Introduction to Data Analytics

Lecture: Descriptive Statistics: Graphical Approaches

NPTEL MOOC

By

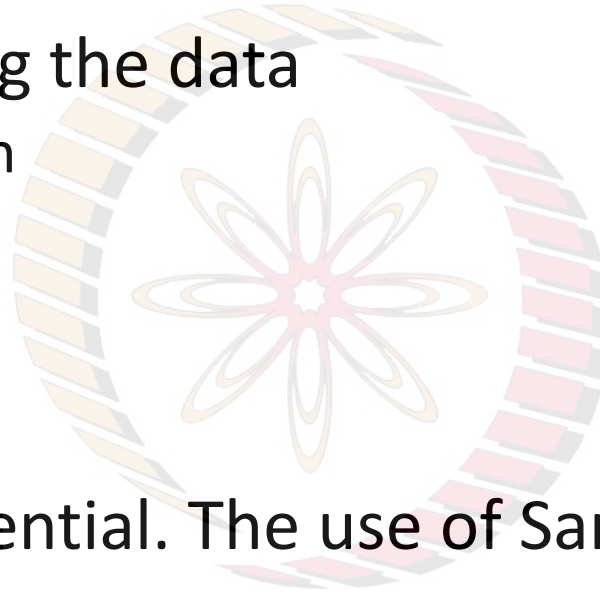Prof. Nandan Sudarsanam, DoMS, IIT-M and

Prof. B. Ravindran, CS&E, IIT-M

Data and Statistics

- Statistics is the science of learning from Data

- Data is essentially numbers (or text/symbols) which represent some information.

- It helps to think of data as 'values' of quantitative and qualitative variables

- What are the variable types:
  - Numerical or Quantitative: (Continuous and Discrete)
  - Categorical or Qualitative: (Always discrete)
    - Nominal
    - Ordinal

Descriptive Statistics

- Quantitatively describing the data
    - Graphical representation
    - Tabular representation
    - Summary statistics

- Descriptive versus Inferential. The use of Sample and Population

- Single and Multiple variables
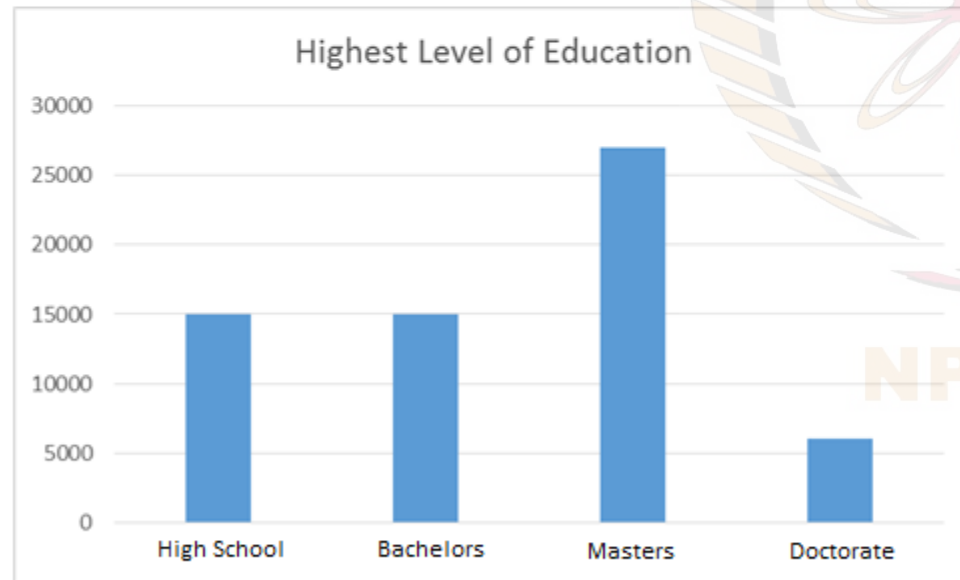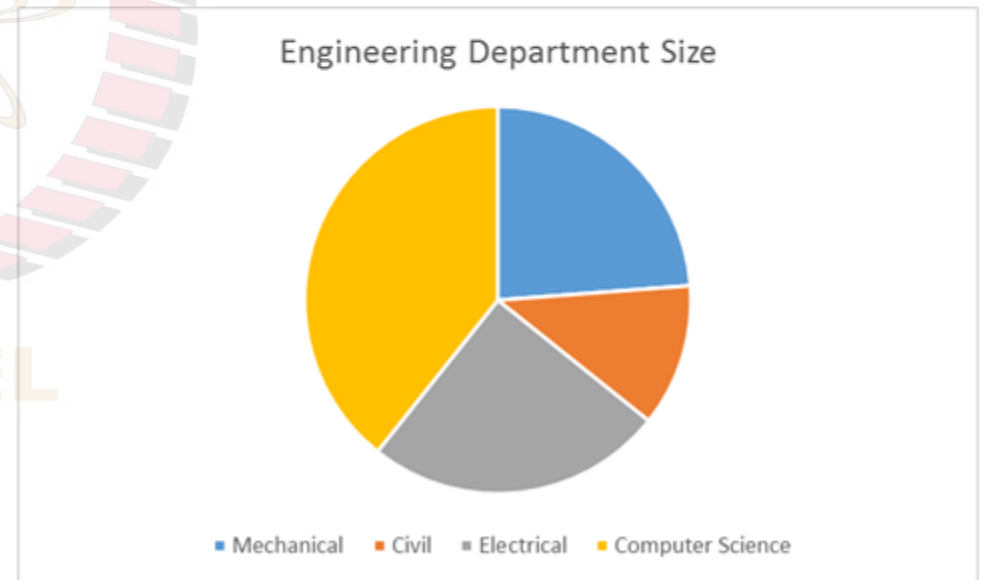   - Distributions and Relationships

Graphical Representation: Single Variable
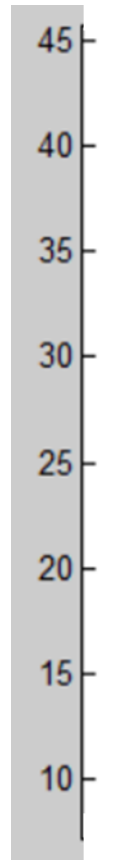
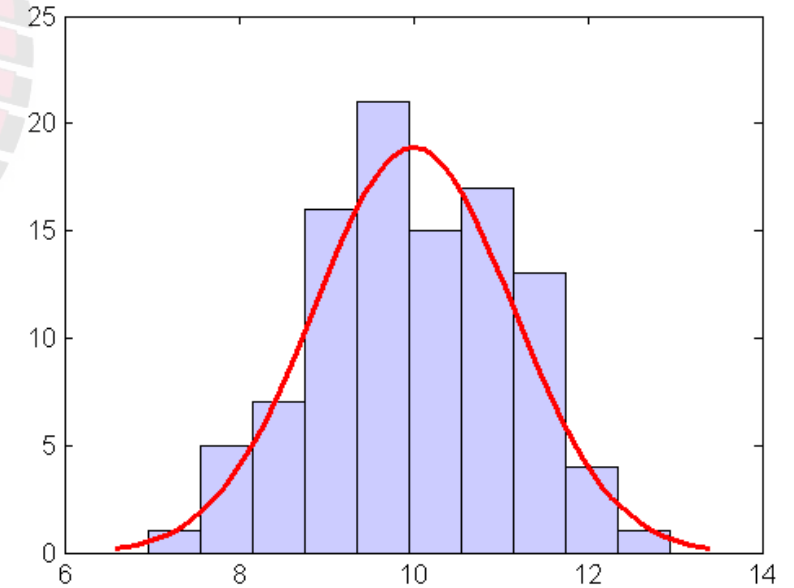- For Categorical Variables

Graphical Representation: Single Variable

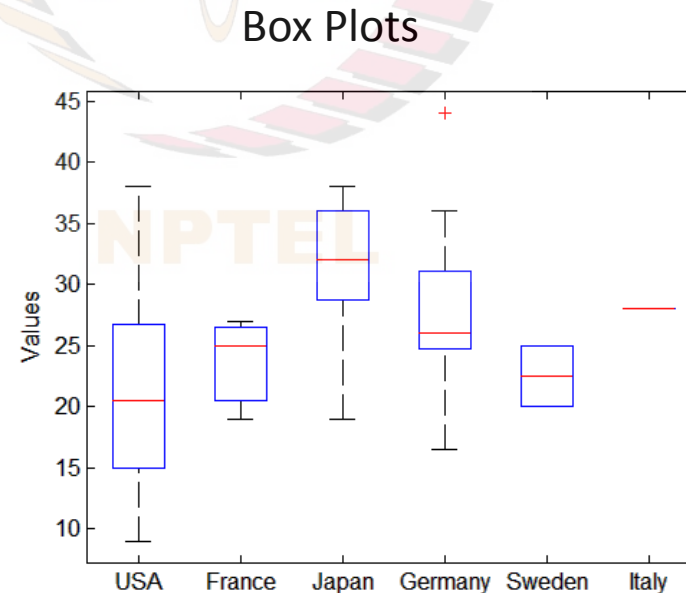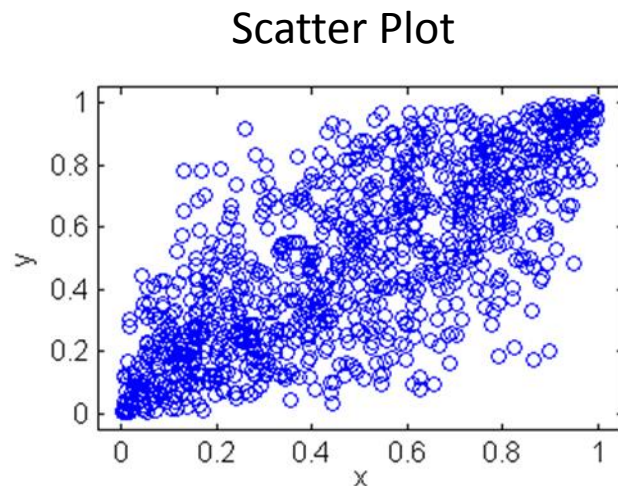- For Quantitative variables

Box Plot

Histogram

Graphical Representation: Multiple Variables

- Scatter Plots: Two quantitative variables

- Box plots – One categorical with one quantitative variable

- Contingency tables - 2 categorical variables with frequency of occurrence as the theme

Scatter Plot

Box Plots

Contingency Table

Results		Work EX	
		Y	N
MBA	Y	22	7
	N	32	17