Q) difference in descriptive, diagnostic, predictive & Prescriptive analytics.

Comparison	Descriptive analytics	Diagnostic analytics	predictive analytics	prescriptive analytics
	It gives us only insight about whether everything is going well or not in your business without explaining root cause	Explain root cause behind the outcome of descriptive analytics.	By utilizing vario- us statistical & machine learning algo to provide recommendation & prov ide ans. to ques. related to what might happen in the future (can't answered by BI)	use to advise on possible outcomes & what should they do to maximize their key Bussines metrics
Complexity	level - 1	level - 2	level - 3	level - 4
To find ?	what is happening ?	Why it is happening ?	What is likely to happen ?	What do I need to do ?
Sectors	marketing, Sales finance, Retail, Construction, Banking, Transportation (58%)	(34%)	Banking, manu- facturing, cyber- security, Transport Healthcare, Retail (36%)	manufacture, utilities, Pulp & paper, mining & metal, oil gas, consumer goods

Q) Descriptive Vs infential stastics

Descriptive statistics	Inferntial statistics.
- It is a branch of statistics in which is concerned with describing the popwlation under study	- It is type of statistics that focus on drawing conclusions about the population on the basic of sample analysis & observation
- organize analyze & present data in a meaningfull way	- Compare, test & predicts data
- final result in form of Charts, graphs & Tables	- final result in form of probability
- It explains the data which is already known to summarize sample	- It attempts to reach the conclusion to learn about population that extend beyond the data avilable

① Descriptive statistics. –
- summarize data
- organize data
- simplify data

Eg:- Tables, graphs average.

② Inferential statistics
- study samples to make generalization about the population
- Interpret experimental data.

1. Descriptive analytics. - process of parsing historical data to better understand data the changes that have occoured in bussiness
- Descriptive analytics can help to identify the area of strength & weakness in an organization
- Descriptive analytics take raw data & process that data to draw conclusion that are useful & understandable

2. Diagnostic analytics - The goal of diagnostic analytics is to understand why something happned.
- At this stage historical data can be measured against other data to answer the question "why something happned?"

3. predictive analytics - It tells what is likely to happens.
- It uses descriptive & diagnostic analytics to detect tendencies, clusters & expectations & to predict future trends.
- It is valueable tool for forcasting.

4. presprictive analytics. - purpose of this analytics is to literally prescribe what action to take to elimate a future problem or take full advantage of a promising trend.

Descriptive statistics

- Classical
  - central tendency
  - frequency table
  - spread (Range, verience, standard devicaon)
- graphical
  - histogram
  - Box plot (whisker)
  - line plot
  - scatter plot
  - pie chart

i) Range < quartile / Interquartile
(max-min)

ii) standard deviation = $\sqrt{\text{varience}}$

iii) varience =

Eg :- P=50 median
P=25 lower quartile LQ
P=75 upper quartile UQ

Interquartile = UQ - LQ.

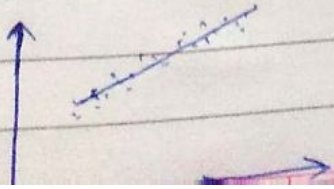
Extrem → can be flase data
outlier → data point greater or smaller than data set

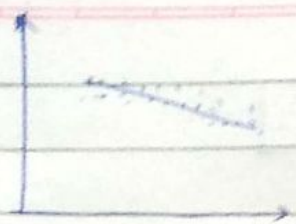
* Contengency Table

Eg

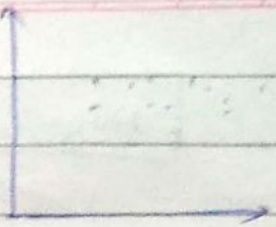
A \ B	0	1
0	00	01
1	10	11

* Correlation (-1 to 1) < +vely / -vely / non-corelated





(-1 to 0)
-very corrth

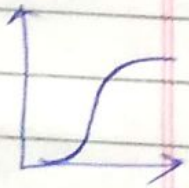


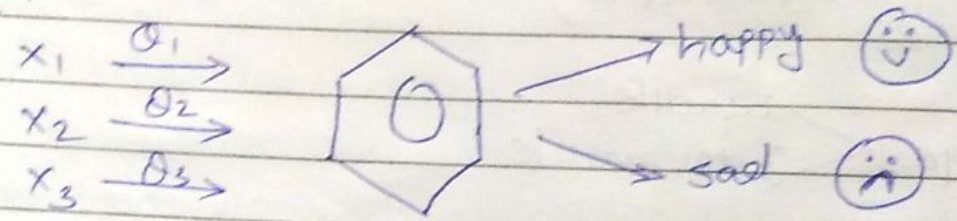
Netural 0
non-corelated

* Regression < logestic
linear

① logestic regression is predective analyse.

Eg:



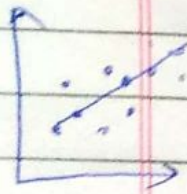


x_1, x_2, x_3 → i/p , $\theta_1, \theta_2, \theta_3$ → weight Happy, sad → o/p

- logestic Regression used to describe data & explain relan between one dependent variable & one or more independent variable
- logestic regression is used when independent variable is binary in ~~region~~ nature

② linear Regression

- linear Regression is used when dependent variable is continious in nature



measure of spread → range, varience, standard deviation
measure of position → mean, median, mode
measure of shape → probability distribution

Present in form	data type
Histogram	Continious, discrite
Box	Continious.
line	Continious
Scatter plot	Continious.
pie	discrite.

* Quantitative → information is recorded as no. & represents an objective measurement or a count.
  Eg Temp., weight.
  Quantitative data also refred as numerical data

Qualitative → Information represents characteristics that we can measured in number
  Eg eye colour, Taste, marital status.

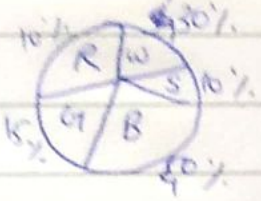
* Type of quantitative data.

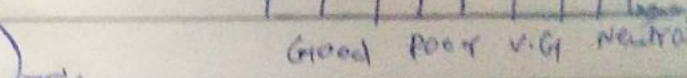
continious
- Can take numeric value, & it can be meaningfully divide into smaller increments, decimal value.
  % fat
  23.9, 32.4, 22.5
  28.8, 25.8

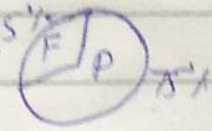
descrite
- It is count of the presence of chart., result item, activity.

* Type of qualitative Data (Pie chart, bar chart)

Categorical	Binary	Ordined
- value that you can put into countable eg: Psychology, Political Science, biology, engineering • Car Colour → White, silver, Black, Gray, Red	- can have only two value. If can place an observation into two category you have Binary value Eg:- pass/fail yes/No male/female	- have atleast 3 categories Eg: very poor, Poor, Neutral, Good, very good.







Inferential statistics

(use to predict data)

Null hypothesis

→ null hypothesis always state that nothing is going on

there is no difference, no relationship, no treatment effect, etc.

- H_0 : $X = Y$

→ The alternate hypothesis states that there is a difference

- H_a : $X \neq Y$ (non-directional)
- H_a : $X > Y$ or H_a : $X < Y$ (directional)

Transformation Technique → standardization, Normalization

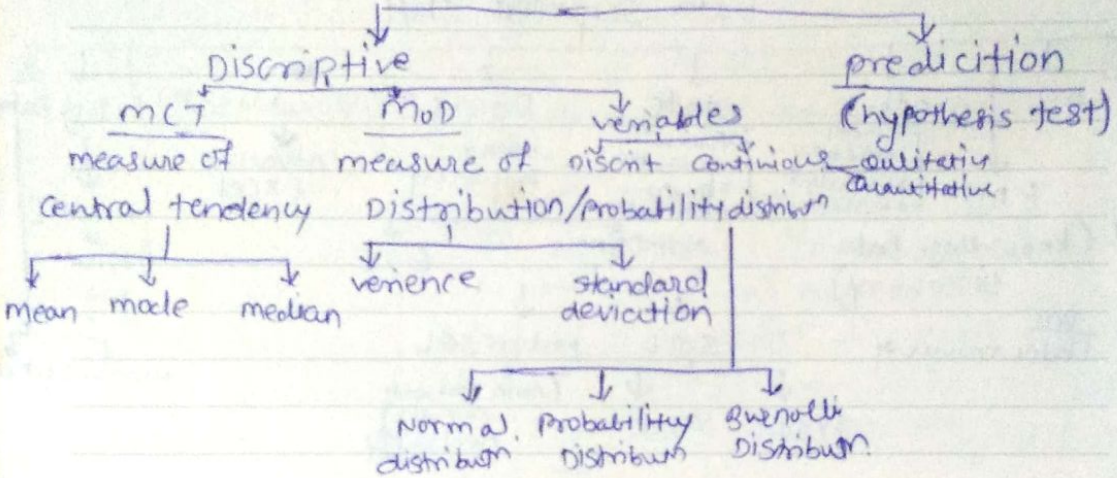
MARCH 2010

15 MON

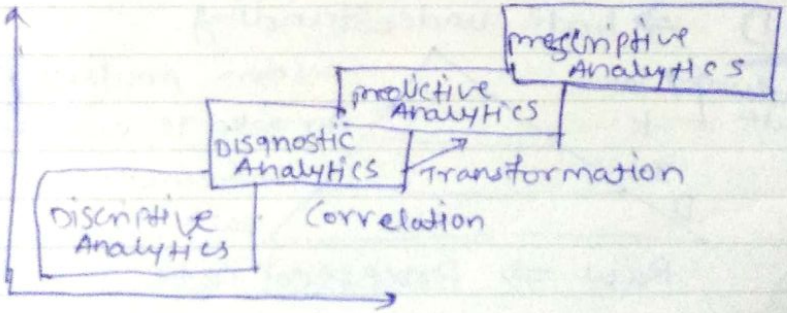
Statistics (Data)

→ To describe data (discriptive statistic)

→ prediction (inferential statistic)



16 TUE



637 nw- about MCT, MOD, variables & Hypothesis test