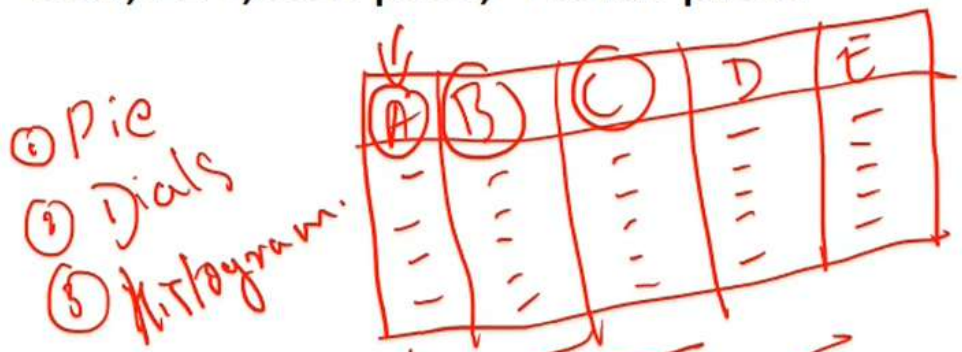
Univariate analysis

- provides summary statistics for each field in the raw data set (or) summary only on one variable.
- Ex:- CDF,PDF,Box plot, Violin plot.



① Pie
② Dials
③ Histogram

① width – x-axis
② Height – y-axis
③ depth – z-axis
④ hue
⑤ size
⑥ shape

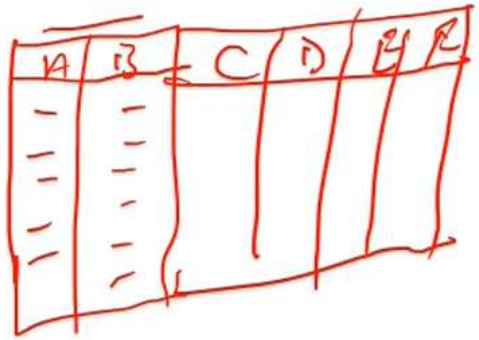
20-22, 23-25

Age- 20, 22, 25, 27 21, 23, 26, 22, 20, 21, 27, 29

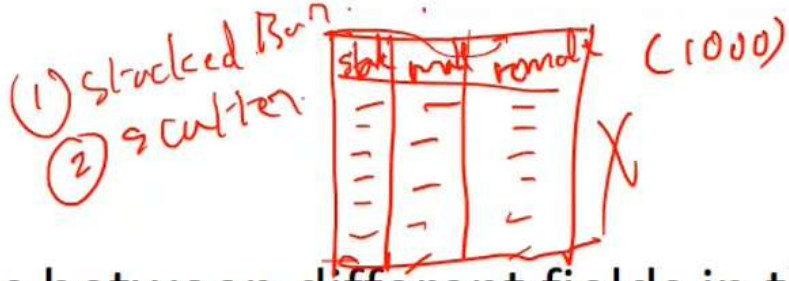
Bivariate analysis

- performed to find the relationship between each variable in the dataset and the target variable of interest (or) using 2 variables and finding the relationship between them.
- Ex:-Box plot, Violin plot., Scatter, Bar, Line.

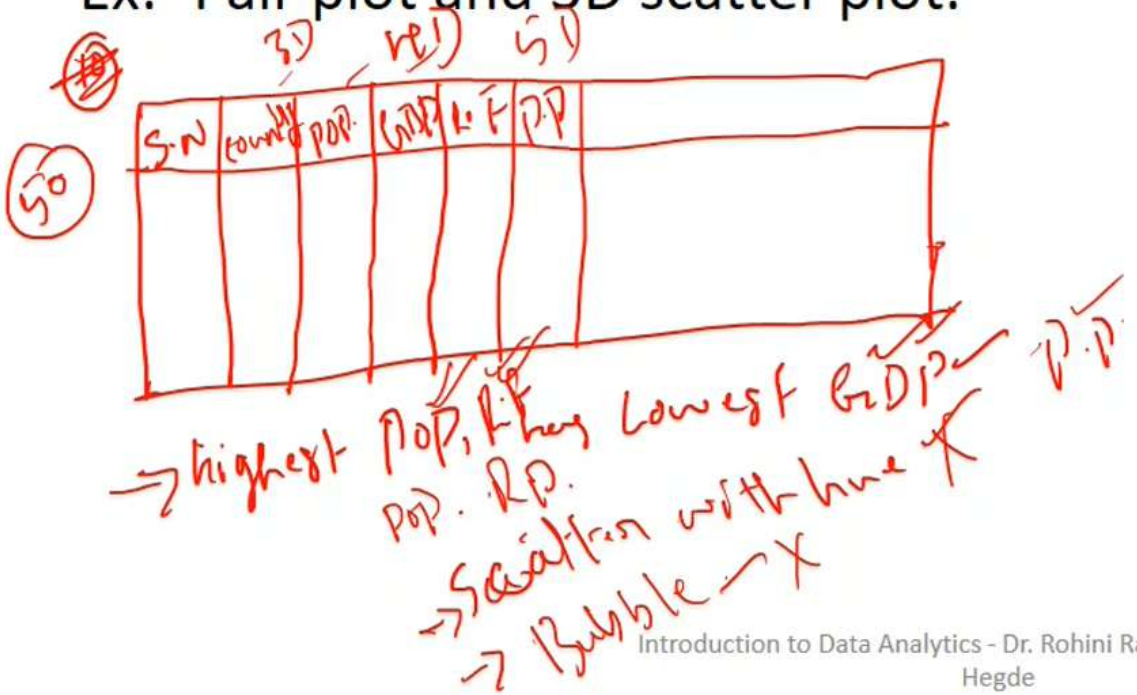
width – x
height – y

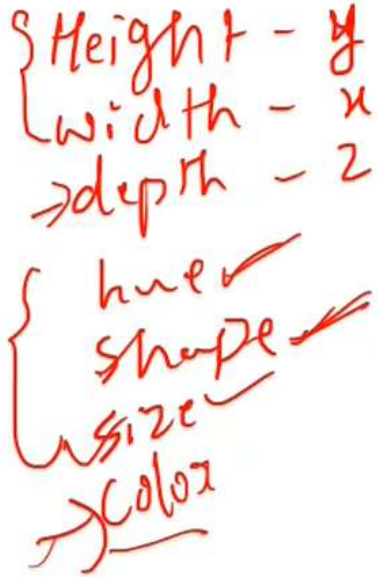


Multivariate analysis



- performed to understand interactions between different fields in the dataset (or) finding interactions between variables more than 2.
- Ex:- Pair plot and 3D scatter plot.





Example

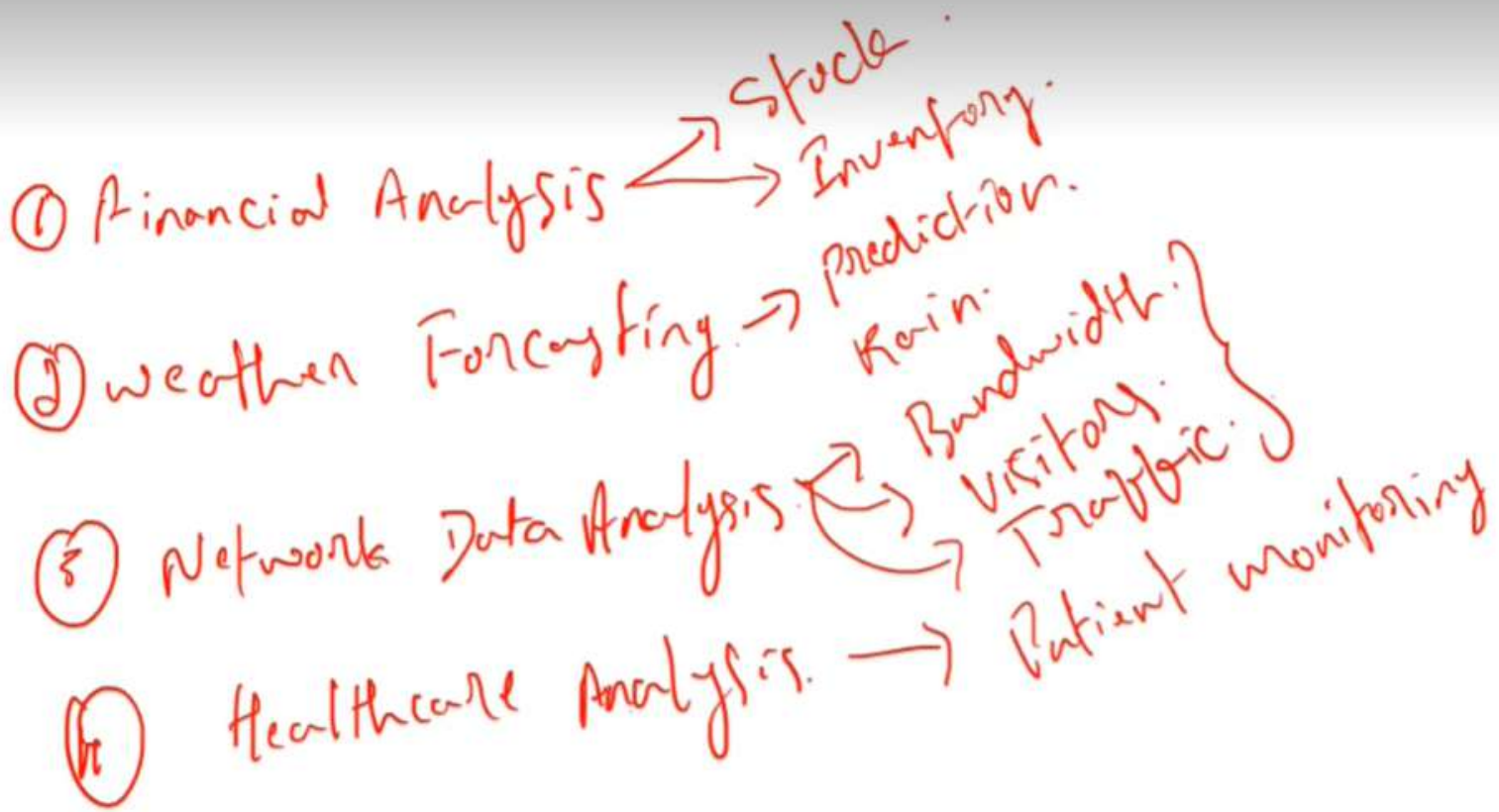
- About Dataset:
  - The dataset contains cases from a study that was conducted between 1958 and 1970 at the University of Chicago's Billings Hospital on the survival of patients who had undergone surgery for breast cancer.
- Attribute Information:
  - Age of patient at time of operation (numerical)
  - Patient's year of operation (year - 1900, numerical)
  - Number of positive axillary nodes detected (numerical). Lymph node in the area of the armpit (axilla) to which cancer has spread.
  - Survival status (class attribute):
    - 1 = the patient survived 5 years or longer
    - 2 = the patient died within 5 year

What is a Time Series Data?

- Anything that is observed or measured at many points in time forms a time series.
- Many time series are fixed frequency, where data points occur at regular intervals according to some rule, such as every 15 seconds, every 5 minutes, or once per month.
- Time series can also be irregular without a fixed unit of time or offset between units.

How you mark?

- How you mark and refer to time series data depends on the application
- One of the following:
  - Timestamps, specific instants in time
  - Fixed periods, such as the month January 2007 or the full year 2010
  - Intervals of time, indicated by a start and end timestamp. Periods can be thought of as special cases of intervals
  - Experiment or elapsed time; each timestamp is a measure of time relative to a particular start time (e.g., the diameter of a cookie baking each second since being placed in the oven)


1) Financial Analysis
Stock
Inventory.
2) weather Forcasting
prediction.
Rain.
3) Network Data Analysis
Bandwidth.
Visitors.
Traffic.
4) Healthcare Analysis.
Patient monitoring