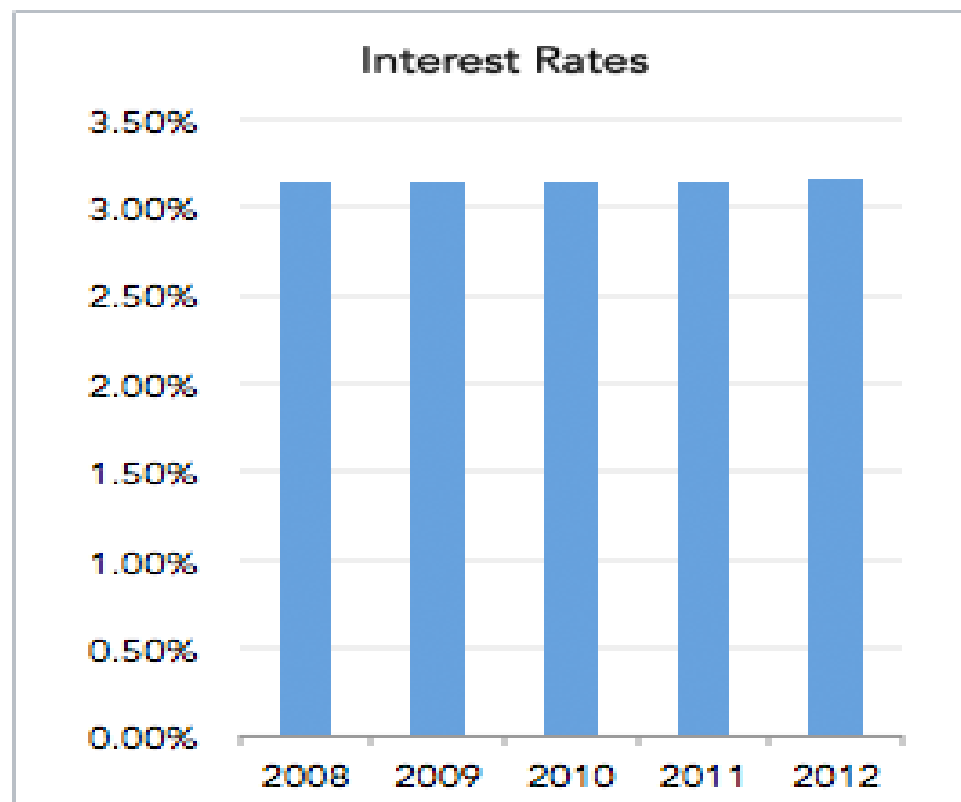
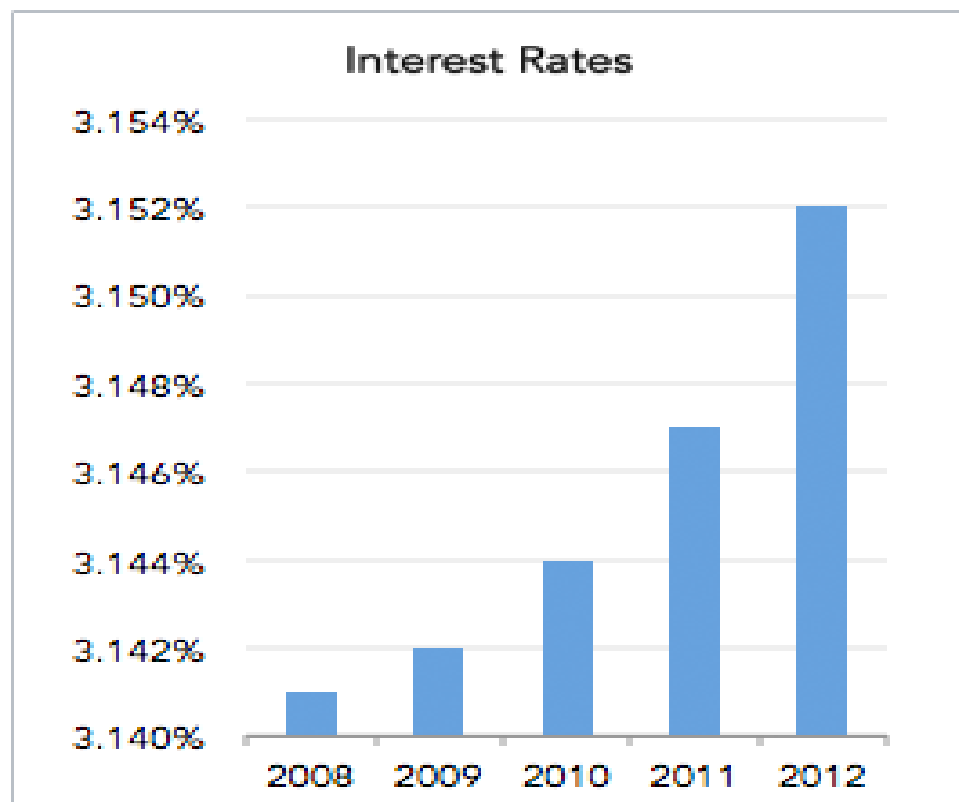


# How to misuse or even lie Data Visualization

Mess around with your scale in the axis

**Changing the scale shows that the rates are the same when they are actually not**

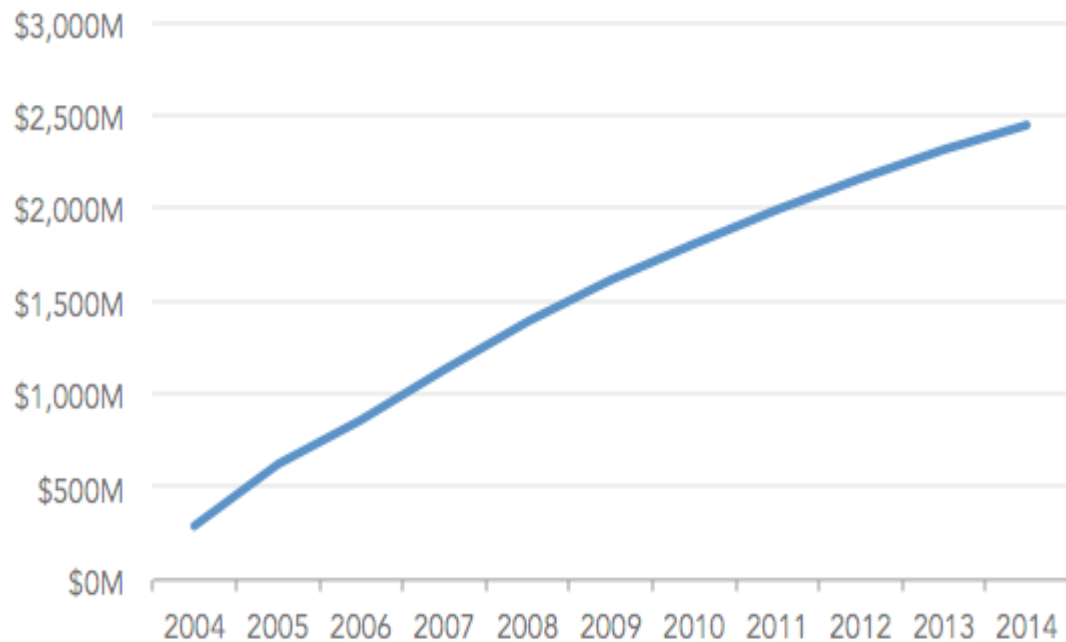
## Same Data, Different Y-Axis



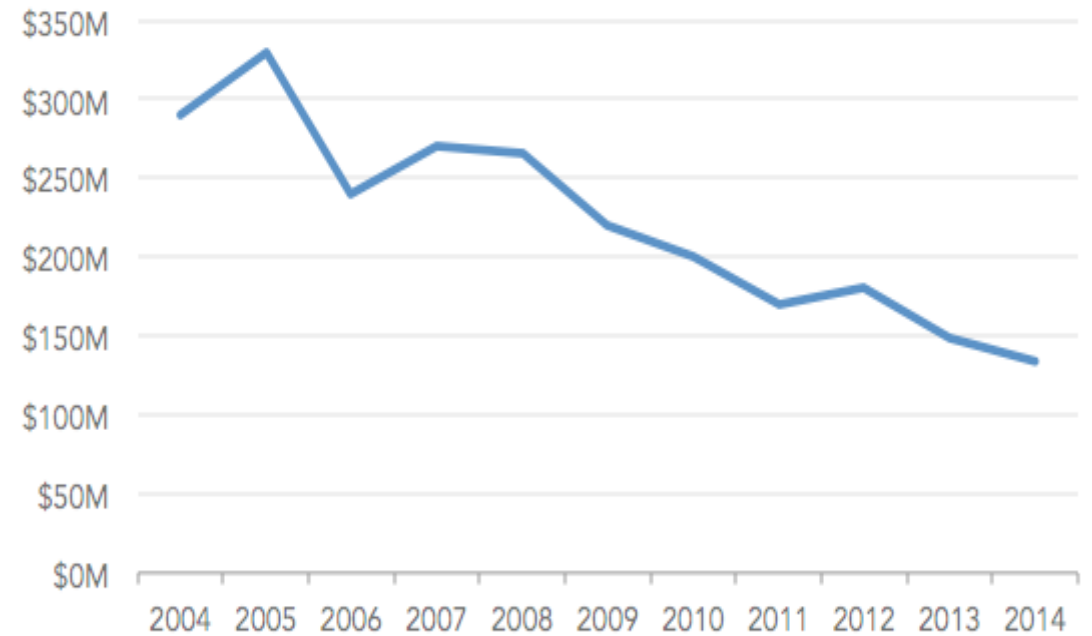
# How to misuse or even lie with Data Visualization

## Use Cumulative graph

**Cumulative Annual Revenue**



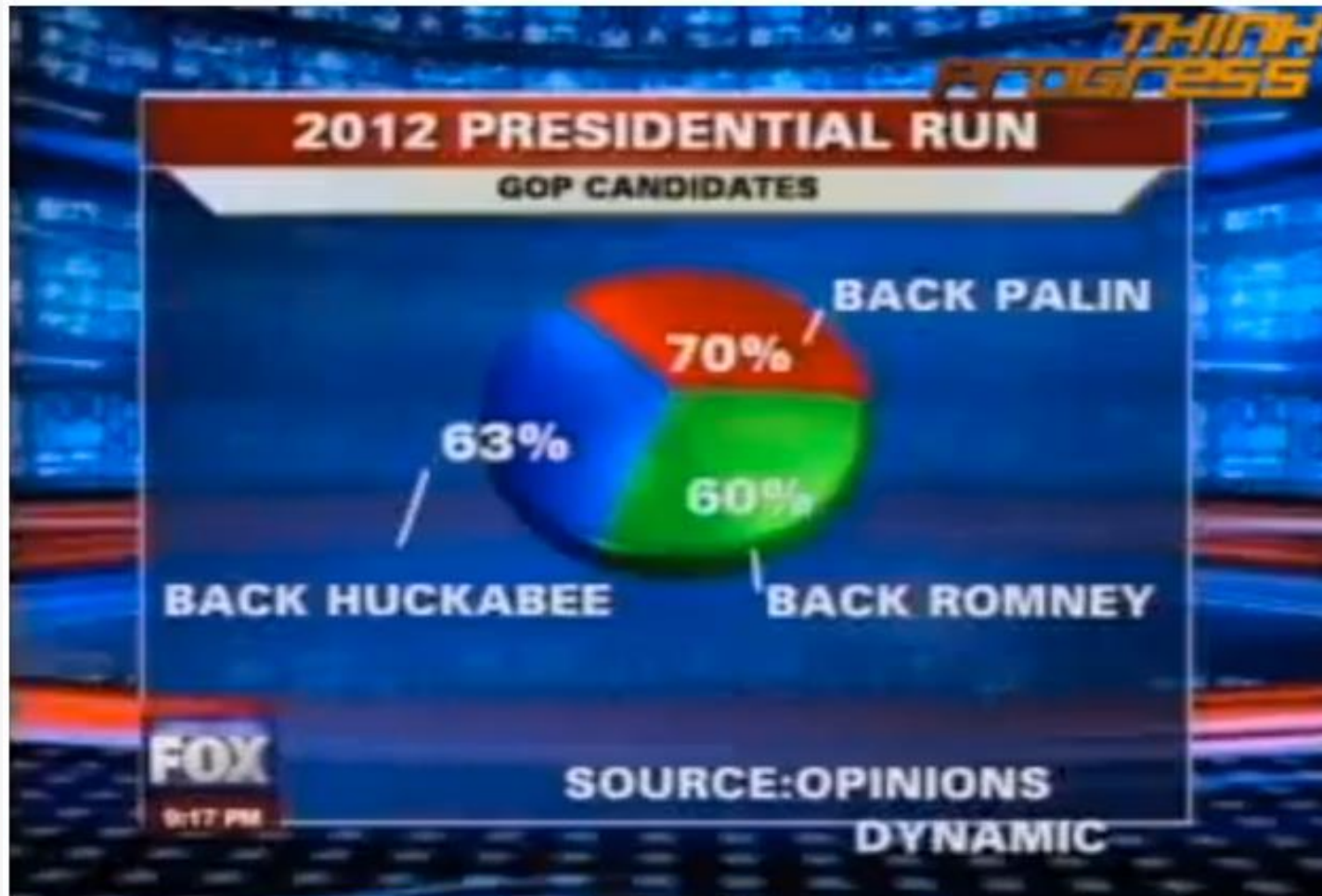
**Annual Revenue**



# How to misuse or even lie with Data Visualization

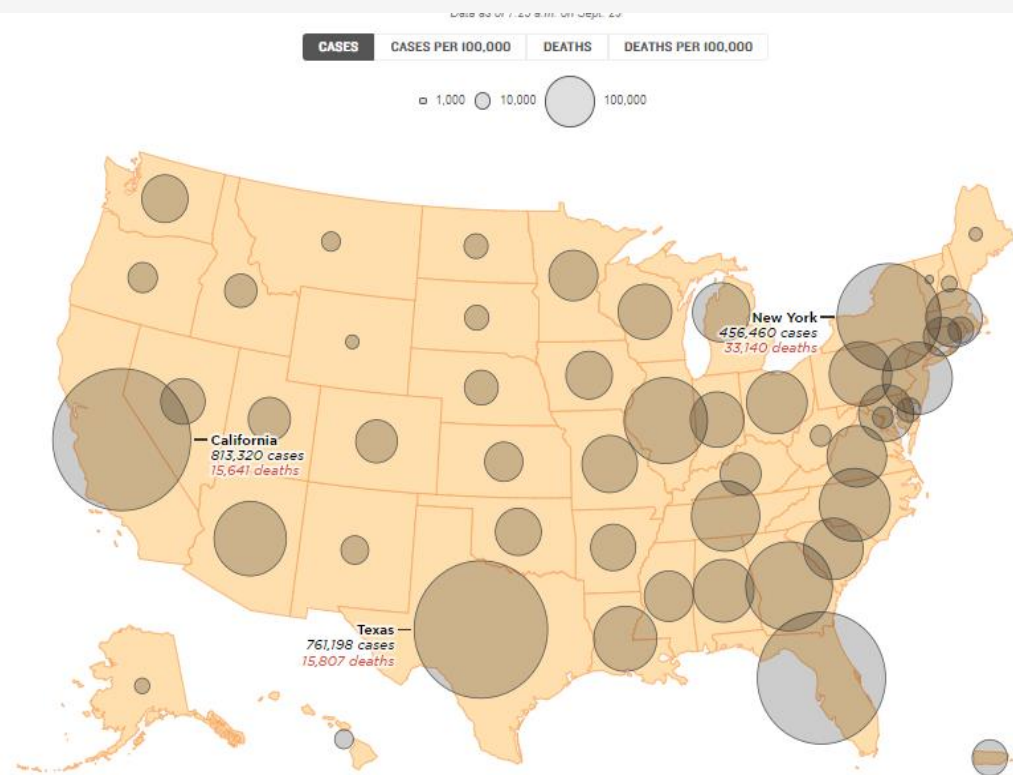
---

Ignore convention



# How to misuse or even lie with Data Visualization

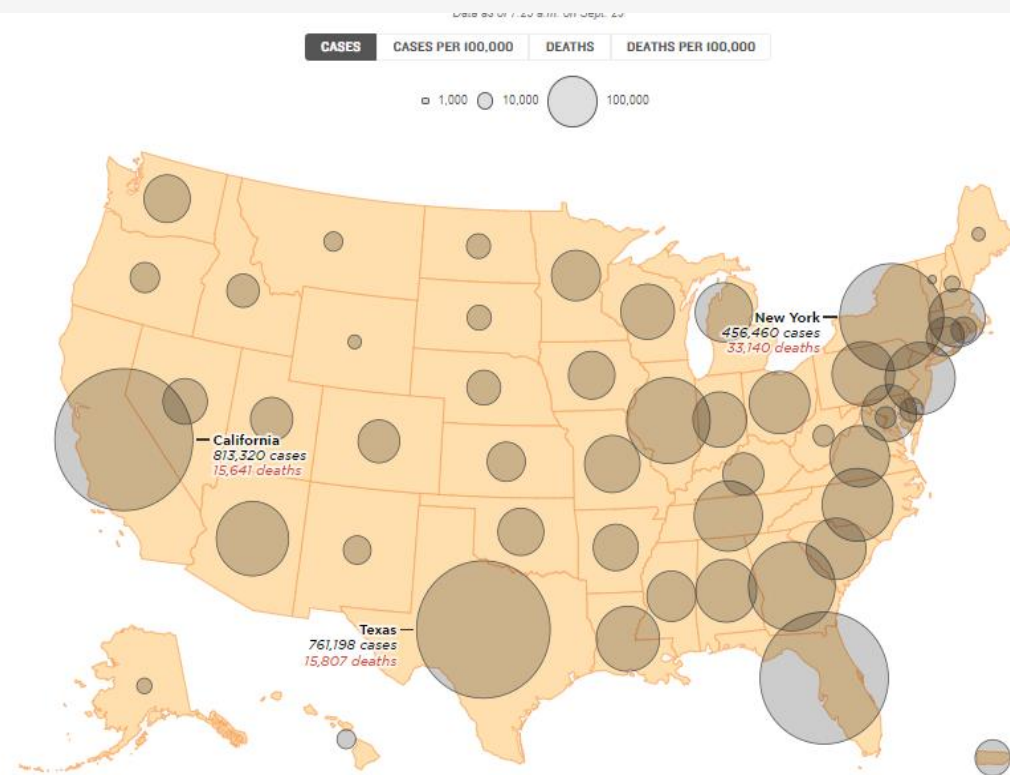
Date 1



On Sept. 2, Massachusetts changed its methodology for counting probable cases, which lowered its overall case and death numbers.

Source: Center for Systems Science and Engineering at Johns Hopkins University

Date 2 looks like nothing had changed, but in fact the numbers had changed because the scale has changed.  
e.g. same circle represents 120,000 instead of 100,000



On Sept. 2, Massachusetts changed its methodology for counting probable cases, which lowered its overall case and death numbers.

Source: Center for Systems Science and Engineering at Johns Hopkins University



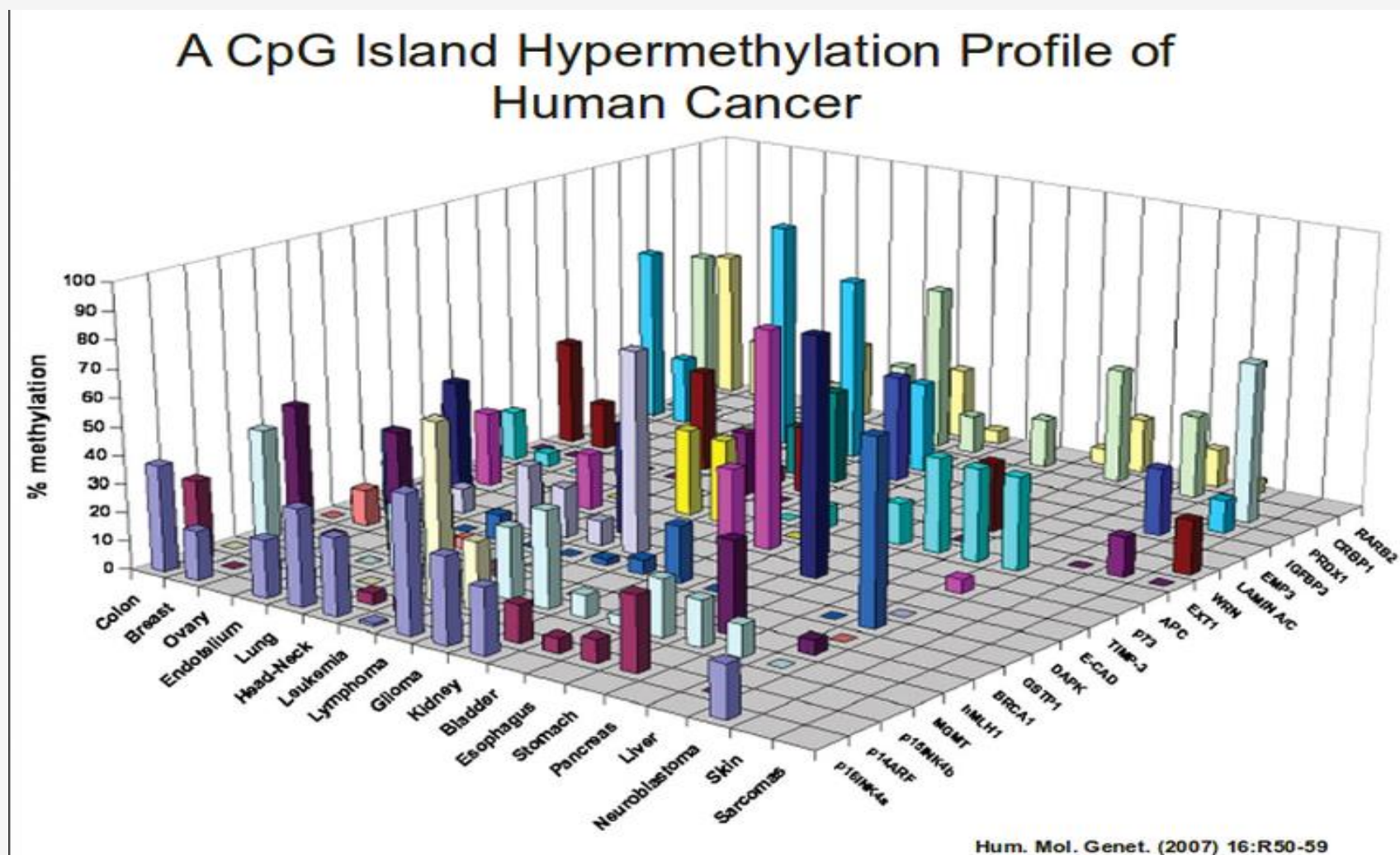
# Bad way to tell a story from Data Visualization

## Criteria for a bad plot:

Take more than 10 min to understand and get the takeaway from the plot

Storytelling with Data

<https://www.amazon.com/Storytelling-Data-Visualization-Business-Professionals-ebook/dp/B016DHQSM2>



## The other extreme

---

# Tim Cook and other Apple leaders use this clever presentation hack to make their slides memorable — and they borrowed it from Steve Jobs

When Apple CEO Tim Cook began talking about a new release of Apple's mobile operating system (iOS 13), he said: "iOS has the highest customer satisfaction in the industry, with an incredible 97%." The slide had one number in large font — 97%. In smaller font beneath the number, a sentence read: "Customer satisfaction for iOS 12." That's it. One number and one sentence.



[https://www.businessinsider.com/apples-leaders-use-this-presentation-hack-to-make-slides-memorable-2019-6?utm\\_content=buffer5934a&utm\\_medium=social&utm\\_source=facebook.com&utm\\_campaign=buffer-bi](https://www.businessinsider.com/apples-leaders-use-this-presentation-hack-to-make-slides-memorable-2019-6?utm_content=buffer5934a&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer-bi)

Best showcase of how to use Data Visualization

---

## Hans Rosling: Master of Data Visualizations

[TED talk on Public health and longevity](#)

<https://www.youtube.com/watch?v=hVimVzgtD6w>

[Income disparities](#)

<https://www.youtube.com/watch?v=DoSTNRhoceY>

<https://www.youtube.com/watch?v=AdSZJzb-aX8Ed/PBS>

Accenture talk at CWRU

<http://www.youtube.com/watch?v=qprHllzhgUk> .

<https://www.gapminder.org/answers/how-does-income-relate-to-life-expectancy/>



# Data Visualizations Toolbox

---

Learning by doing