STATISTICS

Statistics is the science of collecting, organizing, analyzing, presenting and interpreting the numerical data to assist in making more effective decisions.

3 parts

1. Basic statistics
2. Graphs
3. Visualization

Statistics ====== very much important in order to understand the data.

What is statistics ?

1. Mathematics
2. Probability
3. Data
4. Facts and figures
5. Mean median mode
6. Graphs
7. Ratio

A data

Virat kohli average in ODI is 60

I asked to Rashad, is that true?

What Rashad will do?

1. Collect the data
2. Organize the data
3. Analyse the data
4. Interpretation in graphical view

Data:

How many types of data available

String

float

Int

How many types of data :

English ====== maths

1. Numerical data ========== maths

(Quantitative data)

1. Categorical data ========== english

(Qualitative data)

Again numerical data is divided into two parts:

1. Continues data
2. Discrete data

Example:

My salary : 50000

My salary: 50000.45

1. Continues data

What is the temp : 30.5

98.4F==🡺 fever 102

Room temp:<>

1. Discerte data:

How many houses you have: 1 or 2

How many friends you have : 10 or 20

How much is you age : 29 or 30

|  |  |
| --- | --- |
| Numerical data (numbers representation) | Categorical data (English word rep) |
| Quantitaive data | Qualitative data |
| Continues data |  |
| Discrete data |  |

Levels of data:

1. Nominal data

* Categorical type
* Name of the person
* Name of the city
* Hyd blr Chennai
* Very very base level

1. Ordinal data

* Categorical data
* But it maintains some order
* Flop average hit superhit blockbuster
* Primary -- secondary -- teritiary
* High -- medium – low
* First class – second class – third class

1. Interval data

* Numerical type of data
* Extract equal intervals will be there
* It does not have a zero scale
* Temperature is the best example for interval level of data

1. Ratio level data

* The temperature is in Hyderabad twice as benguluru
* Temp(hyd) = 2\*temp(blr)
* Temp in hyd 50c
* Temp in blr 25c

Is this true

Units?

50c = 2\*25c ============

Temp(hyd) in F = 2\*temp(blr) F

50 in c ======== > 122 F

25 in c ======== > 77 F

122 = 2\*77

Weight = 60kgs

Kid = 30 kgs

60=2\*30kgs

If it is having zero scale ============ > Ratio level of data

If it does not have zero scale ========== > Interval data

Numerical data categorical data

Quantitative data qualitative data

Continues data

Discrete data

Interval data nominal data

Ratio level data ordinal data

Open intro sattsistics

Two types of statistics:

1. Descriptive statistics
2. Inferential statistics

Population sample

In india every person if you consider 140cr =========== > population

If you select only 1lakh ======================= > sample

Population means collection of every object

Sample means selection of some objects from corresponding population

Inferential statistics :

Will work on samples, will estimate the outcome on population

Descriptive statistics :

Will work on population, will estimate the outcome on population

99.99 % ====== > 100 %

Whatever you worked sample

You are estimating ============ mean

============ median

============ mode

Here mean is statistic

Median is statistic

Mode is statistic

You are working on a sample to identify a mean ======= > estimated mean of population

Here mean : statistic

Inferential statistics : statistic can be mean\ median\ mode

You are working on population ========== > estimating outcome on population

Outcome : parameter

Descriptive statistics : parameter can be mean\ median\ mode

Little difference between statistic and parameter?

Statistic can mean, median, mode : inferential statistics

Parameter can mean, median, mode : descriptive statistics

When will a mean will be statistic?

If you work on sample ========= estimating population

When will a mean be a parameter?

If you work on population ======= estimating population