91 SPRINGBOARD
cowork . network . grow

School of AI
Hyderabad

Data
Aesthetics

1st June 2019 | 10am - 1pm

Venue : 91Springboard,
Hitech City

presented by

Srihari Pramod

Ex Lead - 3D Visualization and Technical
MYOS Property.com

“The greatest value of a picture is when it forces us to notice what we never expected to see”

- John Tukey

Data Visualization

Applications:

- Exploratory Data Analysis
  - Understanding Distributions of data
  - Understanding correlations
- Data Mining – Finding hidden insights
- Presenting the results

Dataset

	Name	Subject	Score	score_ranked	group_rank
0	Alisa	Mathematics	62	6.0	1.0
1	Bobby	Mathematics	47	8.0	3.0
2	Cathrine	Mathematics	55	7.0	2.0
3	Alisa	Science	74	4.0	2.0
4	Bobby	Science	31	10.0	3.0
5	Cathrine	Science	77	3.0	1.0
6	Alisa	History	85	2.0	1.0
7	Bobby	History	63	5.0	2.0
8	Cathrine	History	42	9.0	3.0
9	Alisa	Economics	62	6.0	3.0
10	Bobby	Economics	89	1.0	1.0
11	Cathrine	Economics	85	2.0	2.0

Data Types :

- Numeric (Continuous | Discrete)
- Categorical (Binary | Multi class)
- Nominal, Ordinal etc.

Classification of Charts based on datatypes

Categorical vs Numerical	Numerical Distribution	Numerical vs Numerical
• Bar • Line • Area • Pie • Geographical Map	• Box • Histogram	• Scatter

Classification of Charts based on dimensions

1-D

Numerical Distribution
- Box
- Histogram

2-D

Categorical vs Numerical
- Bar
- Line
- Area
- Pie
- Geographical Map

Numerical vs Numerical
- Scatter

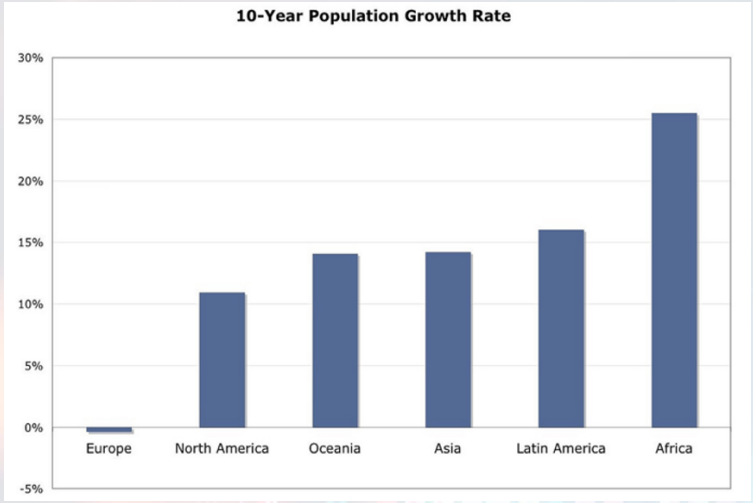
3-D

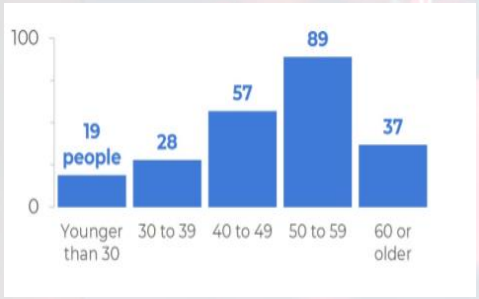
Numerical as a function of two (or more) categorical
- Heat Map
- Tree Map

Three Numerical
- Surface
- 3D scatter

Bar Charts/ Column Charts

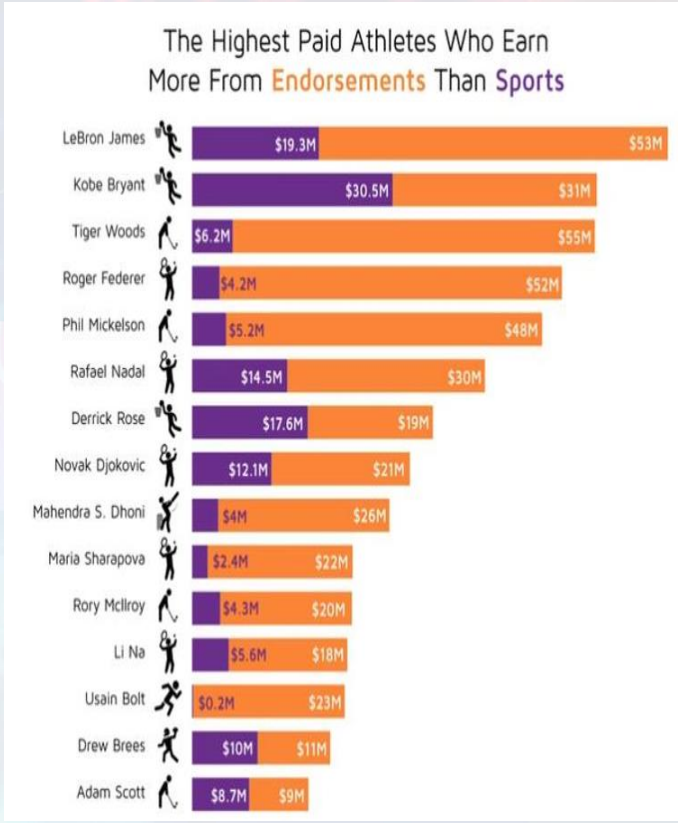
Independent Variable – Categorical
Dependent Variable - Numeric





Bins

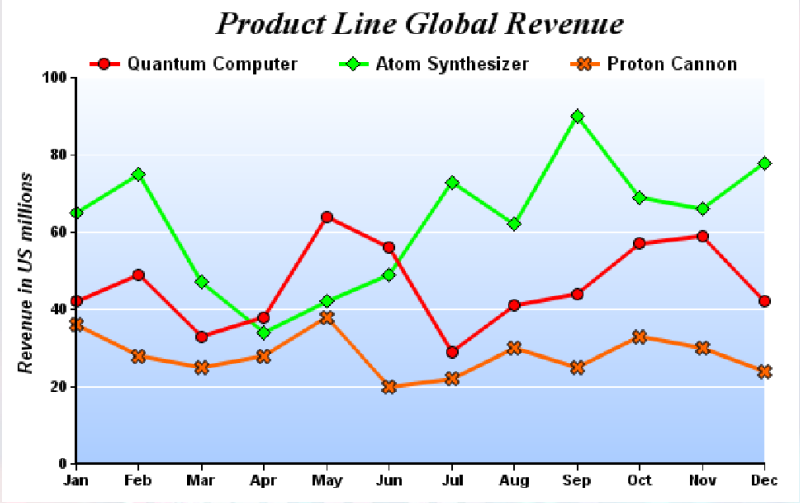
Grouped



Horizontal Bars (Stacked)

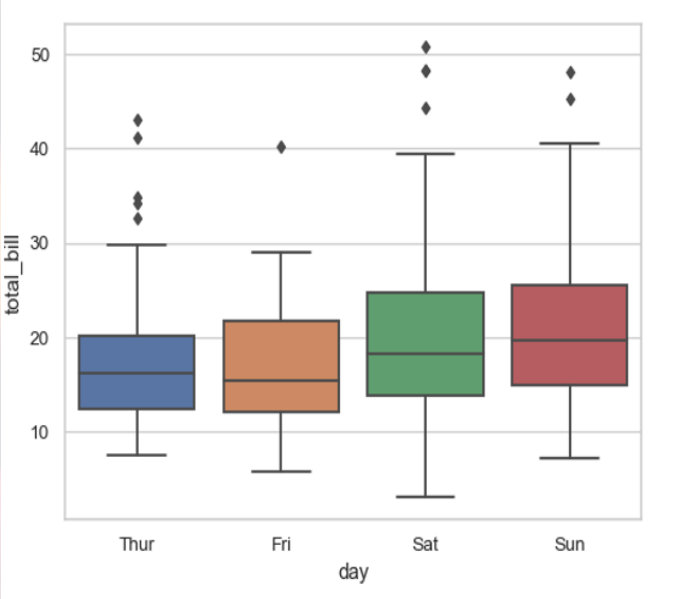
Line Charts

- *Trend in numeric data*
- *X axis is time based, chronological or sequential*



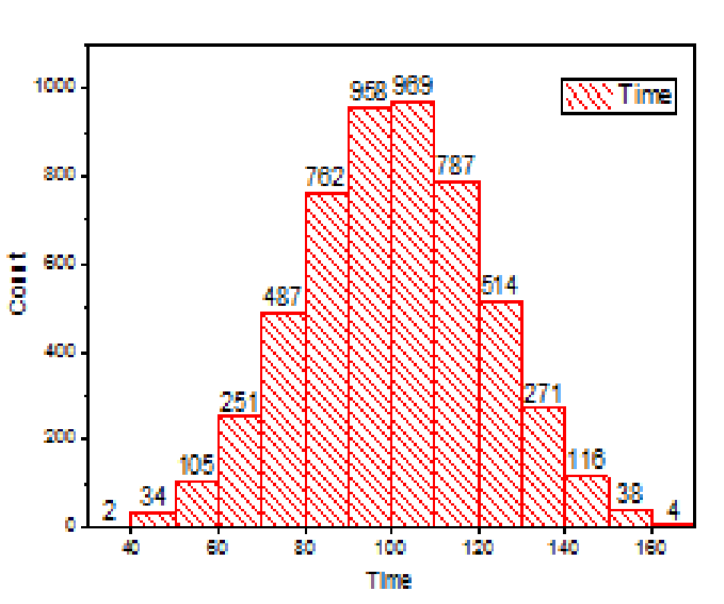
Box Plots

- *Helps Comparing distributions among the columns*
- *Helps to detect outliers*



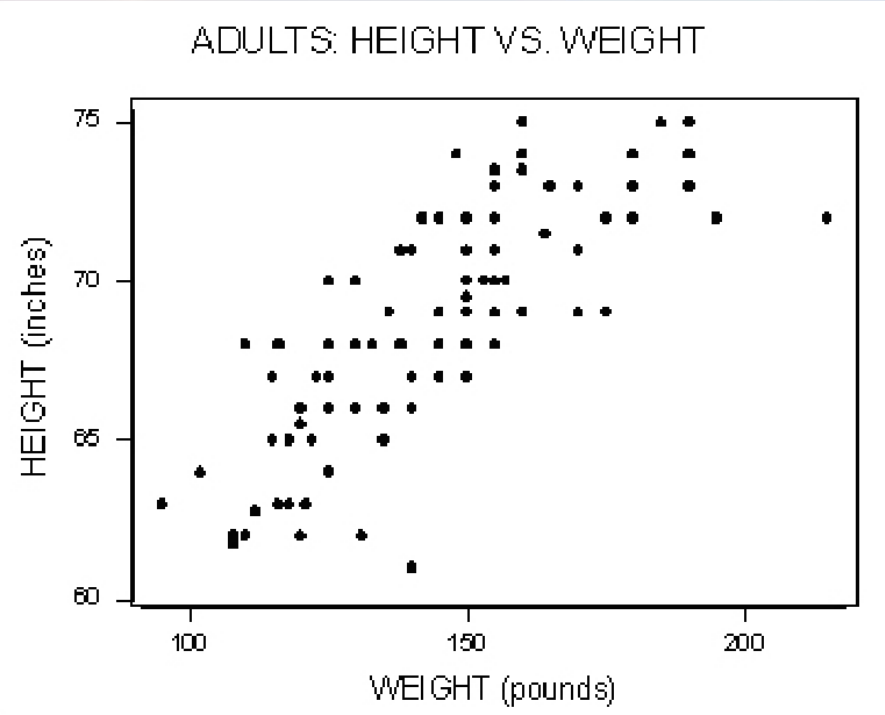
Histograms

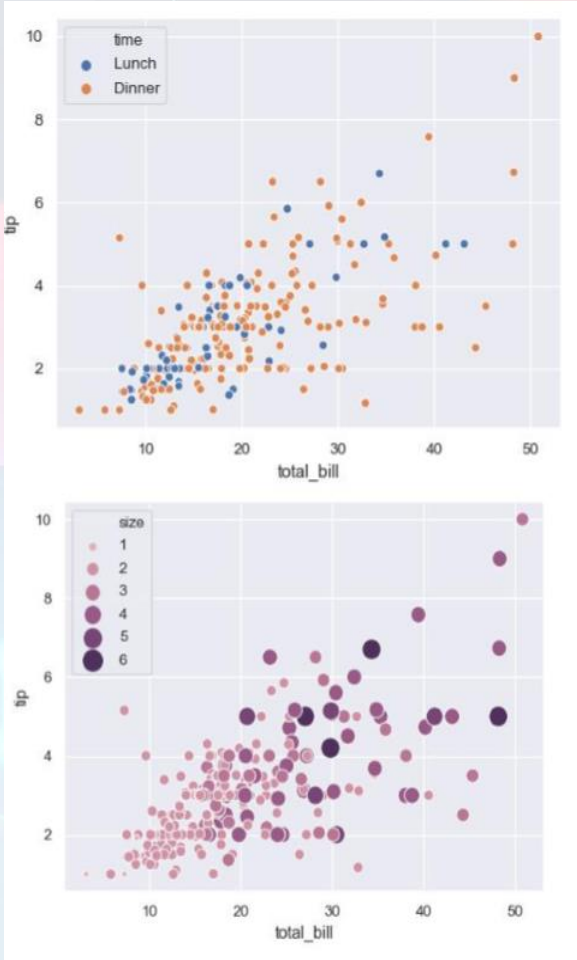
- *Helps understanding distribution of single feature*
- *1 Dimensional*



Scatter Plots

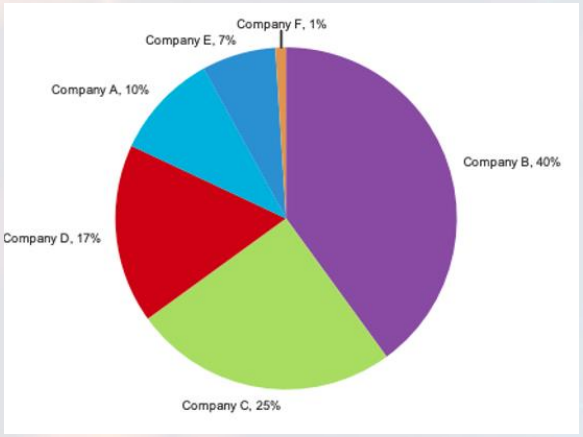
- *Numerical variable on both axis*
- *Plot shows correlation between X and Y*



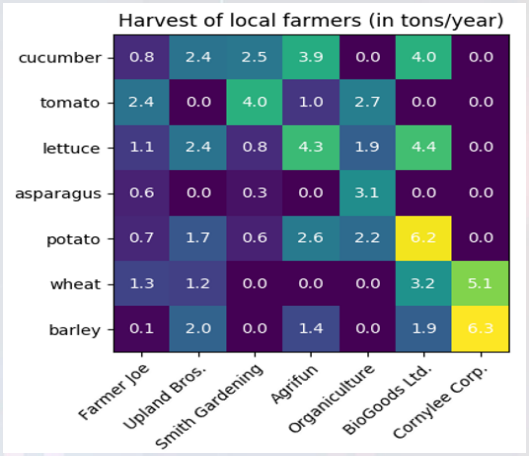


Colour|Shape|Size

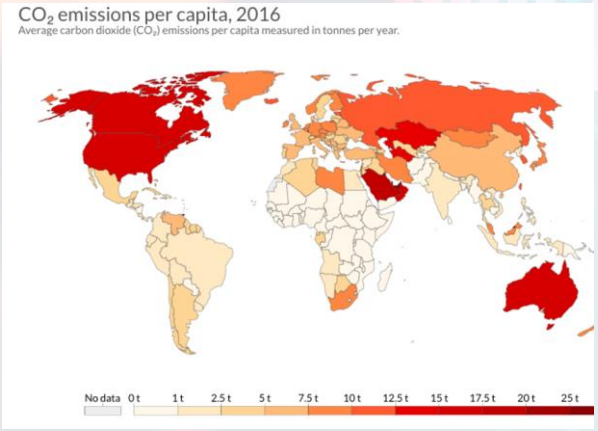
Other Plot types



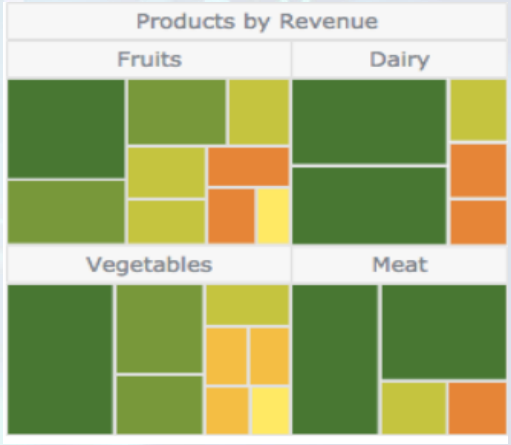
Pie Charts



Heat Maps



Map plots



Tree maps

Summary

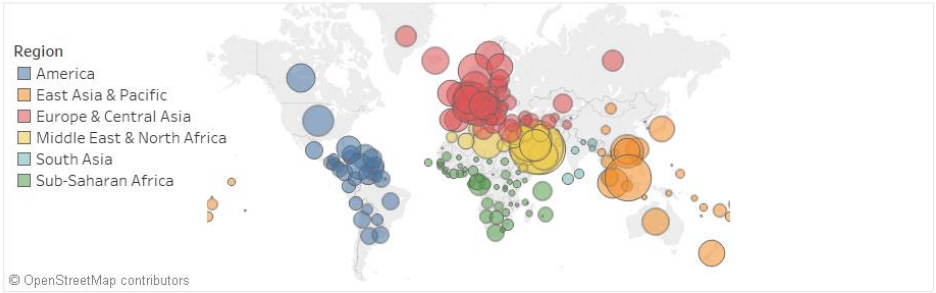
Graph Type	Usage	Additional Comments
Bar graphs	Comparing data values within or across categories; Discrete data	Consider Line graphs for Continuous data
Histograms	Distribution of values across a possible range	
Line graphs	Continuous data; Display trends	
Time series	Data with a time dimension	
Pie graphs	Comparing fractions of a whole; Very few fractions and precision is not important	Avoid to the extent possible
Gauge charts	Comparing values between a small number of variables	
Scatter plots	Understanding correlations between two quantitative dimensions of data	3 or 4 dimensions possible by encoding data points as bubbles, etc.
Heat maps	Area graphs that use colour or brightness to indicate values (or changes in value) of large data sets; Show relationships between 2 factors	
Box plots	Understanding distribution of a numerical data; Comparing distributions across categories; Identify outliers	
Tree maps	Display hierarchical data in rectangles	

Tableau Hands-on workshop

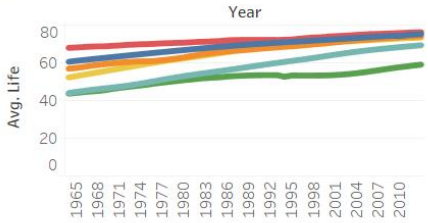
Global INdicators

Gap-minder_data

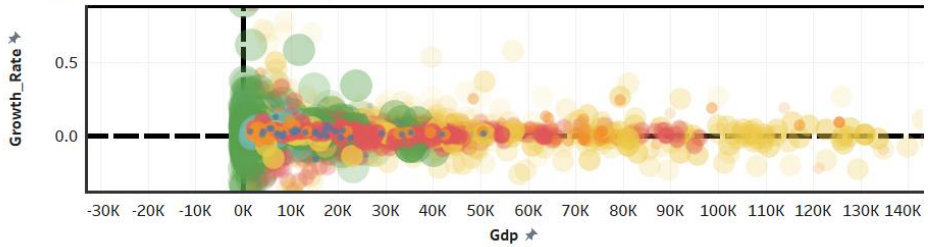
Global GDP Scenario



Global Life Expectancy over the years



GDP_Growthrate over the years - 2012



Global Fertility rate

Components & Characteristics of a Good Data Visualization

Components		
Justification	**Information**	**Conclusion**
• Right plot choice	• Right variable and numbers	• Right insights. Bottom line or final learning

Characteristics		
Simplicity	**Readability**	**Integrity**
• Minimum ink with maximum information	• Labels, legends, title, fonts	• Concluding only what exists

"In good information visualization, there are no rules, no guidelines, no templates, no standard technologies, no stylebooks ... You must simply do whatever it takes."

—Edward Tufte

Thank you !

https://www.kaggle.com/sriharipramod

https://medium.com/@sriharipramod

https://twitter.com/@sriharipramod

https://www.instagram.com/sriharipramod

https://github.com/SrihariPramod

Srihari Pramod

Ex Lead – 3D Visualization and Technical

MYOS Property.com

Dataset :
https://www.kaggle.com/sriharipramod/gap-minder-gdpgrowth

Further Reference if interested:

https://www.gapminder.org/data/

https://www.gapminder.org/videos/hans-rosling-on-cnn-us-in-a-converging-world/#

https://www.edwardtufte.com/tufte/