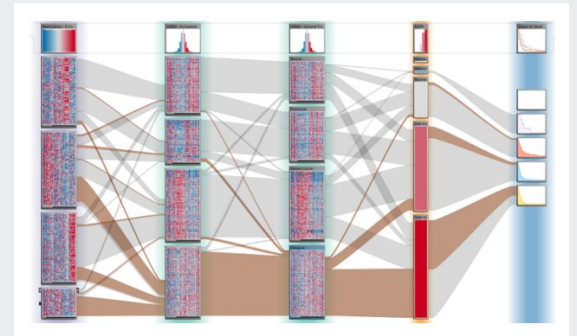




# Visualización

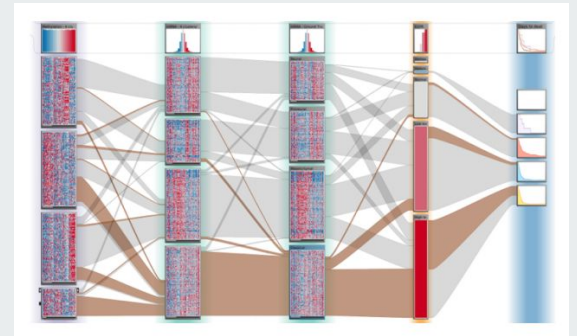
75.06 Organización de Datos





# Visu

75.06 Organización de Datos



## What is data visualisation?



The representation and presentation of data to facilitate understanding



# Visu: 4 Elementos

- Datos
- Objetivo
- Metáfora Visual
- Historia

## What Makes a Good Visualization?





# Objetivos

- Como parte del proceso de análisis de datos
- Para comunicar resultados o conclusiones

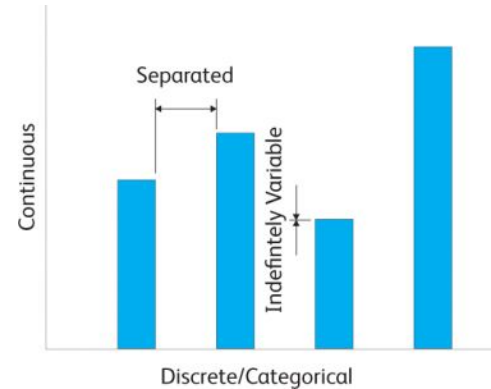
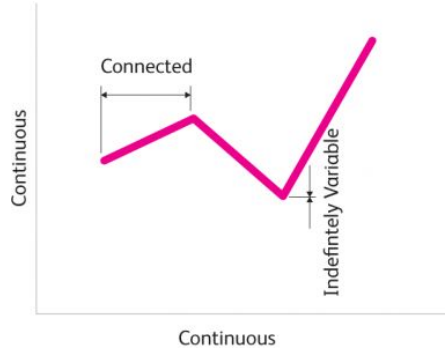


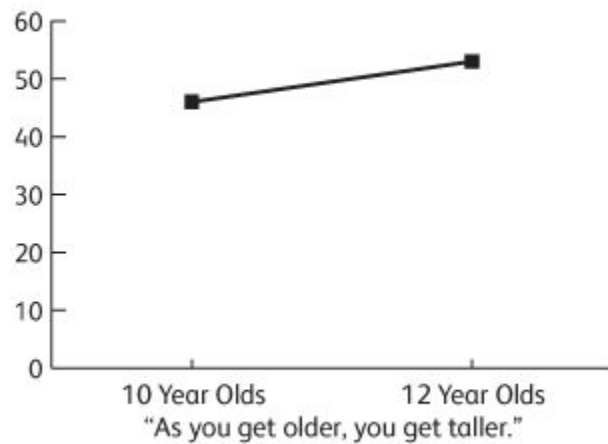
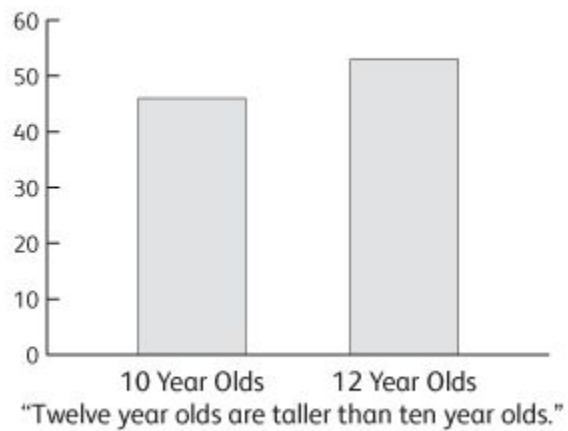
# Visu: Principios Básicos

- Simple es mejor que complejo
- Realizar una buena elección del color
- Manejo de Planos (Oclusión)
- Realizar un buen uso del plano
- 7 +/- 2
- Ceguera al cambio
- Foco en la funcionalidad

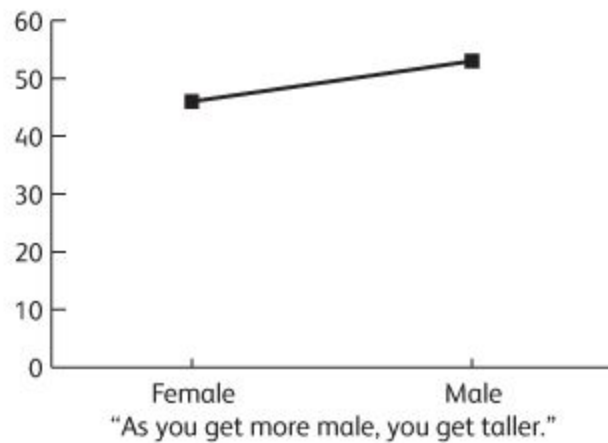
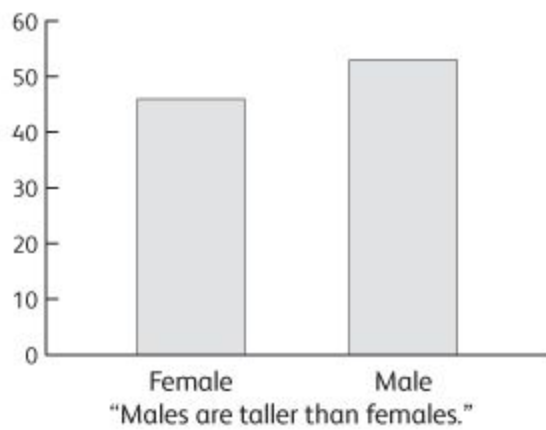
# Visualización Principios Básicos

- Diferenciar atributos continuos de categóricos
  - Gráficos de líneas vs Gráficos de barras







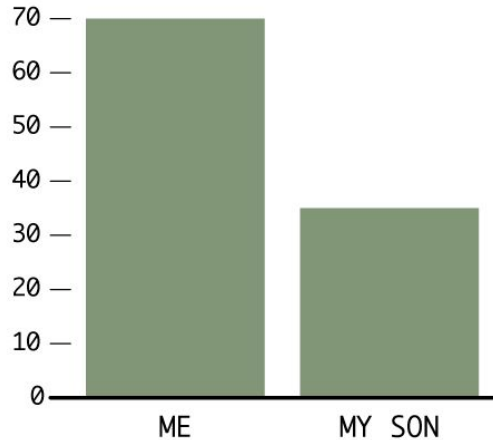




## Los Ejes Deben Comenzar en 0

Height

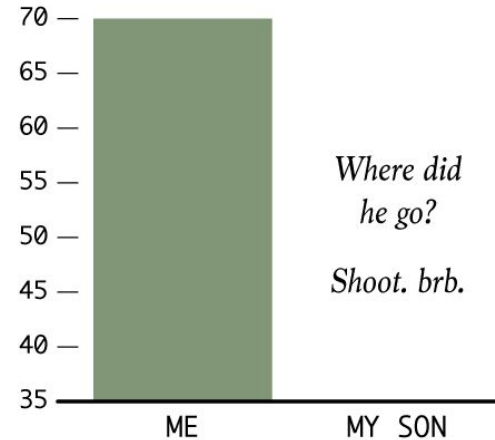
INCHES



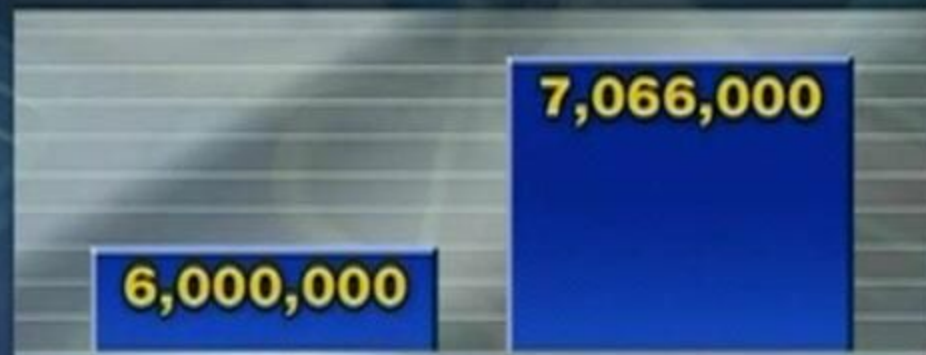
vs.

Height

INCHES



# OBAMACARE ENROLLMENT



AS OF  
MARCH 27

MARCH 31  
GOAL



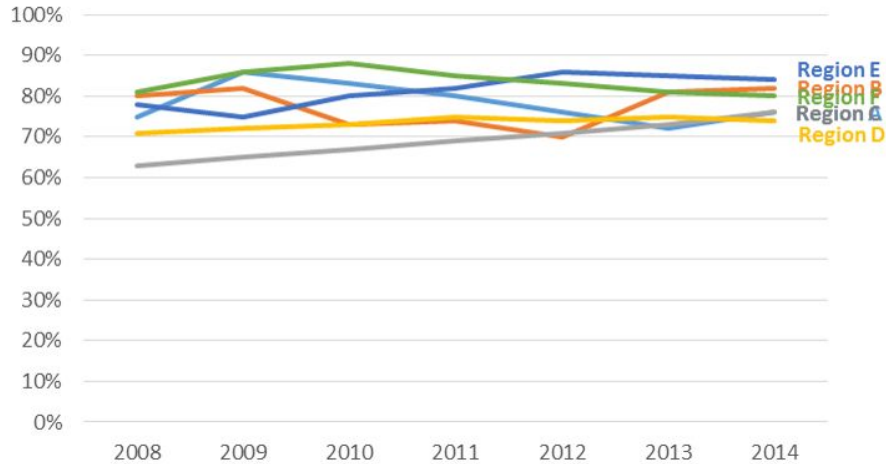
SOURCE: HHS

[mediamatters.org](http://mediamatters.org)

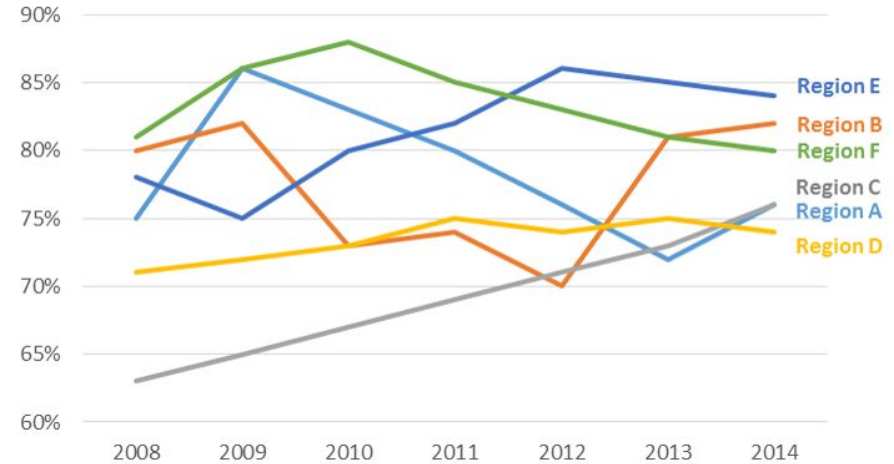
IE IN SECOND PLACE WITH \$26.5 MIL, WHILE "MUPP DOW FUT 16,325.00

# Como en todo, hay excepciones...

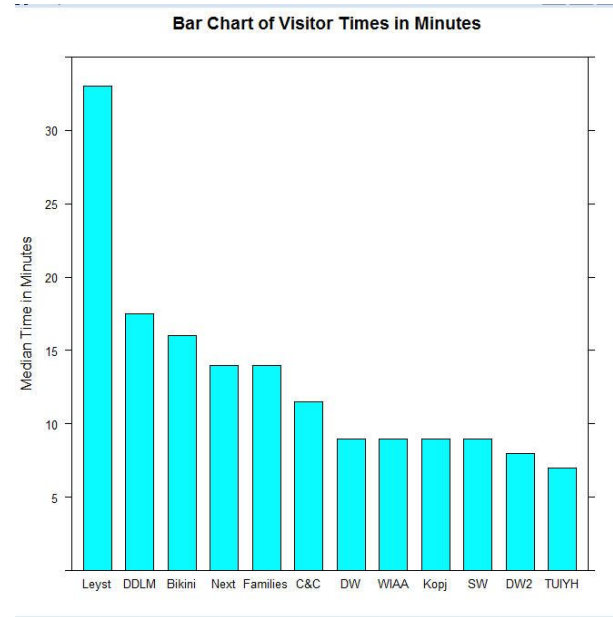
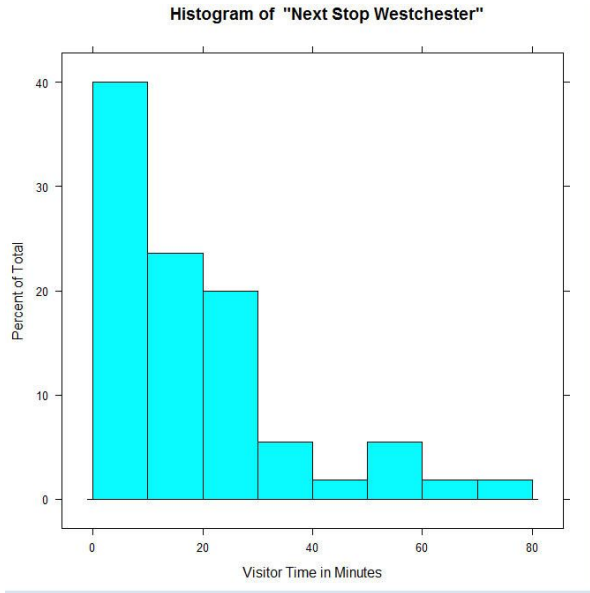
Region progress since 2008



Region progress since 2008



# Histograma vs Bar Chart





# Histograma vs Bar Chart

## Histograma

- Muestran la distribución de una variable.
- Eje x: Variable numérica discretizada en buckets (bins)
- Eje y: Cantidad / Porcentaje
- No se puede reordenar el eje X
- No hay espacios entre la barras

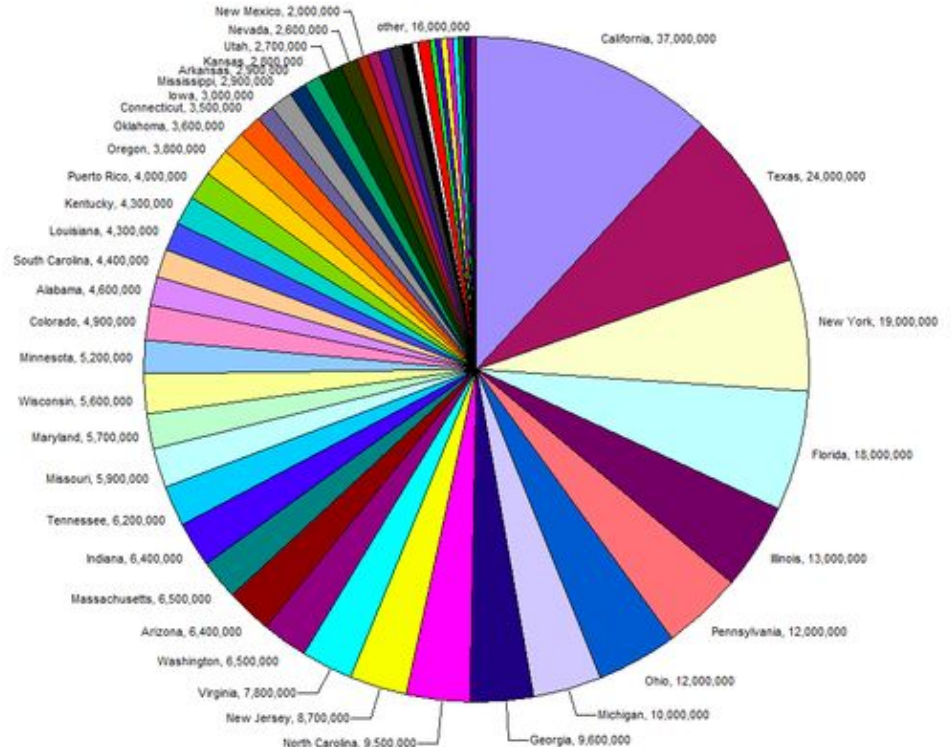
## Bar Chart

- Comparación entre variables
- Eje x: Variable categórica
- Eje y: Variable numérica
- El eje x puede tener cualquier orden
- Barras espaciadas

# Pie charts

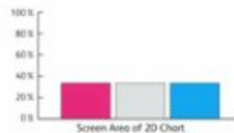
Por qué usarlos con cuidado...

- Con muchos datos se vuelve difícil (o imposible) de interpretar



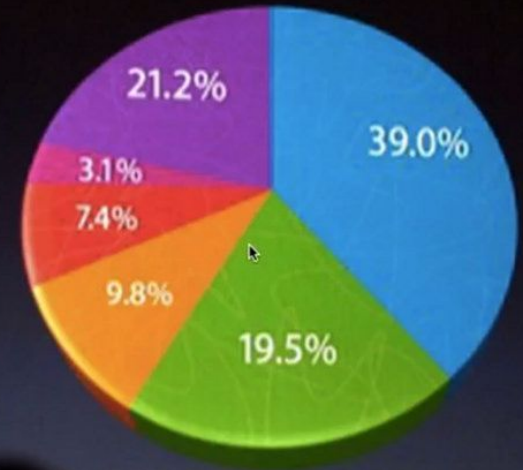
# Pie charts

- Cuando incluimos 3D modificamos las áreas, cambiando las proporciones



## U.S. SmartPhone Marketshare

● RIM  
● Apple  
● Palm  
● Motorola  
● Nokia  
● Other

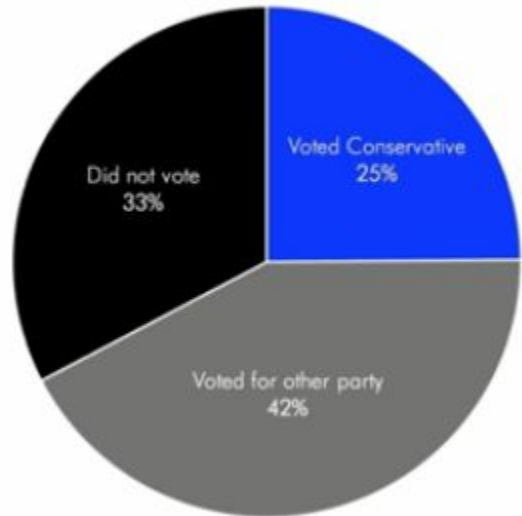




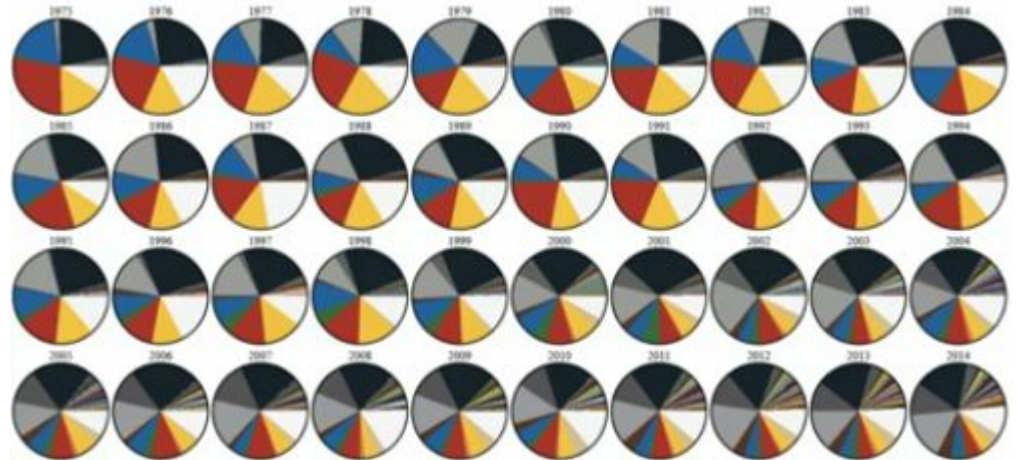
# Pero como todo, tienen una función

Para mostrar una distribución  
entre 2 o 3 valores

Summary of eligible voters in the UK General Election 2015



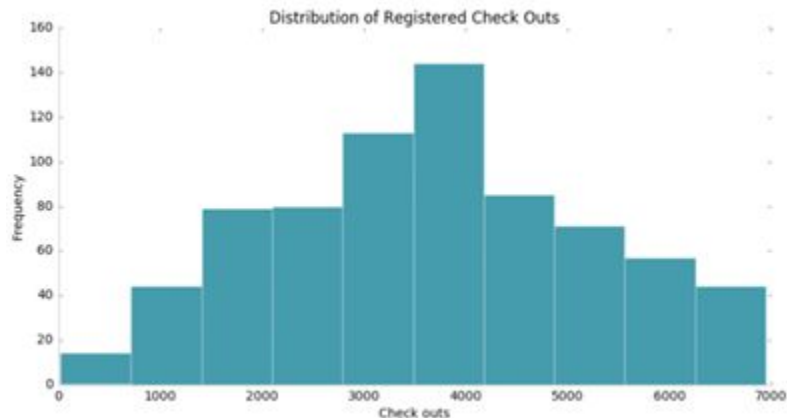
Para mostrar la evolución de un  
set de valores



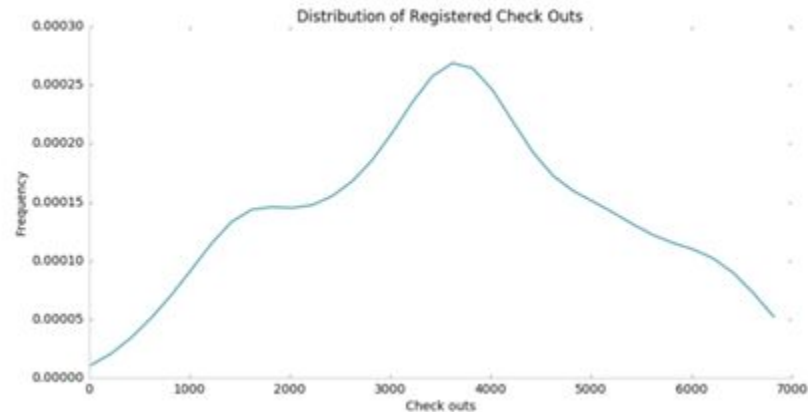
# Algunas visualizaciones comunes

Para analizar la distribución de una variable:

Histograma



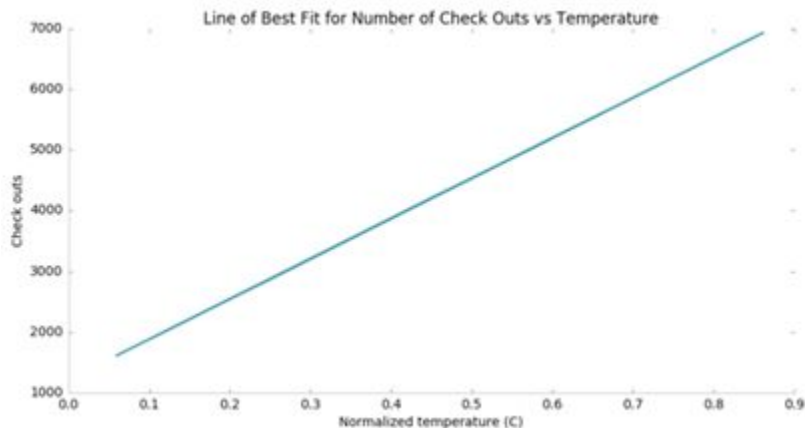
Density Plot



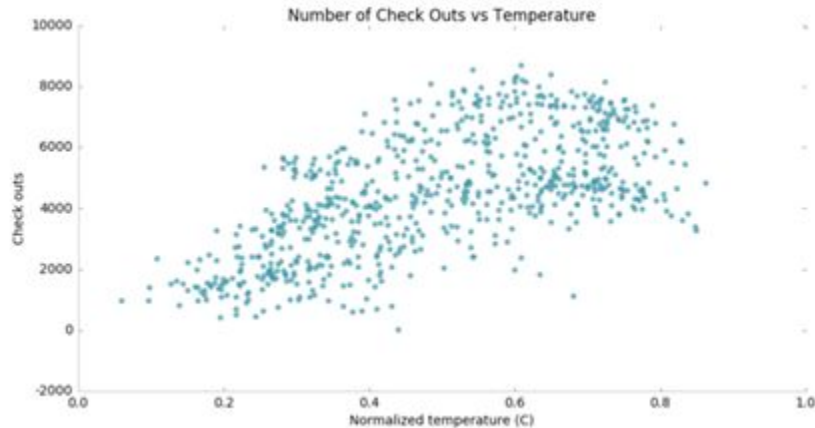
# Algunas visualizaciones comunes

Para explorar la relación entre distintas variables:

## Line Plot



## Scatter Plot

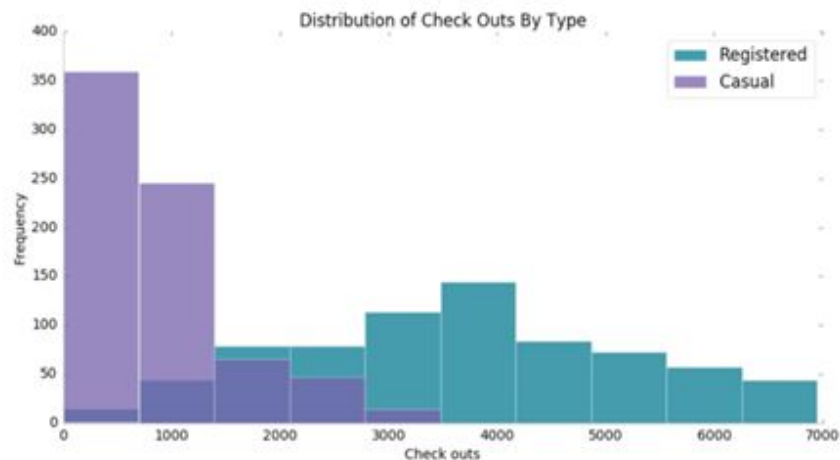
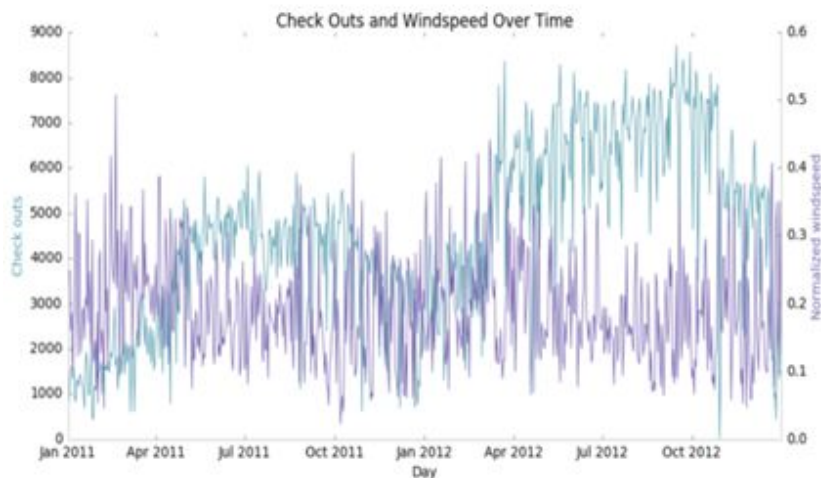


# Algunas visualizaciones comunes

Para explorar la relación entre distintas variables:

Line Plot con dos ejes y

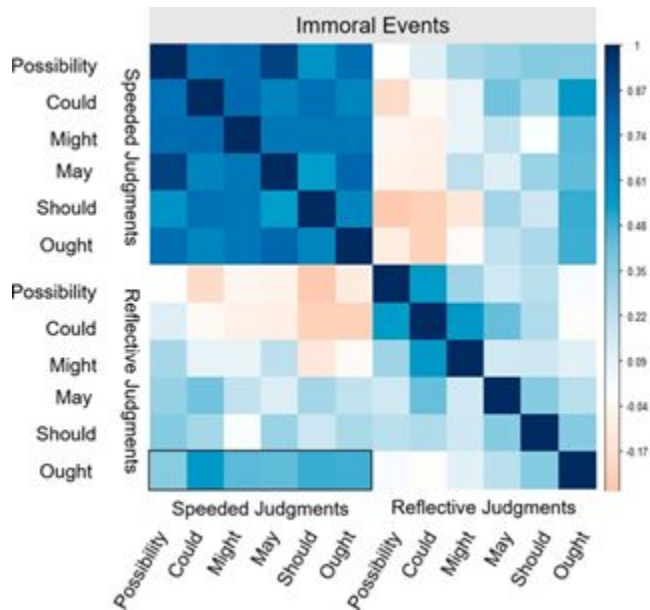
Histogramas superpuestos



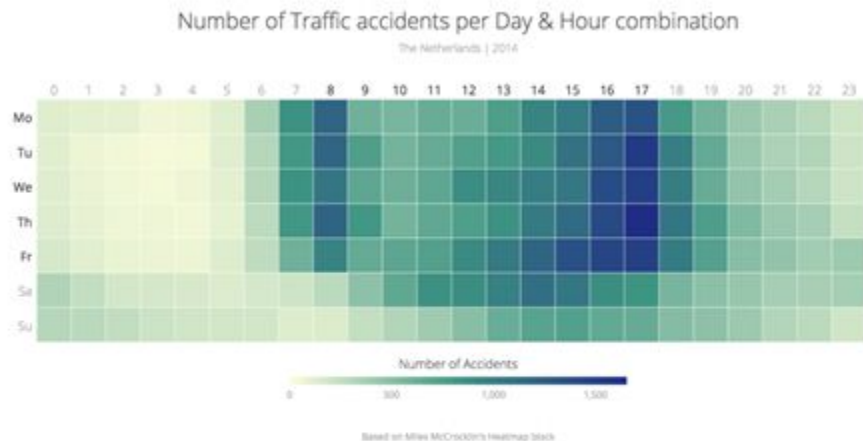
# Algunas visualizaciones comunes

Para explorar la relación entre distintas variables:

## Matriz de correlación



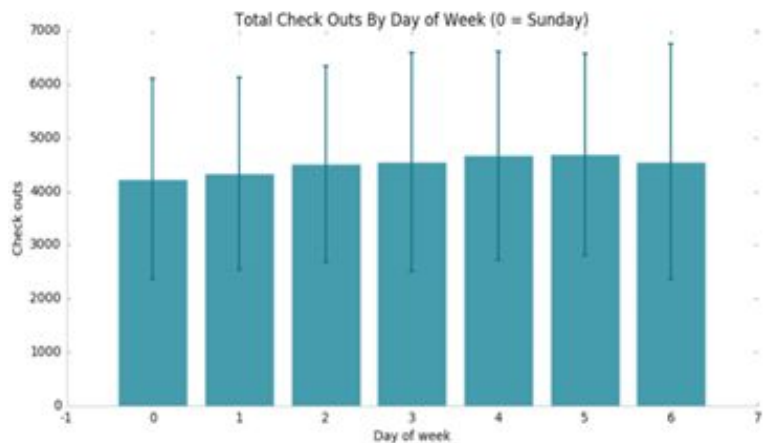
## Heat Map



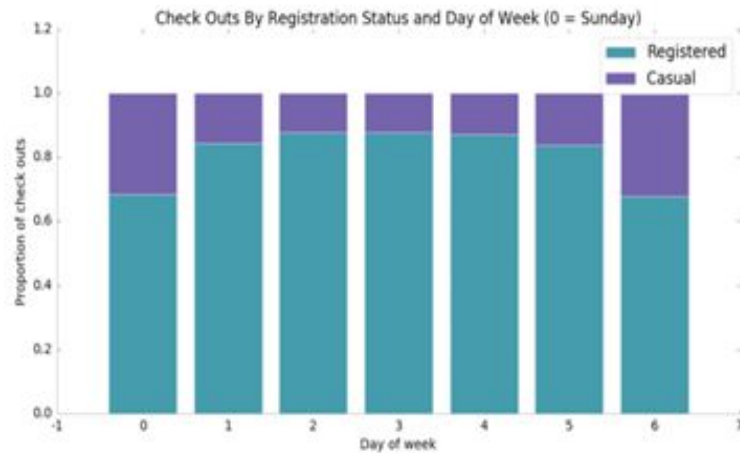
# Algunas visualizaciones comunes

Comparando grupos o categorías:

Bar Plot



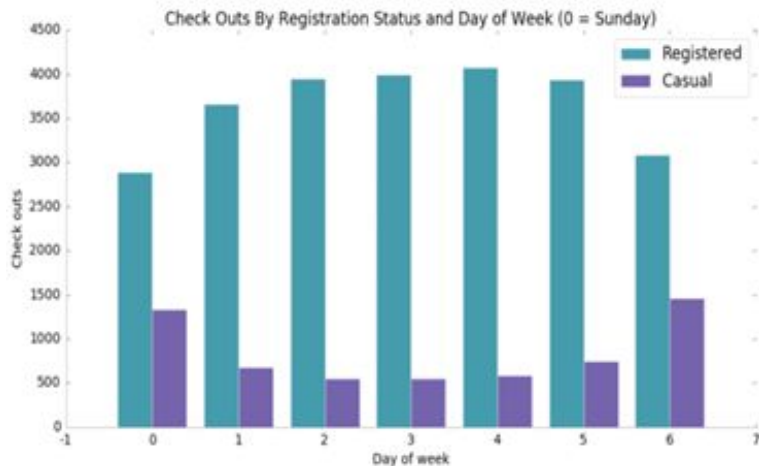
Stacked Bar Plot



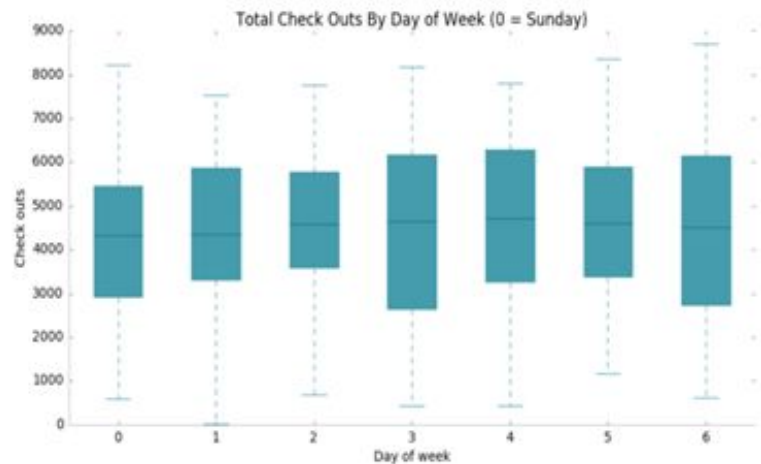
# Algunas visualizaciones comunes

Comparando grupos o categorías:

Grouped Bar Plot

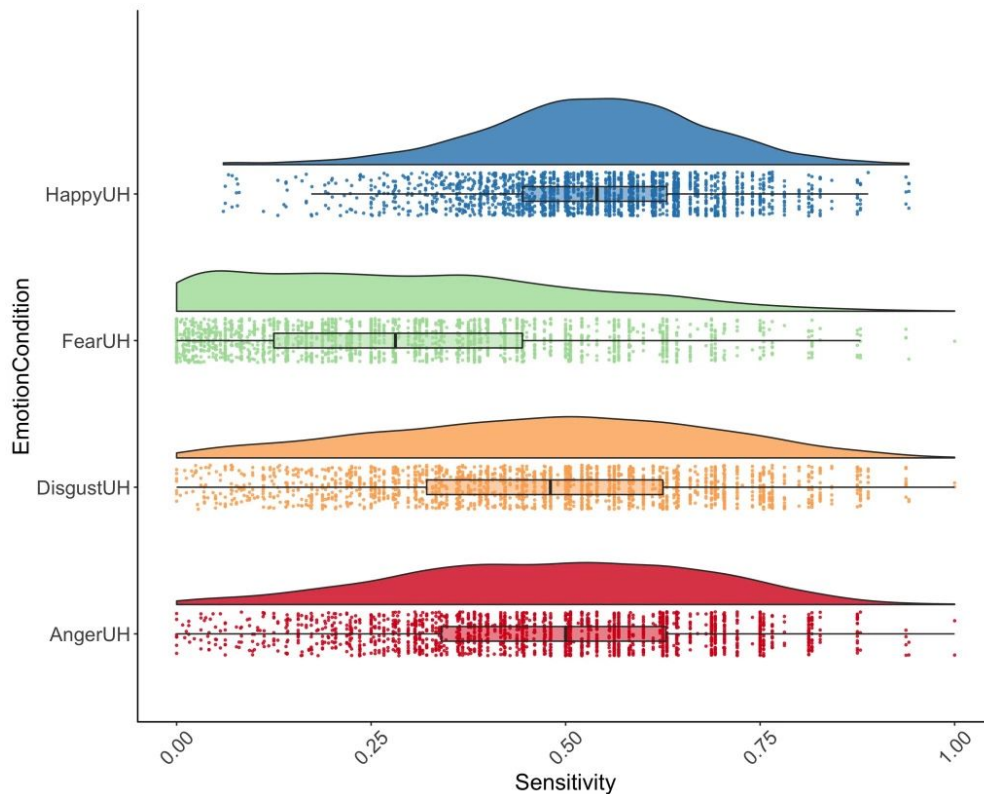


Box Plot



# Explorando otras opciones

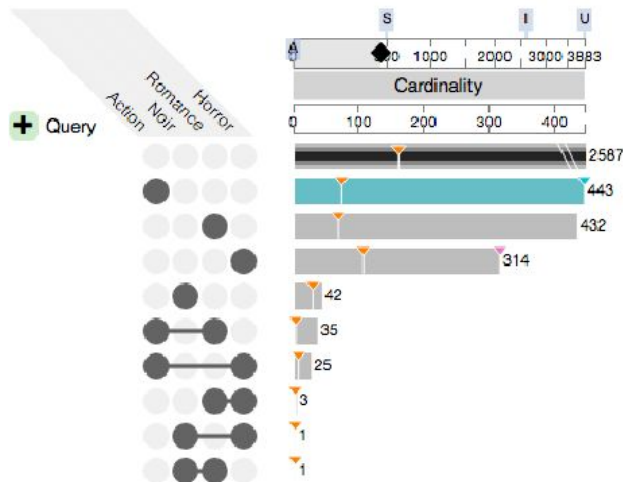
## Raincloud Plot



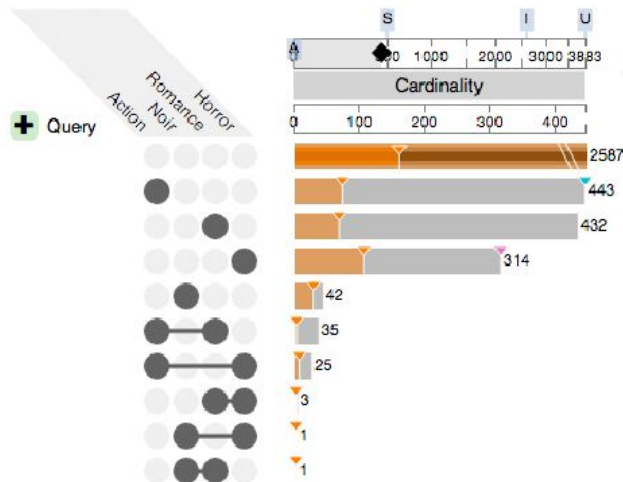


# Explorando otras opciones

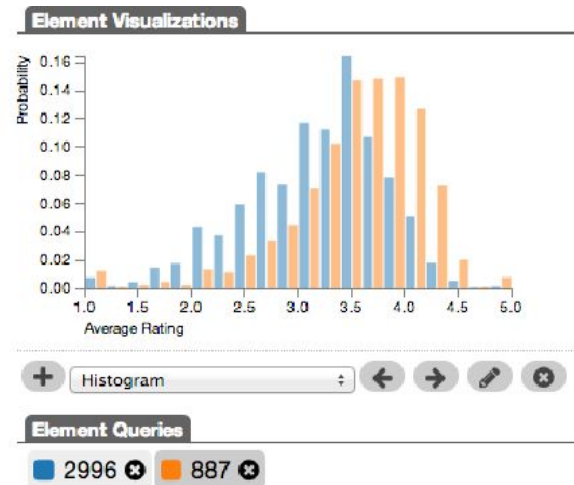
## UpSet



(a) Element query in set space



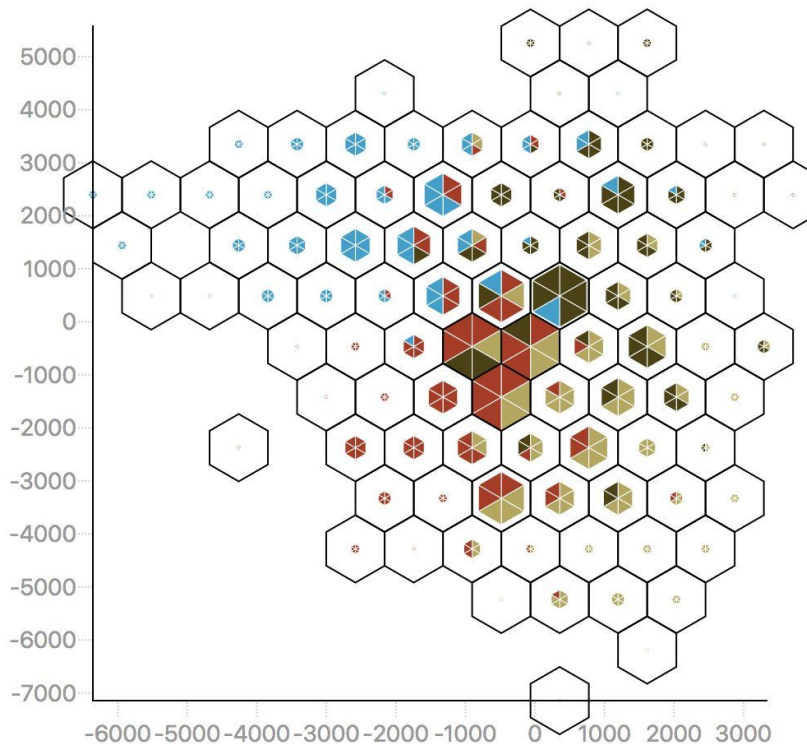
(b) Element query in element space



(c) Element visualization

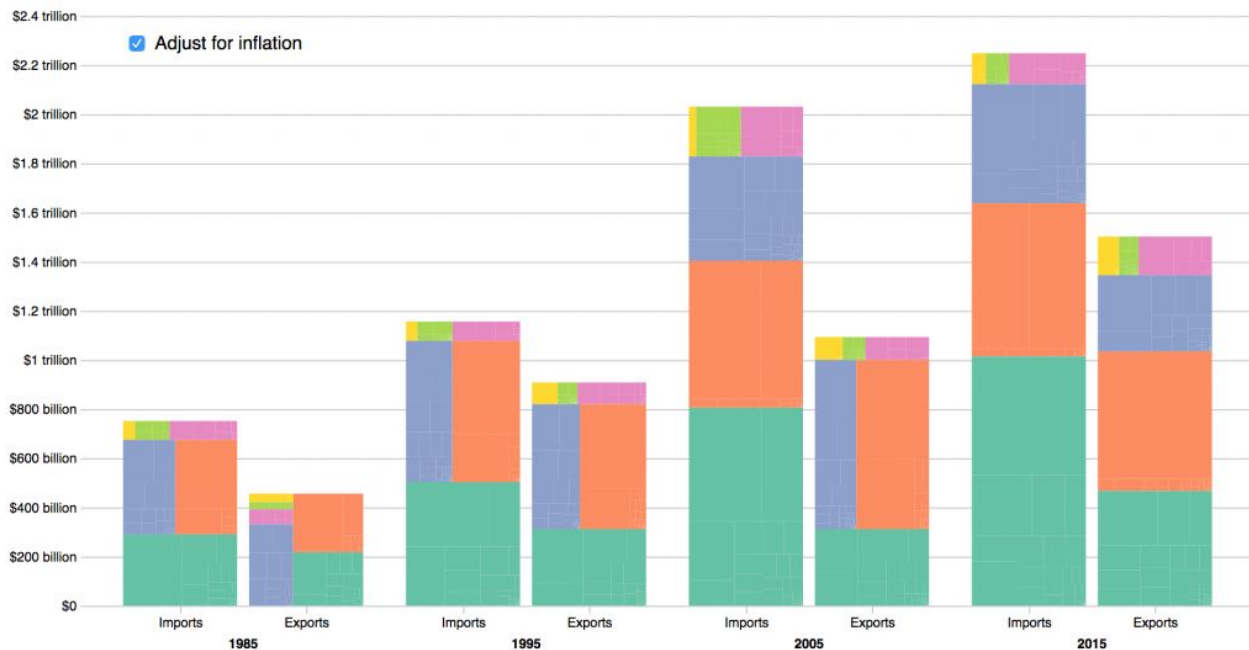
# Explorando otras opciones

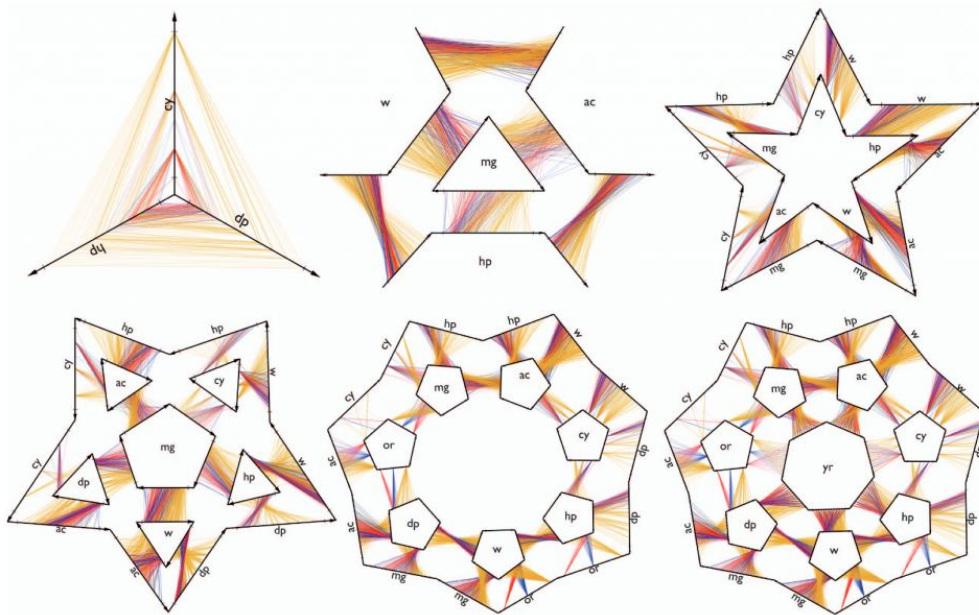
## Multi-Class Hexbin



# Explorando otras opciones

## Tree map bar chart





# Best in Show

The ultimate data-dog

INTELLIGENCE

dumb clever sml med lrg

TYPE

Herding Hound Non-sporting Sporting Terrier Toy Working

public popularity

Inexplicably overrated



Bulldog



Hot dog!



OUR DATA SCORE

intelligence + longevity + ailments + costs - grooming - appetite = data score

Rightly ignored

Overlooked treasures

our data score

# What is the world's biggest cash crop?

## MOST PLANTED

Area Harvested (million km<sup>2</sup>)

## MOST FECUND

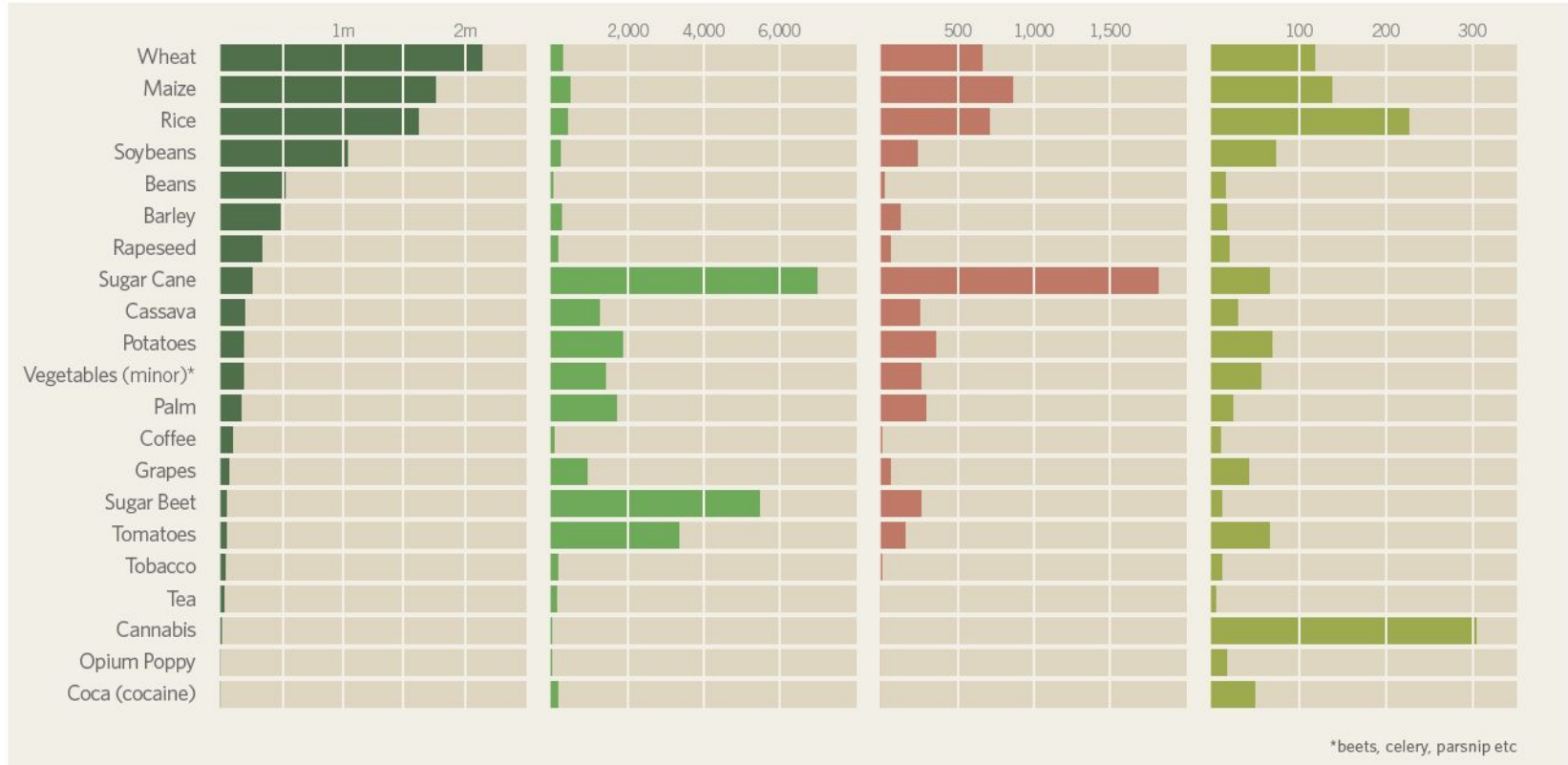
Yield per km<sup>2</sup> (tonnes)

## MOST POPULAR

Gross yearly production (mtonnes)

## MOST REVENUE

Gross production value (\$bn)

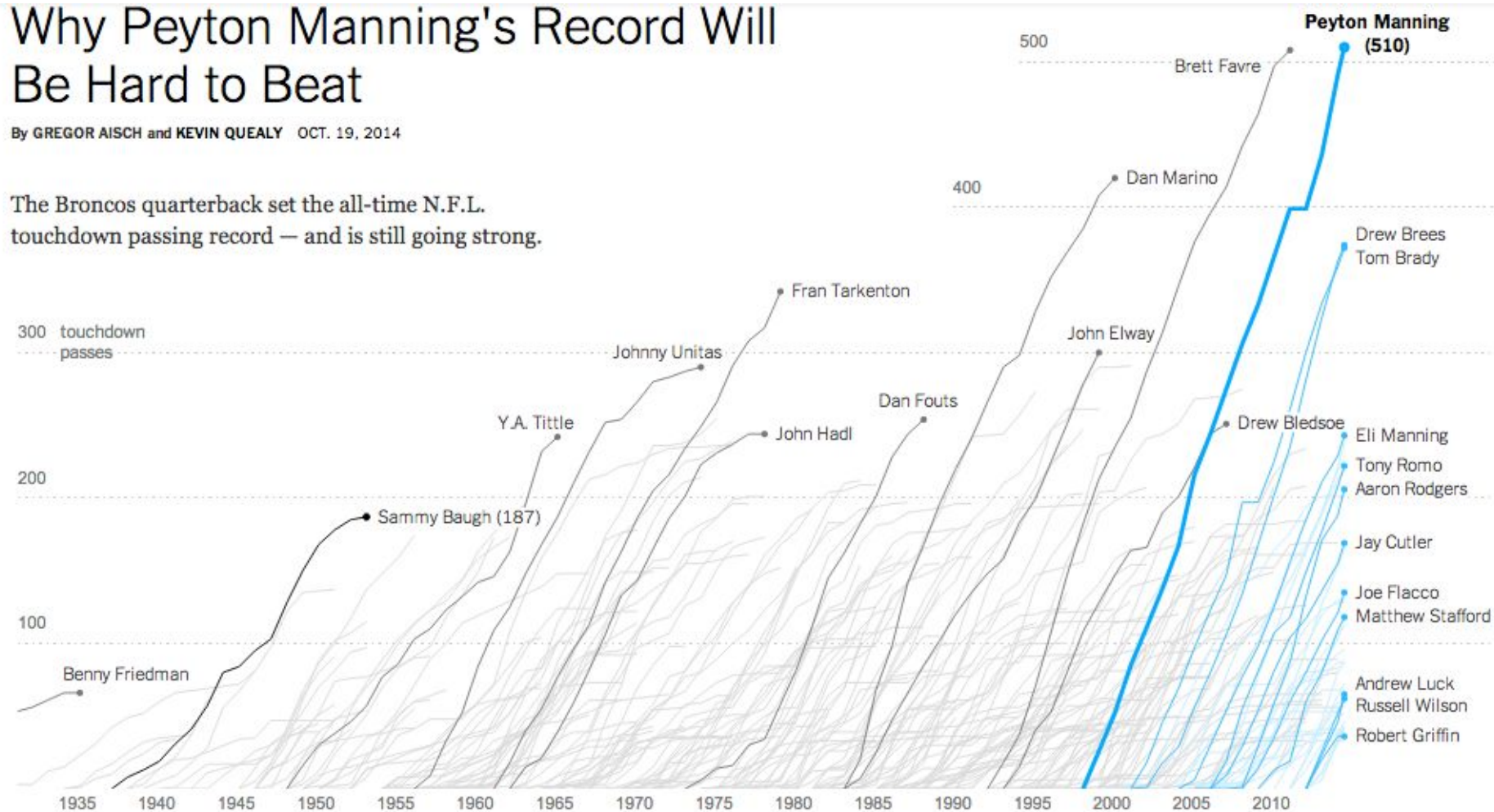




# Why Peyton Manning's Record Will Be Hard to Beat

By GREGOR AISCH and KEVIN QUEALY OCT. 19, 2014

The Broncos quarterback set the all-time N.F.L. touchdown passing record — and is still going strong.

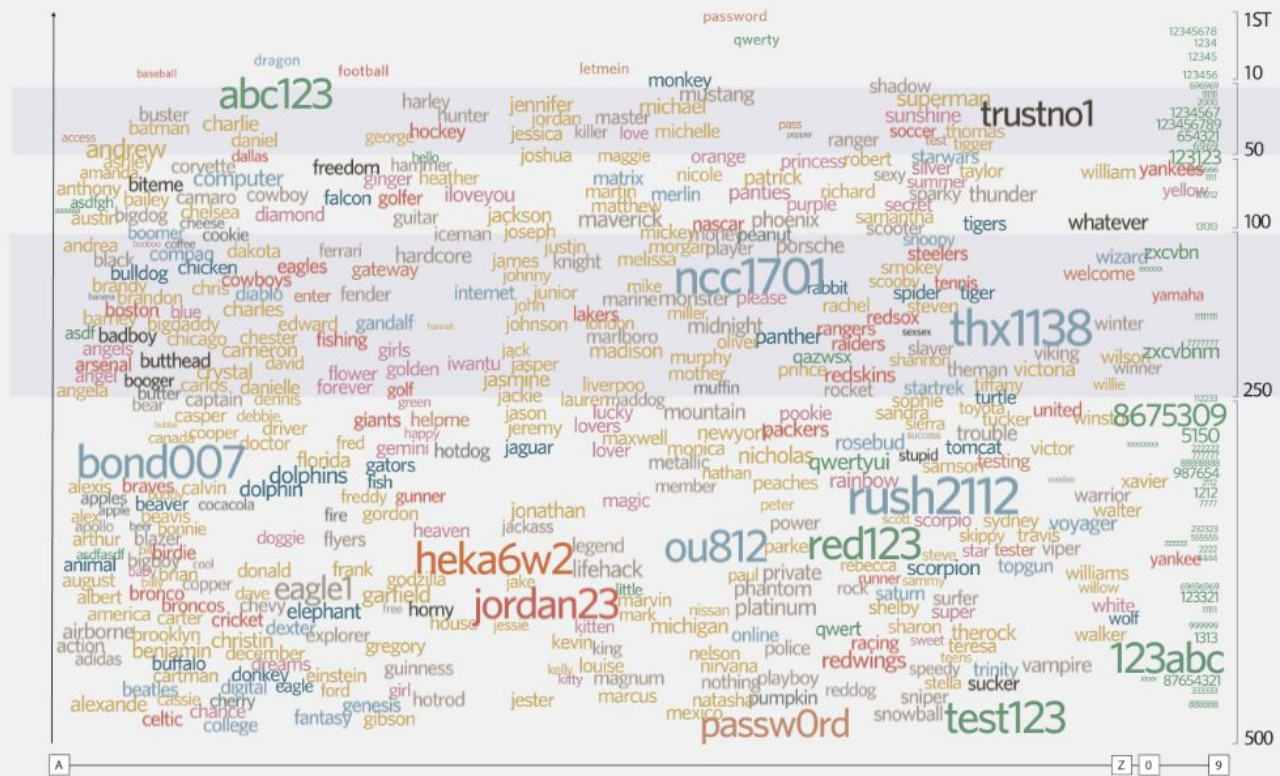


Is yours here?

Clear All

Rebellious/Rude Simple Alphanumeric Sport Truly Random

RANK



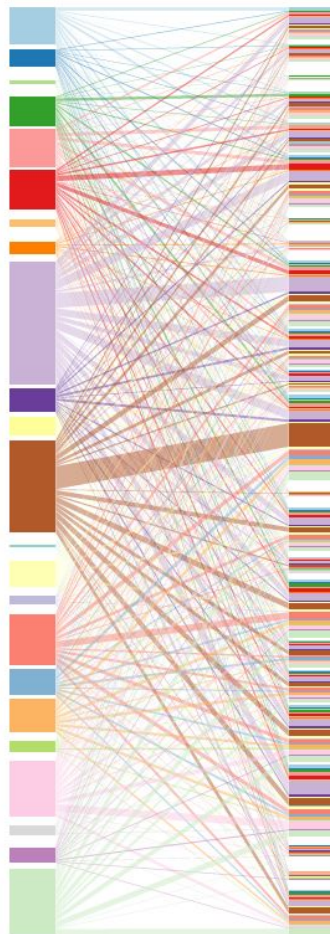


# Father-son occupation pairs

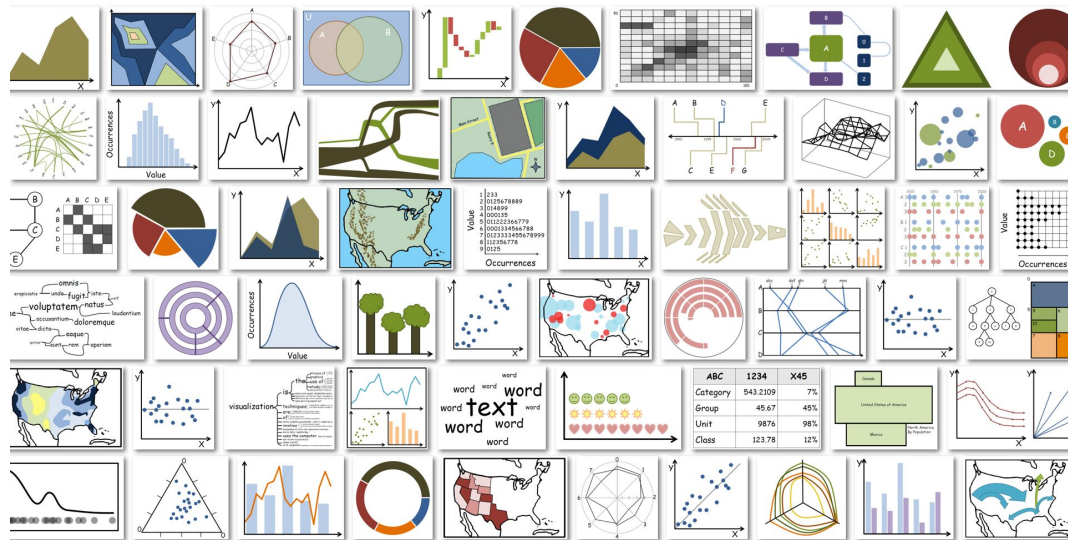
Father

Son

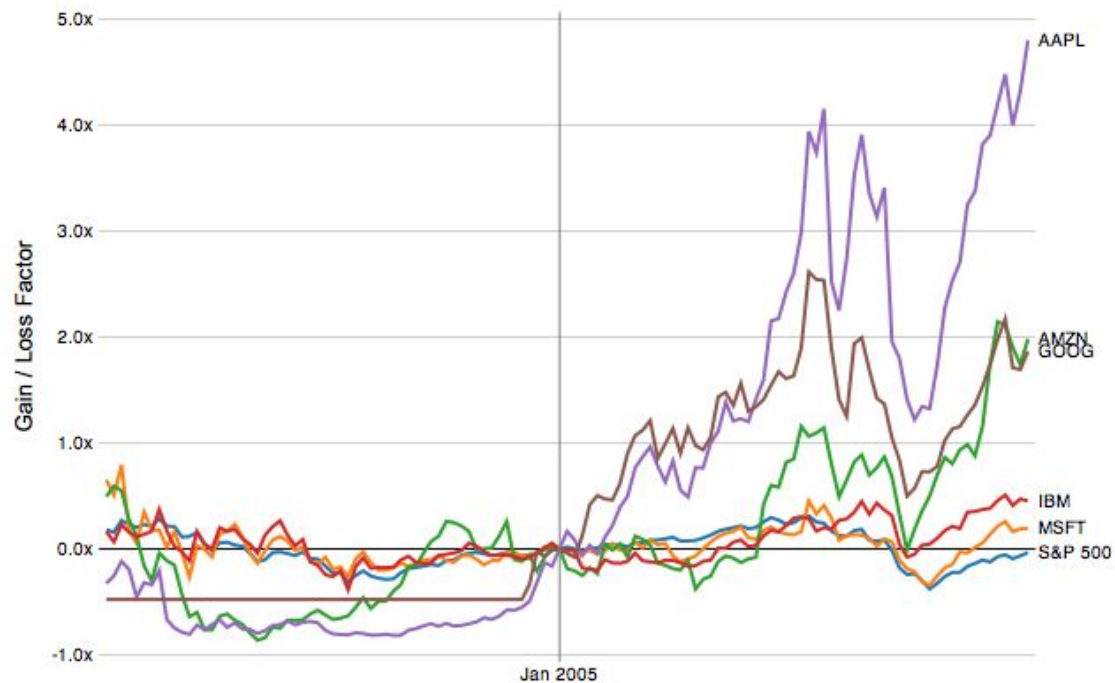
architecture\_engineering  
business\_finance  
cleaning\_maintenance  
construction  
education  
entertainment  
farming  
legal  
  
management  
  
math\_CS  
medicine  
  
military  
  
nursing  
office  
personal\_care  
production  
protective\_service  
repair  
restaurant  
sales  
science  
social\_service  
transportation



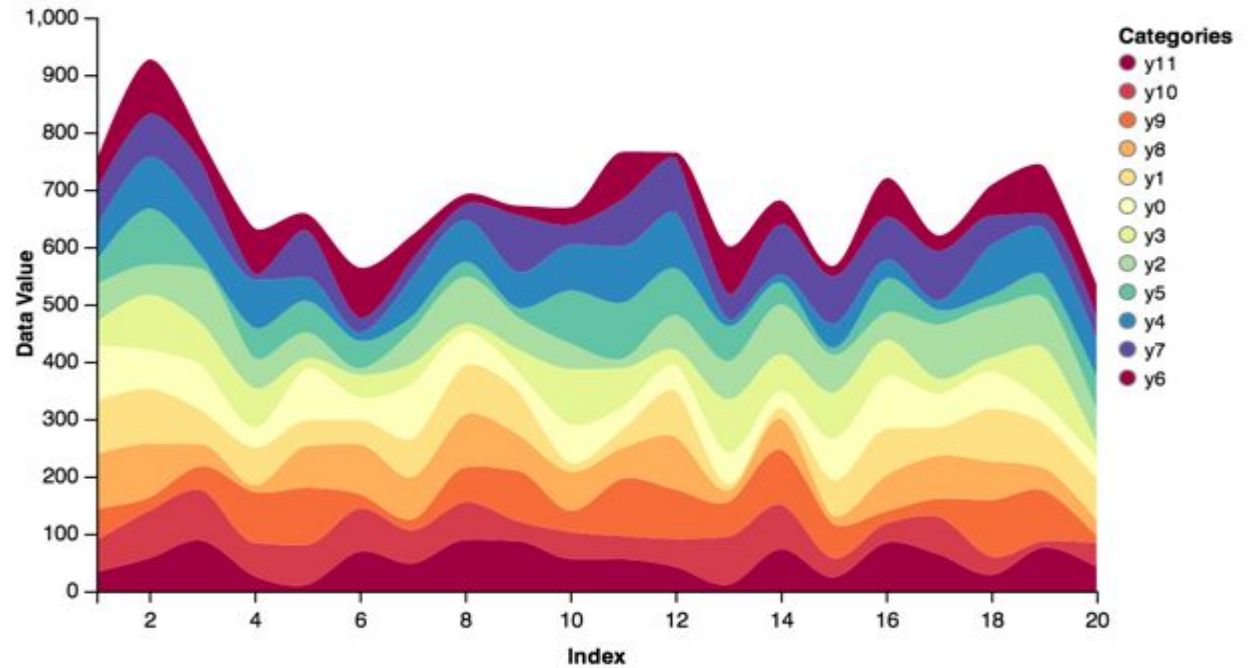
architecture\_engineering  
business\_finance  
cleaning\_maintenance  
construction  
education  
entertainment  
farming  
legal  
  
management  
  
math\_CS  
medicine  
  
military  
  
nursing  
office  
personal\_care  
production  
protective\_service  
repair  
restaurant  
sales  
science  
social\_service  
transportation



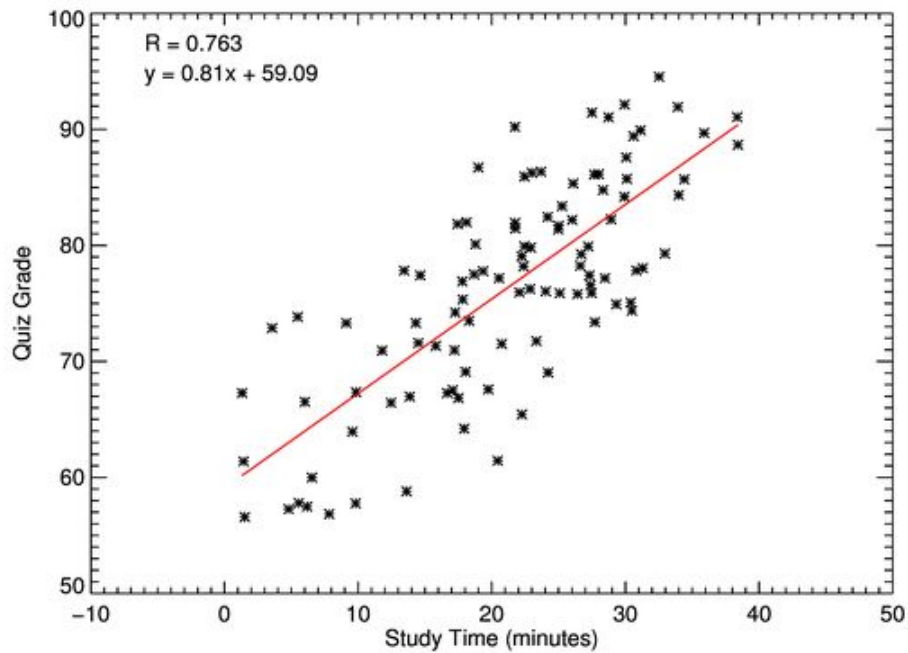
# Line Chart



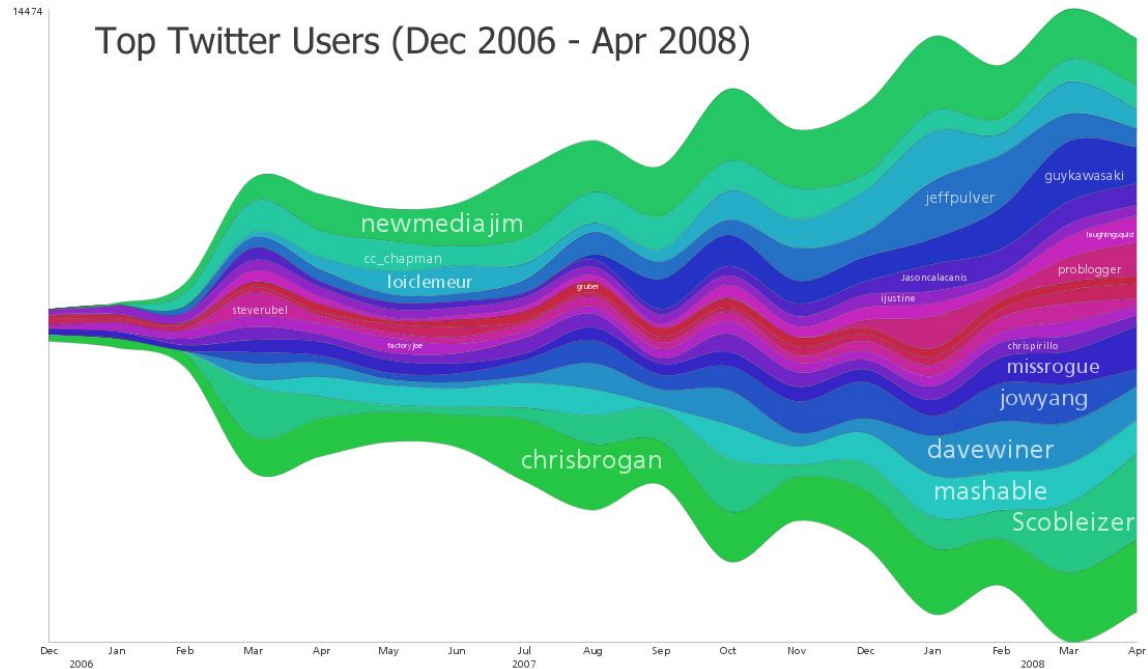
# Stacked Area



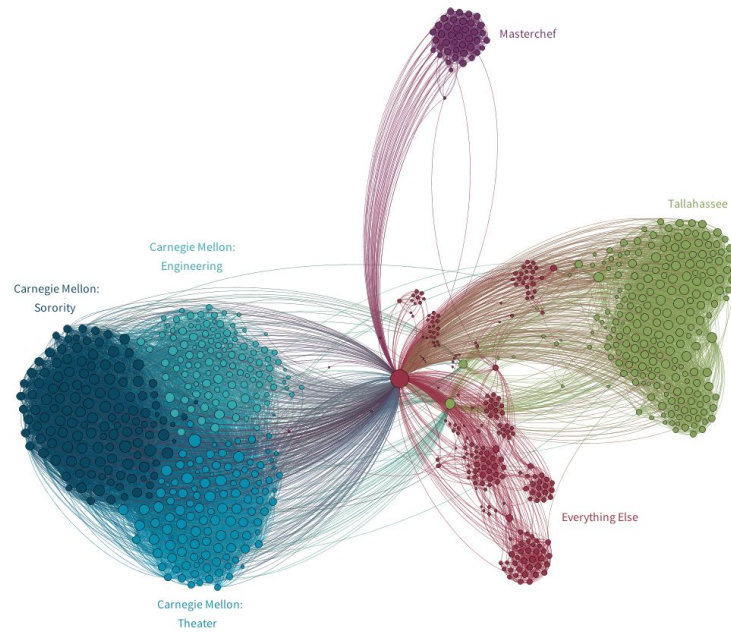
# Scatter



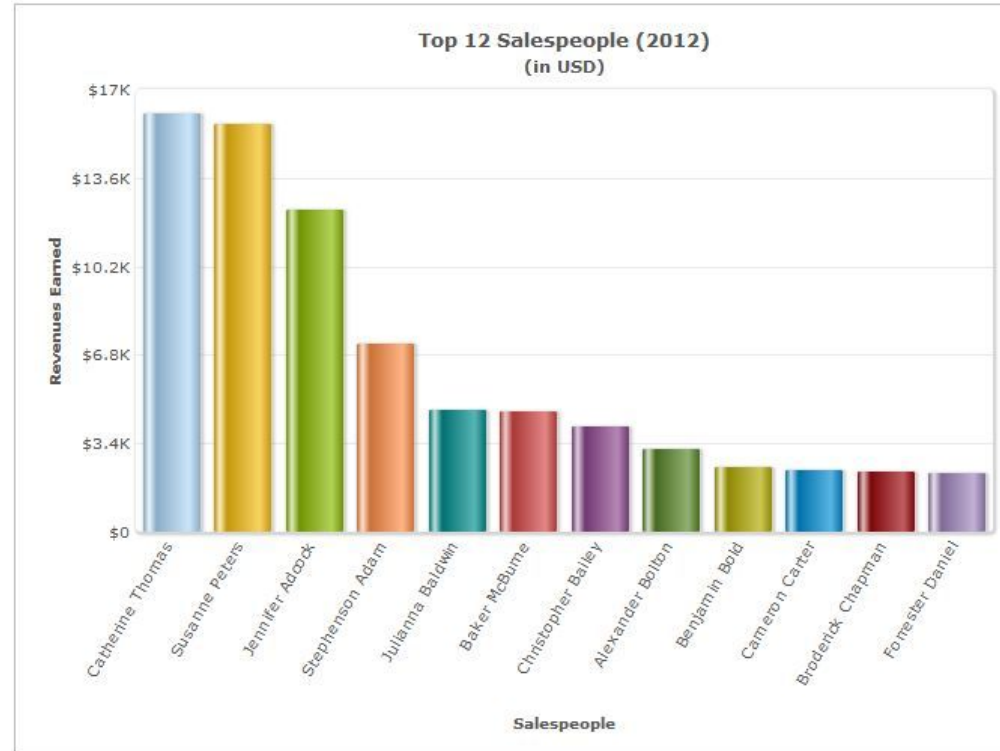
# Streamgraph



# Network Graph

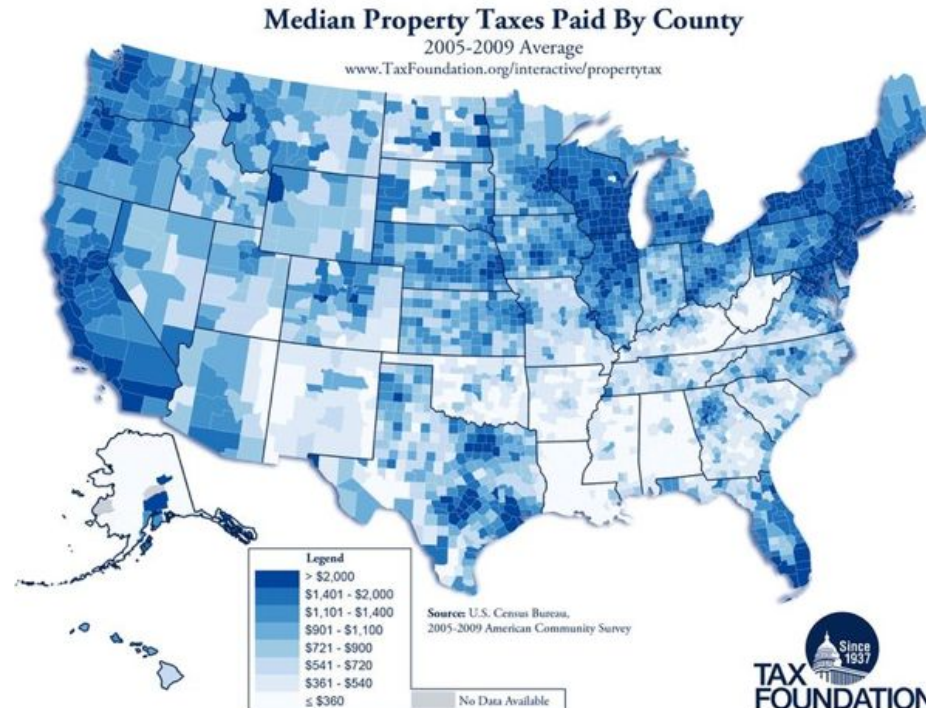


# Bar Chart

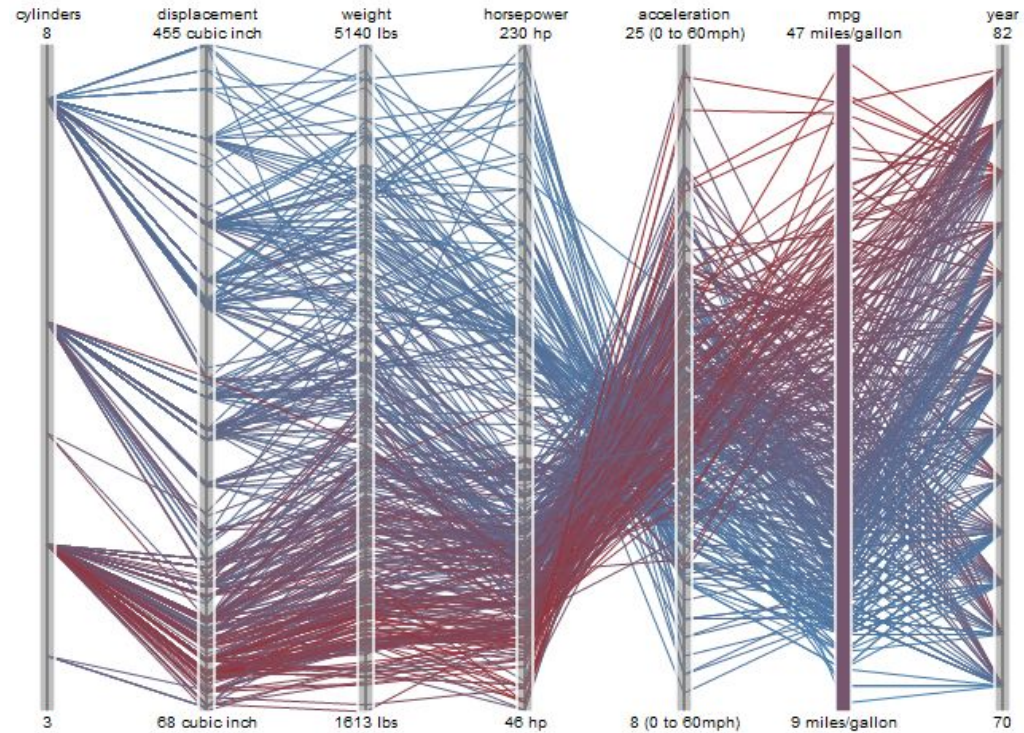




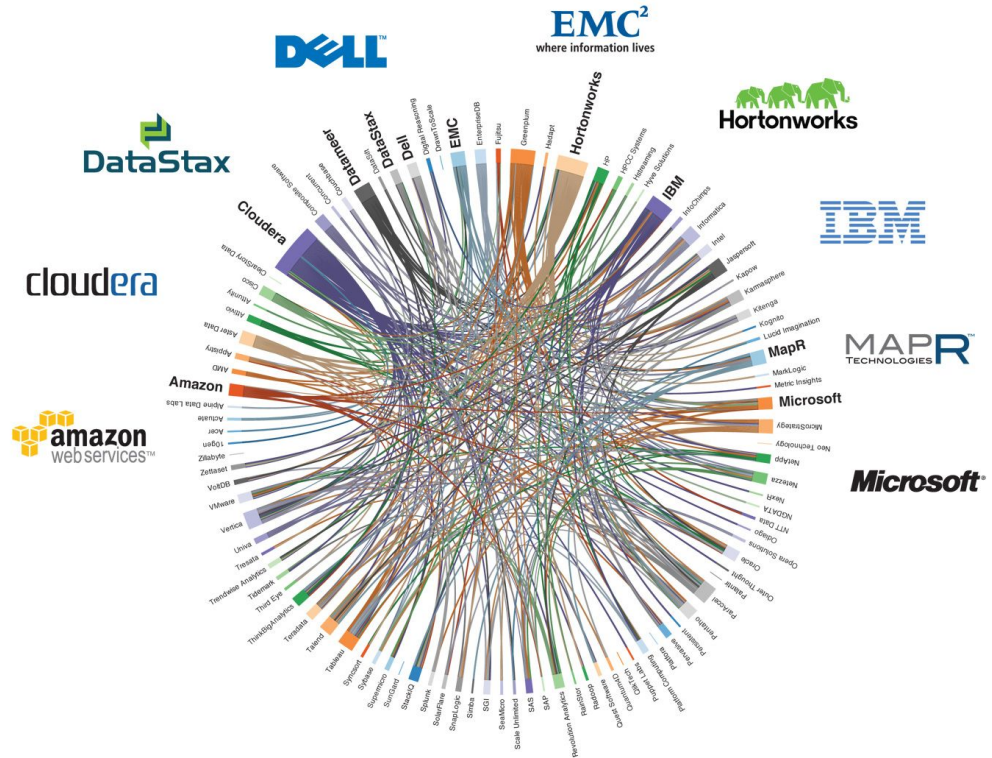
# Choropleth



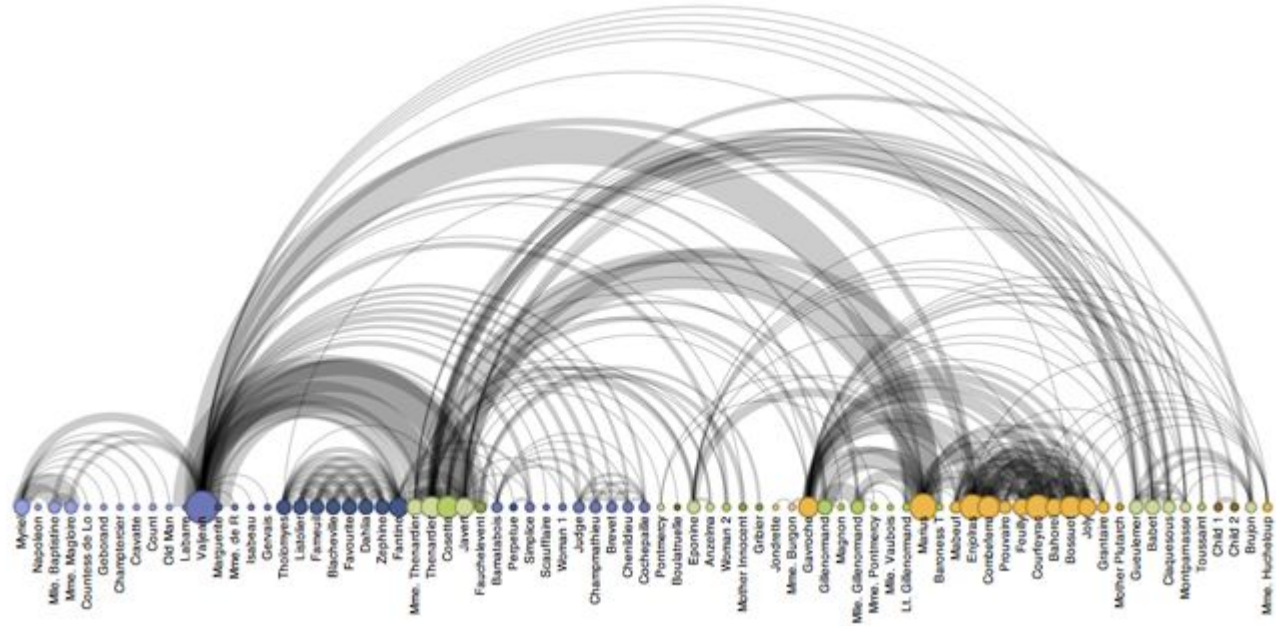
# Paralell Coordinates



# Radial Cluster

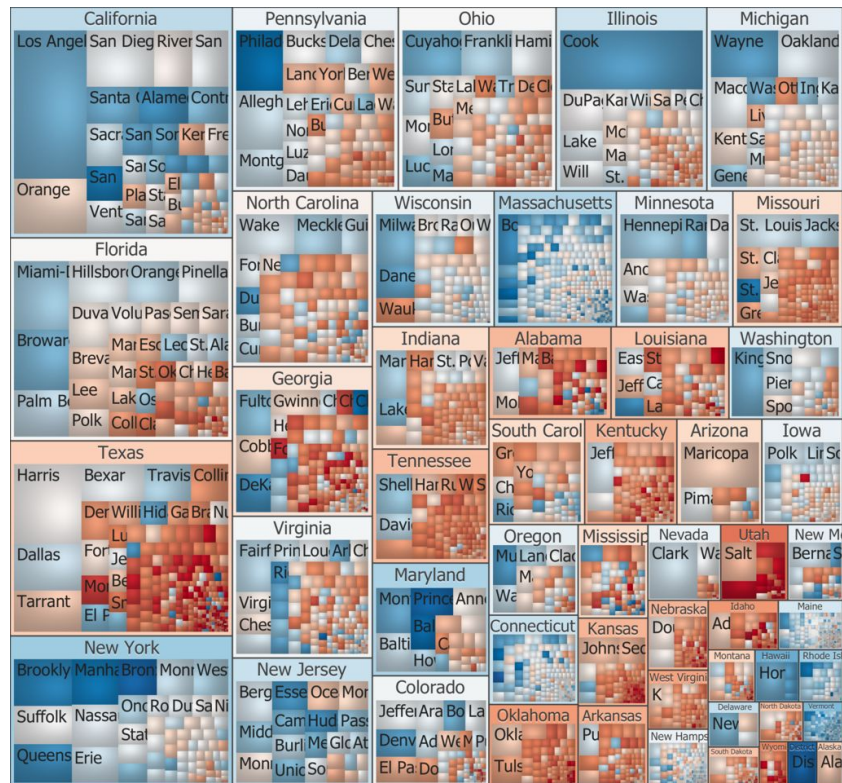


# Arc Diagram

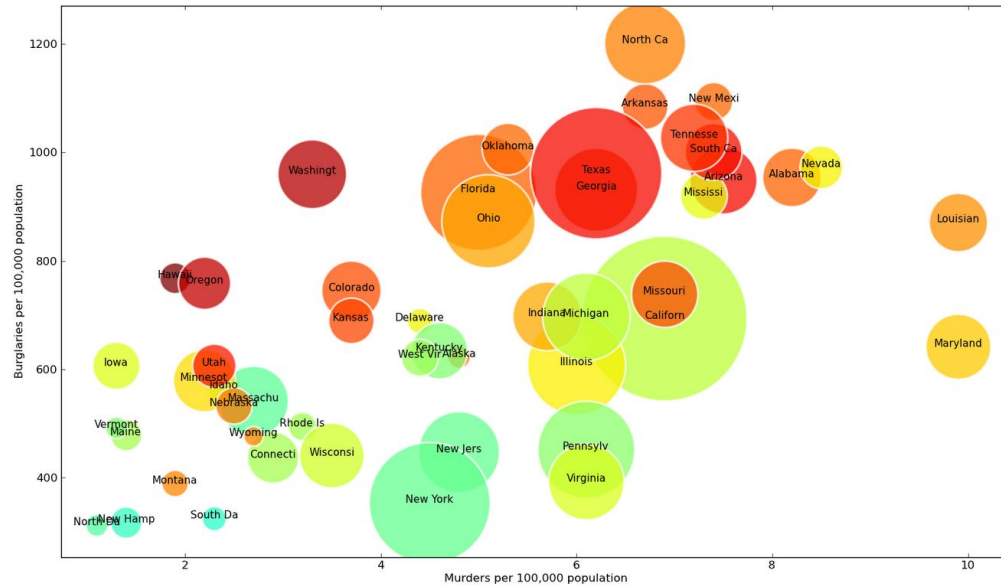


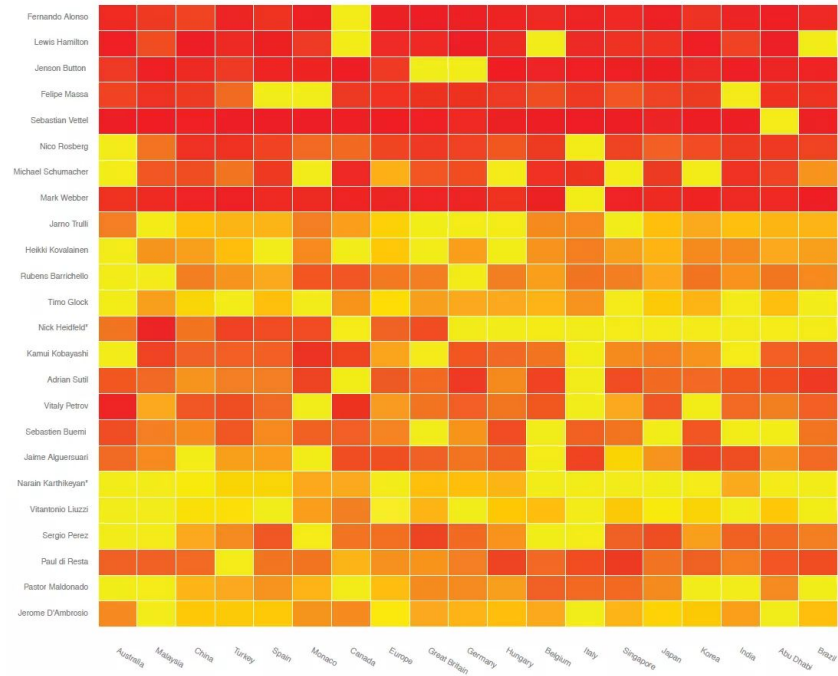






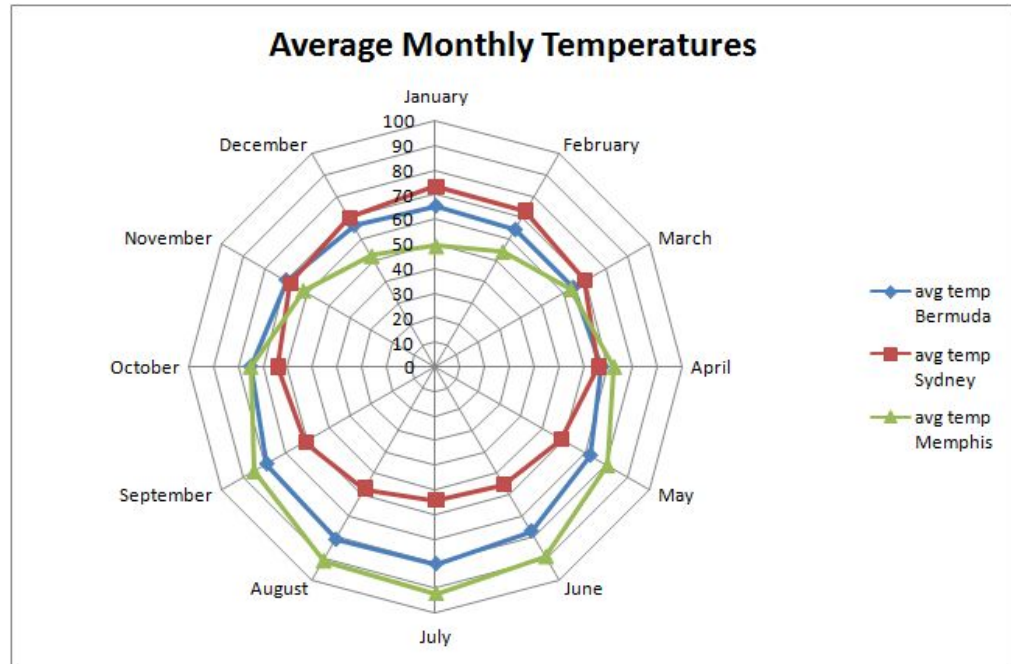
# Bubble Charts



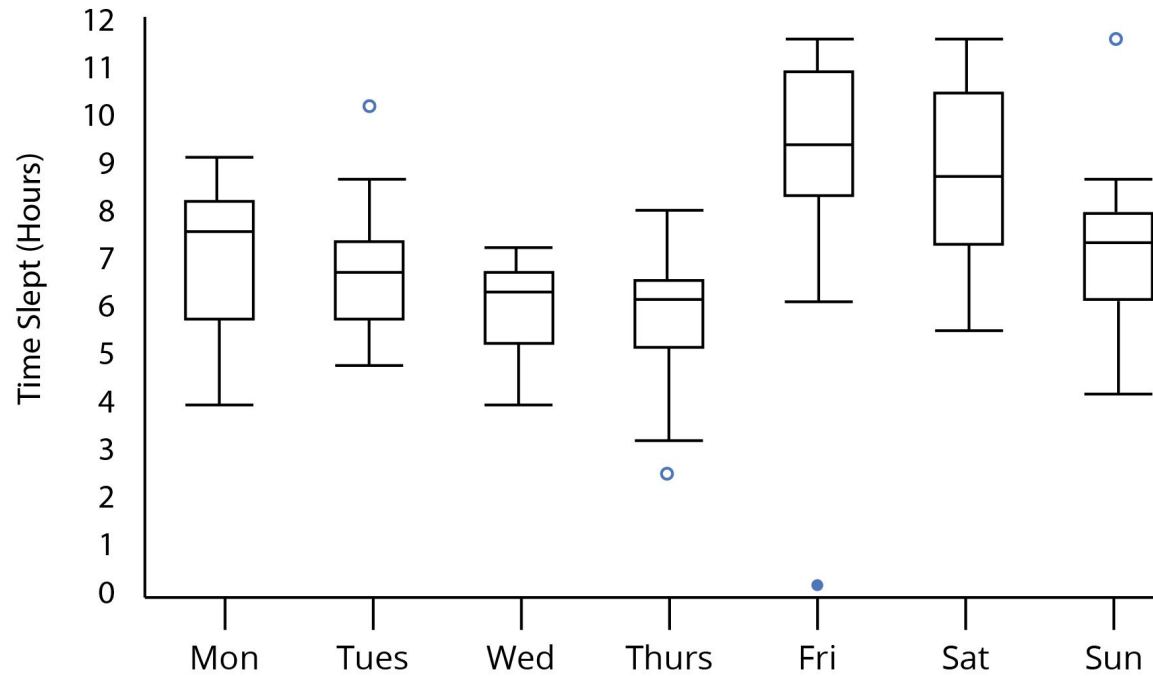




# Radar Chart



# Box Plot



# Violin Plot

