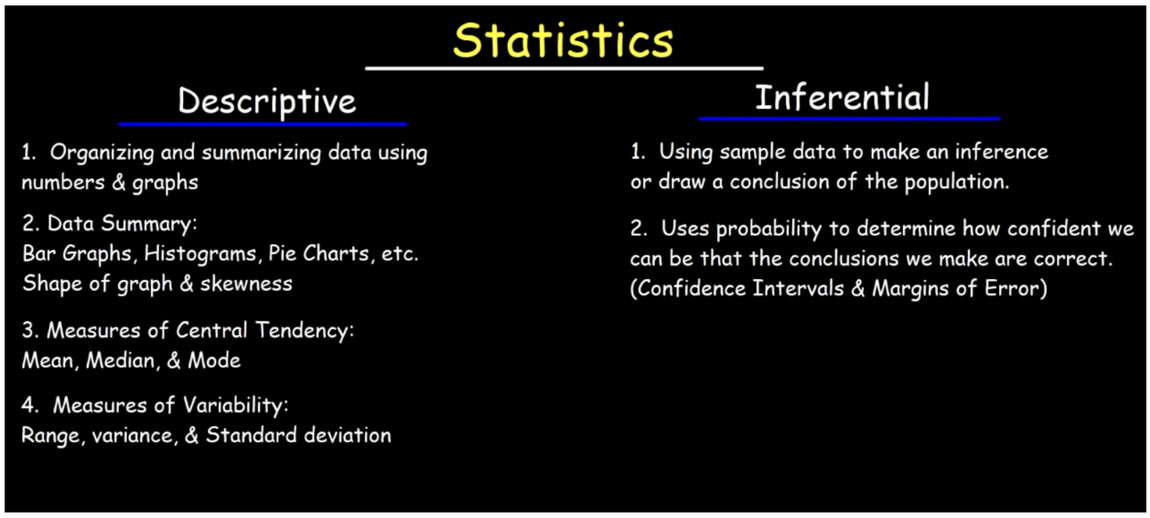
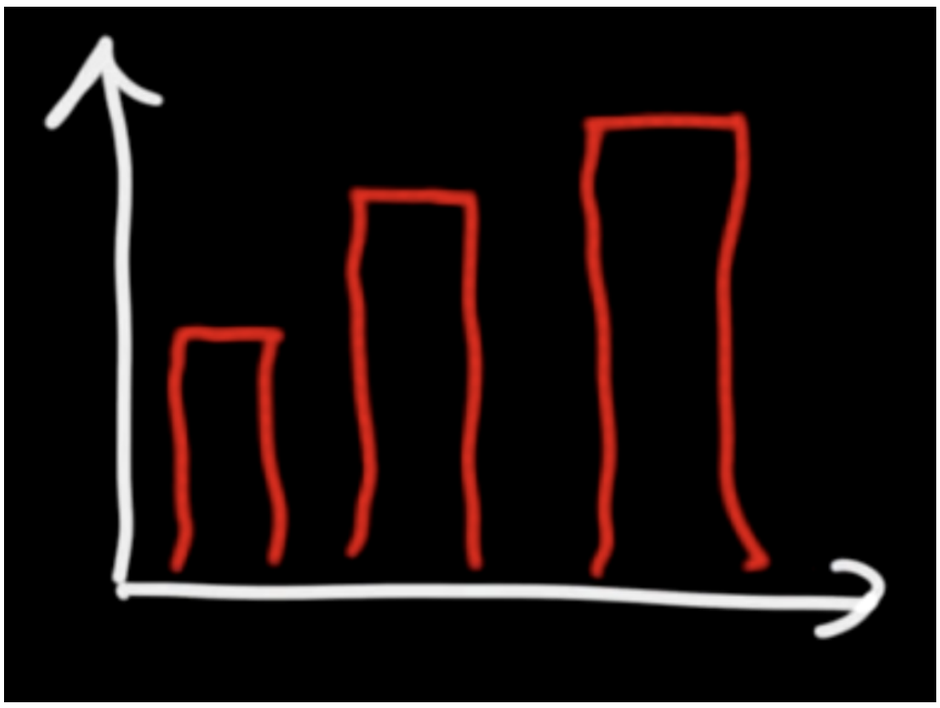
✅

Descriptive Statistics vs Inferential Statistics



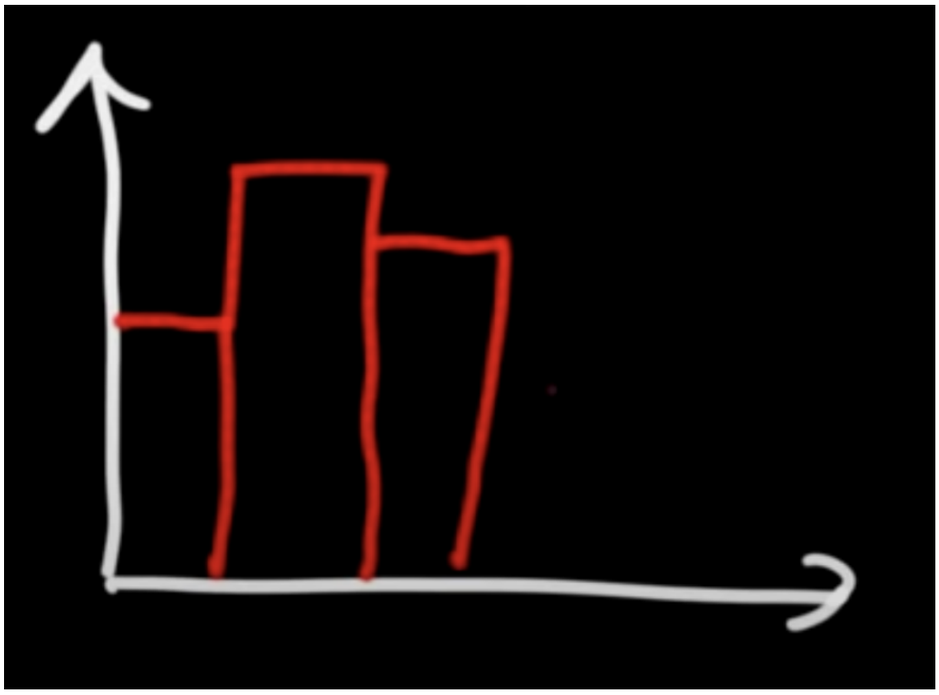
Slide 1

- Statistics
  - The science of different functions applied to the data is statistics like
    - Analysing data
    - Collecting data
    - Interpreting data
- Types of statistics
  - Descriptive statistics
    - It is used to organise and summarise data using numbers and graphs
    - For example
      1. A bar graph



A bar graph

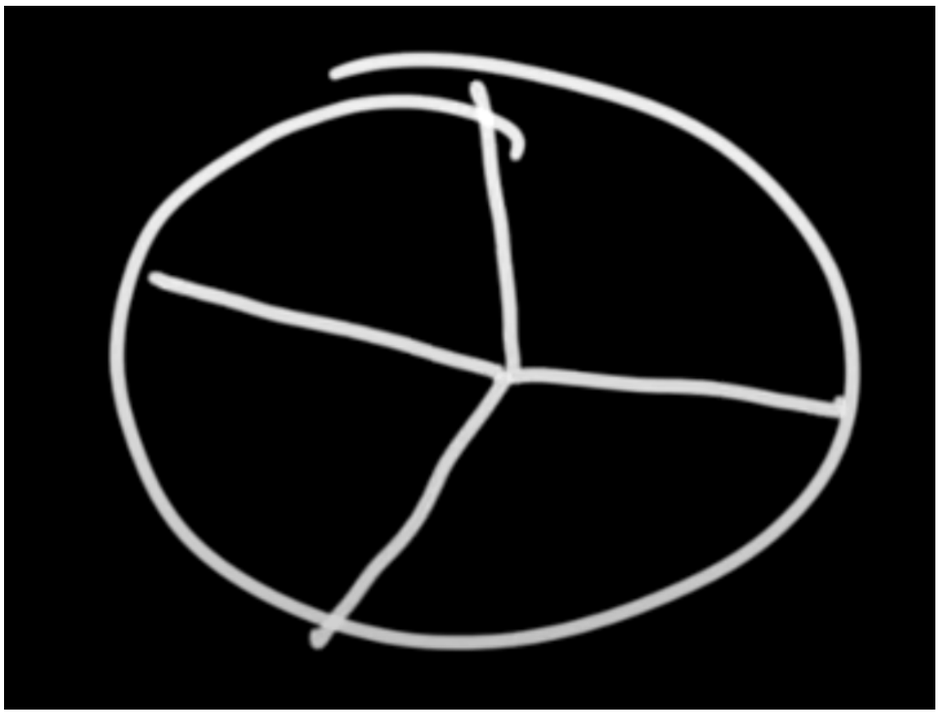
2. A histogram



A histogram

The bars are attached

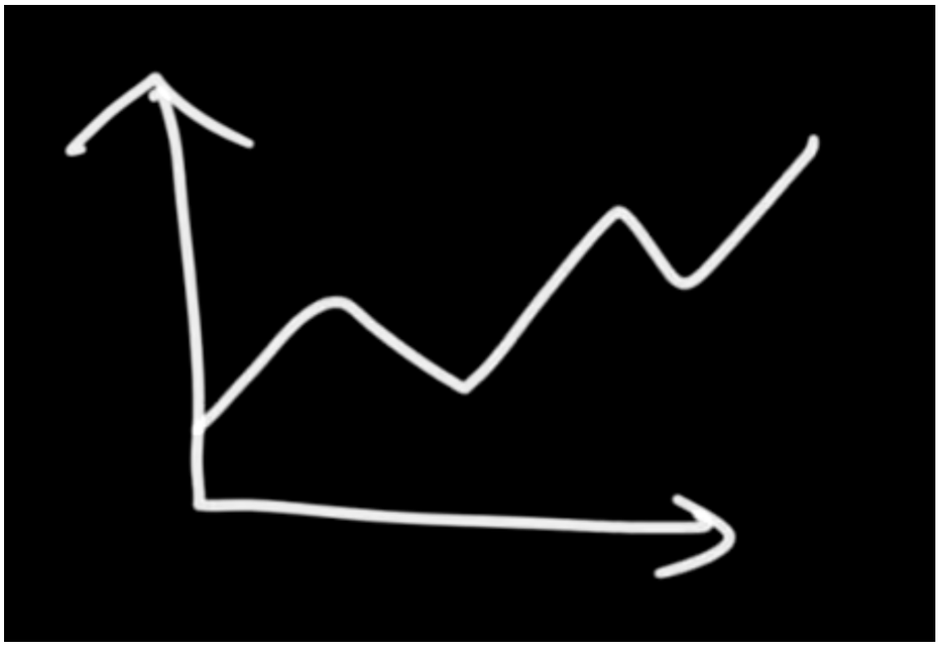
3. A pie chart



A pie chart

It is useful when data is divided based on categories and their respective percentages

4. A line graph

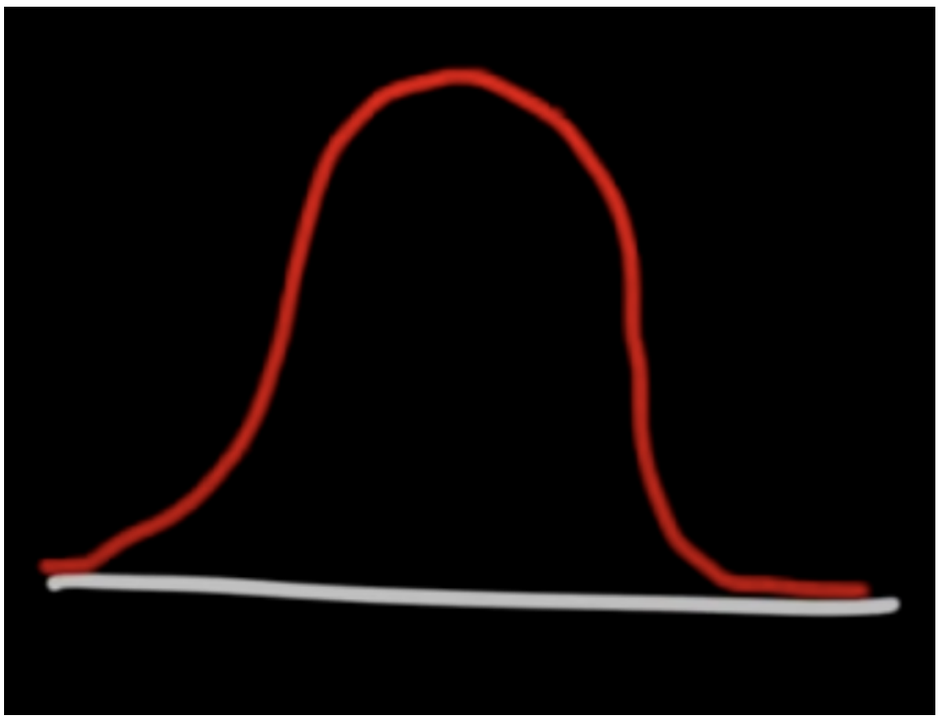


A line graph

5. A table

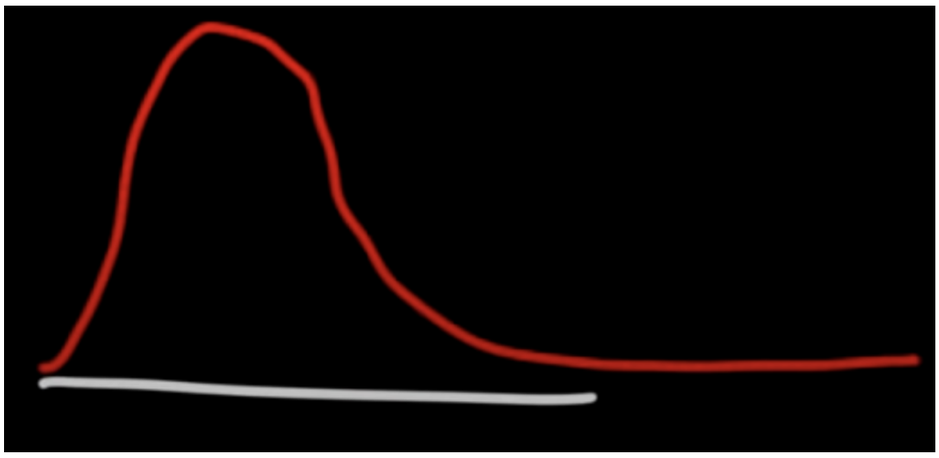
   Tables like frequency tables, frequency distribution tables, and many more are used to organise data

- These fall under descriptive statistics
- It can be used to describe the shape of the graph
  1. It can check if the data is symmetrical



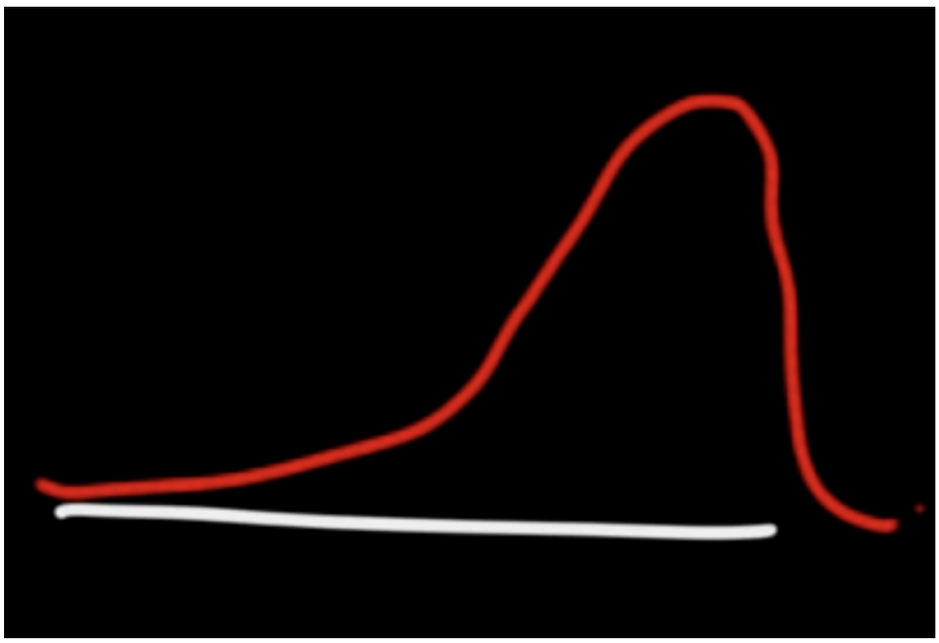
Symmetrical data

2. It can check if the data is skewed to the right



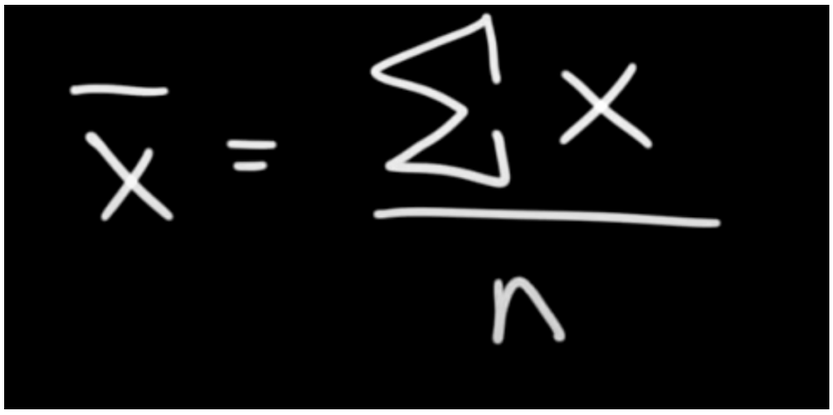
Data skewed to the right

3. It can check if the data is skewed to the left



Data skewed to the left

- It can be described using different numbers
    1. Measuring the central tendency is one such way
        a. Measuring the mean which is of two types
            i. Sample Mean



Sample Mean

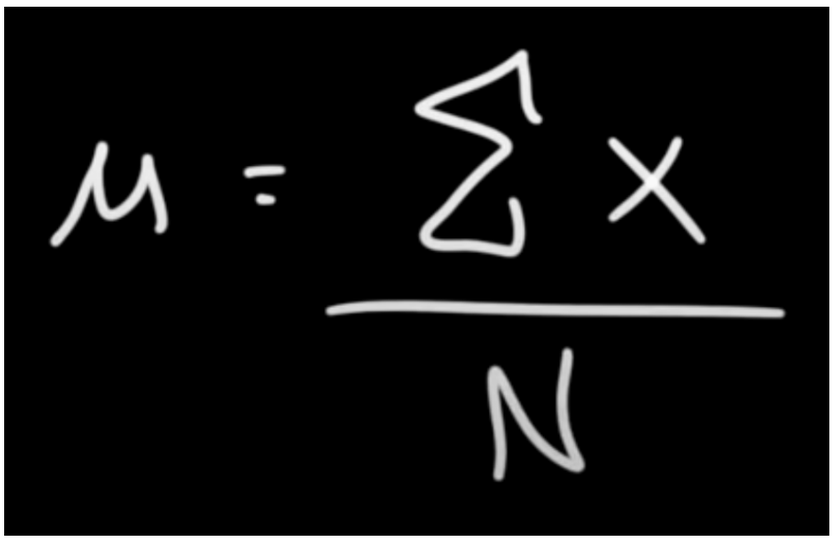
In the above formula

$\Rightarrow \bar{X}$ is the sample mean

$\Rightarrow \sum X$ is the sum of all the samples

$\Rightarrow n$ is the number of samples

ii. Population Mean



Population Mean

In the above formula

$\Rightarrow \mu$ is the population mean

$\Rightarrow \sum X$ is the sum of the data points

$\Rightarrow N$ is the number of data points in the population

b. Measuring the median which is the middle value of the dataset after sorting it

c. Measuring the mode which is the most frequently occurring value in the dataset

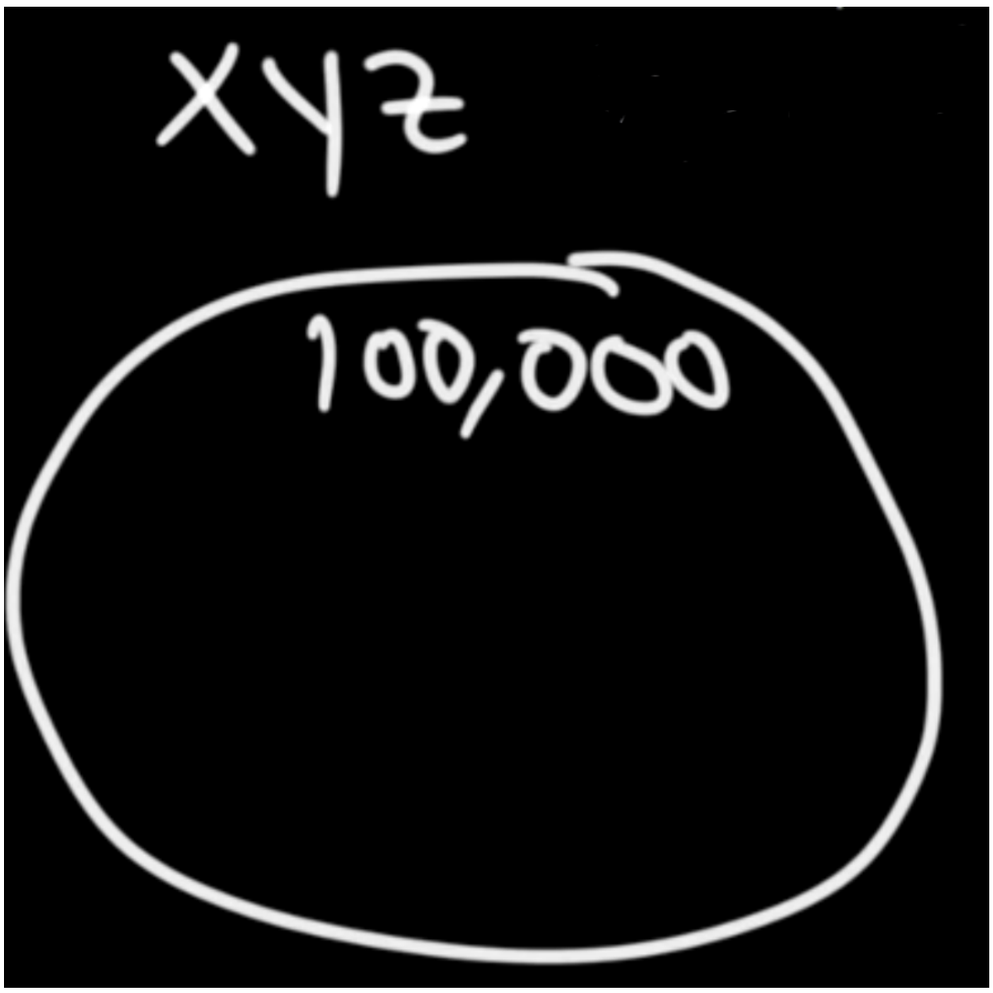
2. Measuring the variability is another way

a. Range

b. Variance

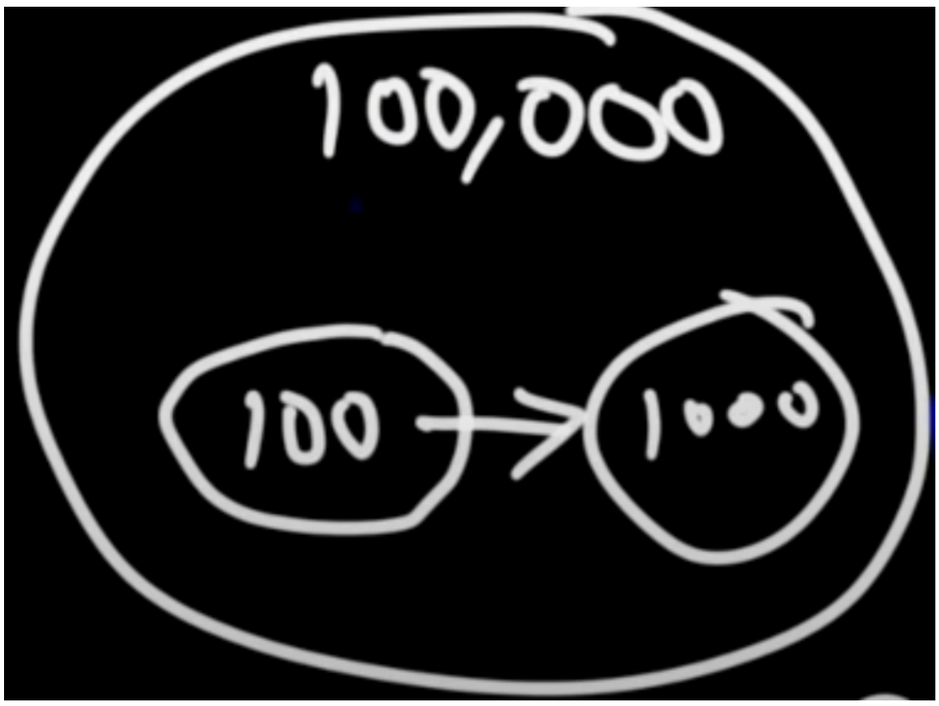
c. Standard Deviation

- Inferential statistics
  - It uses sample data to make an inference about the population
  - It can also use probability to determine how confident are the conclusions made
  - For example



A city XYZ having a population of 100000

1. A city XYZ has population of 100000 and a test is conducted to find how many prefer blue cars
2. It is not possible to ask everyone if they like blue cars
3. Take a small sample of 100 individuals and ask the chosen people if they like a blue car or not
4. If 20 people out of the sample like blue cars then the percentage of people liking a blue car is $\frac{20}{100} * 100 = 20\%$
5. A conclusion can be made about the entire population around $20\% \pm 2\%$ like blue cars where 2% is the margin of error
6. It is descriptive statistics when only the sample is considered
7. It is inferential statistics when the sample is used to conclude about the entire population
8. The confidence interval of concluding that around $20\% \pm 2\%$ is 95% so this is the confidence with which this is provided
9. This confidence interval can be increased by increasing the number of data points in the sample (i.e. choosing 1000 people to provide an opinion and it results in around $21\% \pm 1\%$ where 1% is the margin of error and this is better than the previously recorded value of $20\% \pm 2\%$)


100,000
100
1000