->statistics(stats) : collect , organise , analyse data and predict future outcomes

->descriptive:

1.central tendency (mean , median , mode)

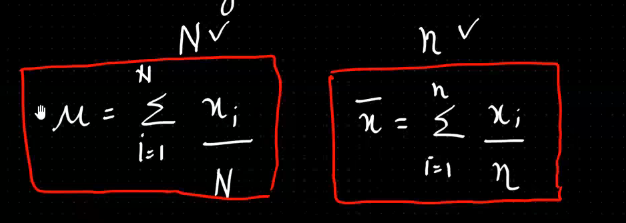
2.measure of dispersion (variance , standard deviation)

->inferential:

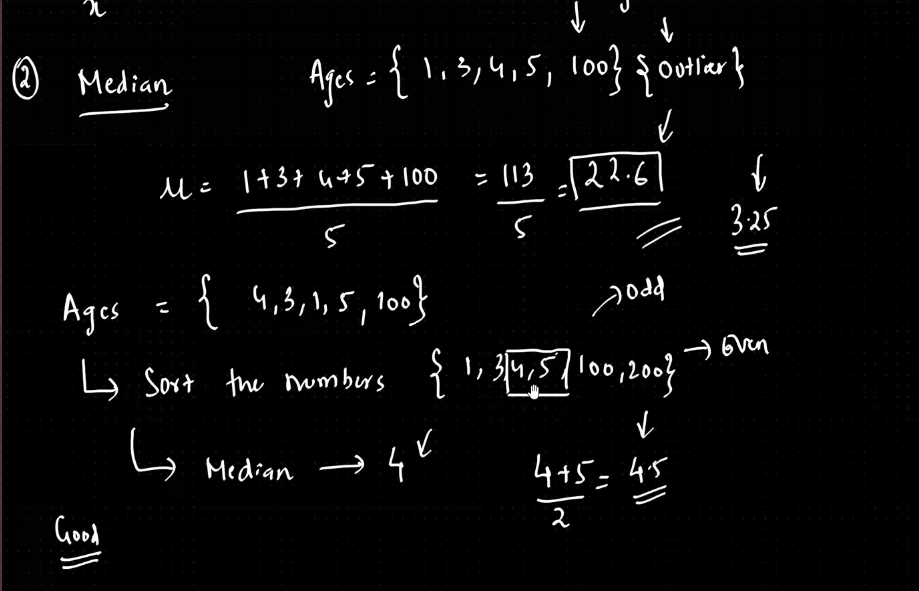
->collect data -> conclusions or inferences -> other data

->CENTRAL TENDENCY:

- population mean(N) and sample mean(n):



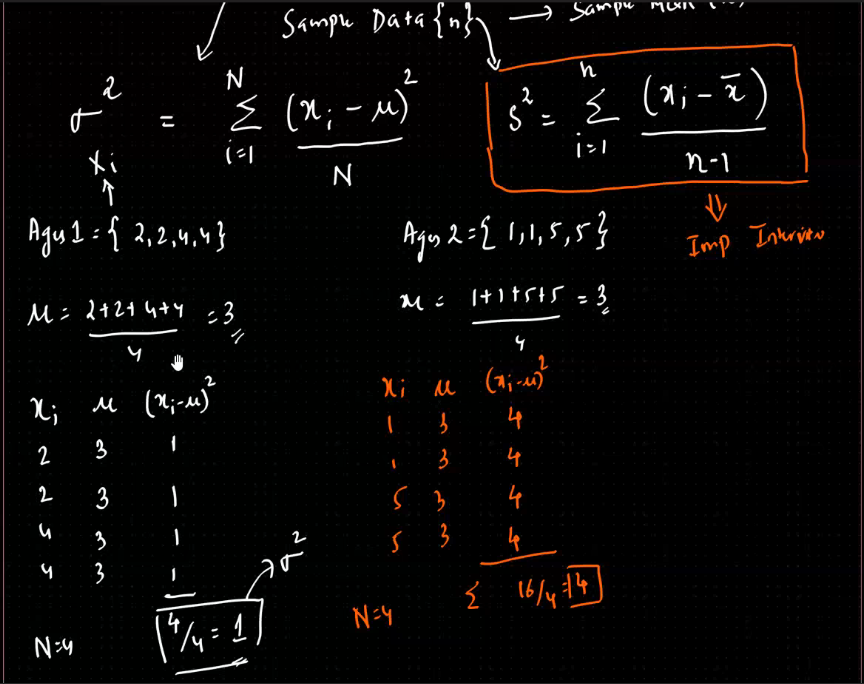
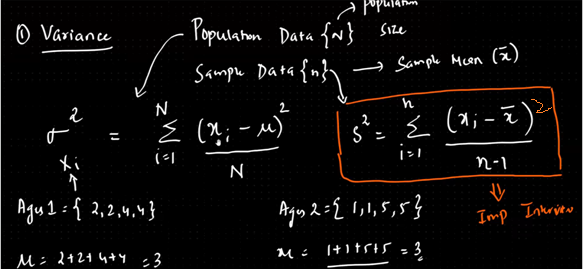
-mean:

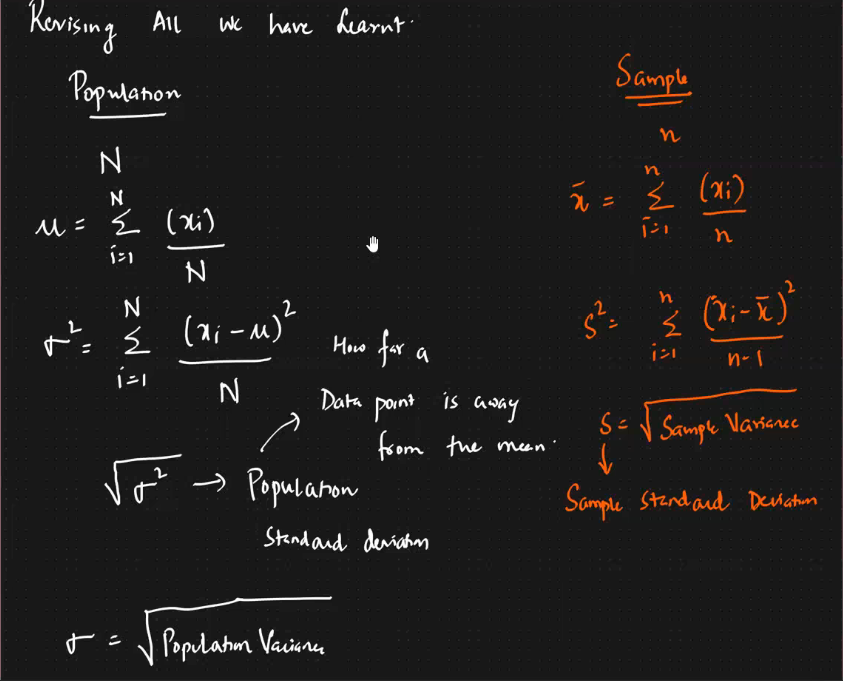


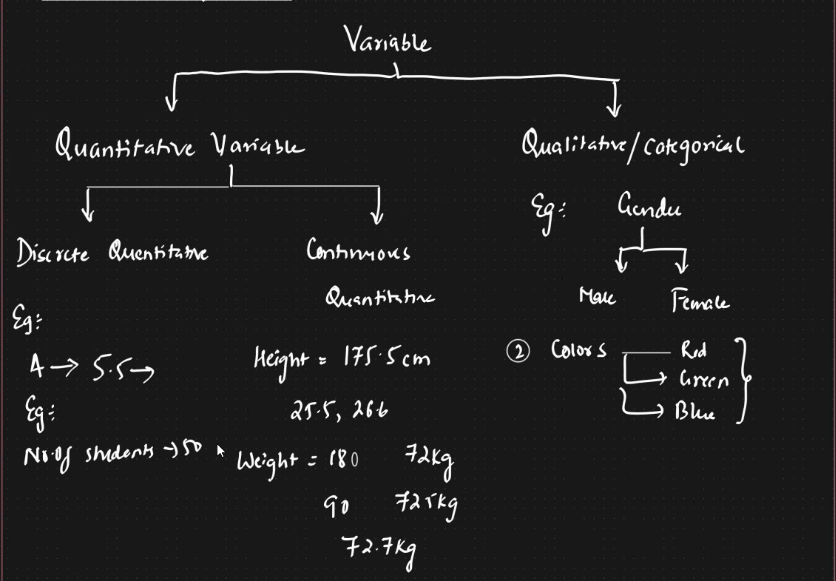
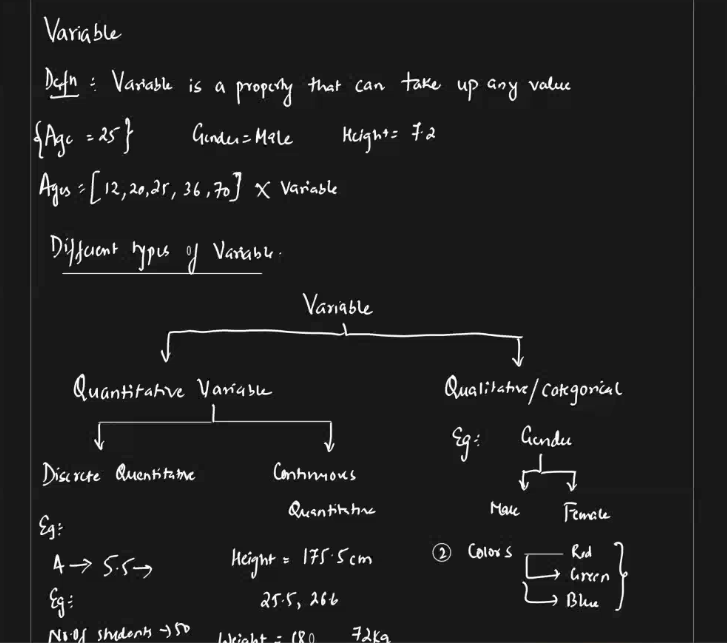
-mode

->MEASURE OF DISPERSION:

-variance:



-standard deviation: 



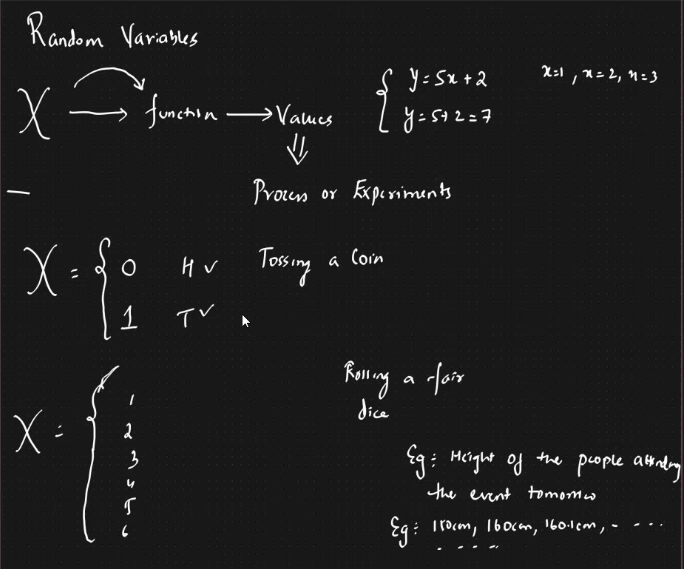
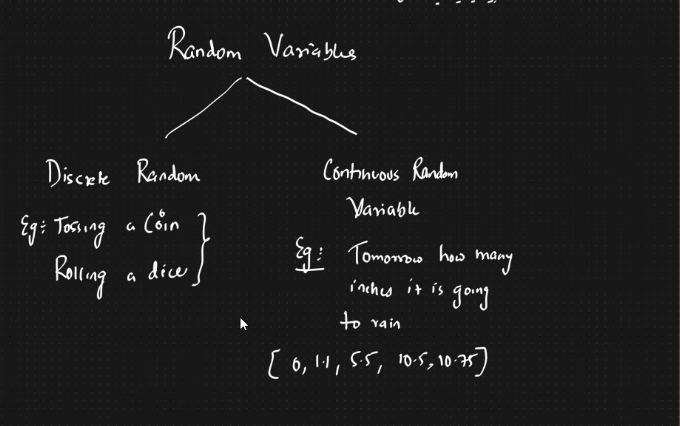
RANDOM VARIABLES:

**random variable** is just a way to represent an outcome that is uncertain or happens by chance.

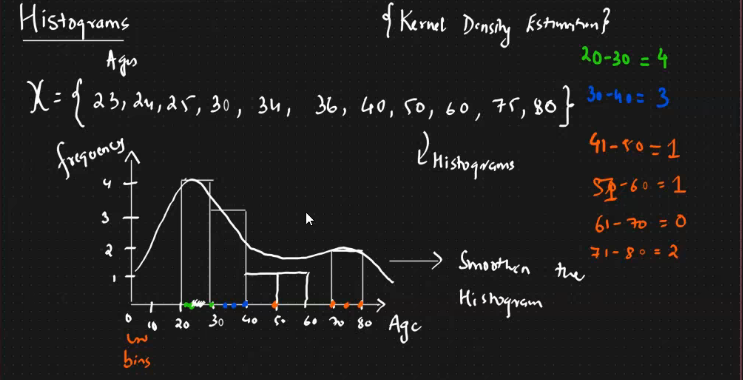
* **For example**, if you roll a die, the result is uncertain. The **random variable** would represent the outcome of that roll (like 1, 2, 3, 4, 5, or 6).

There are two types:

1. **Discrete random variables**: These take specific values (like the roll of a die: 1, 2, 3, etc.).
2. **Continuous random variables**: These can take any value within a range (like your height, which could be any number like 5.7 feet, 5.71 feet, etc.).



-HISTOGRAM:



-> **Kernel Density Estimation (KDE)** is a way to estimate the shape of the distribution of data, like finding out how the data is spread out. It's like creating a smooth curve to represent the data, instead of using bar charts (like histograms).

* **Imagine** you have a set of data points (e.g., test scores), and you want to know how the scores are distributed.
* Instead of just drawing bars like a histogram, KDE draws a smooth curve, making it easier to see the overall pattern of the data.

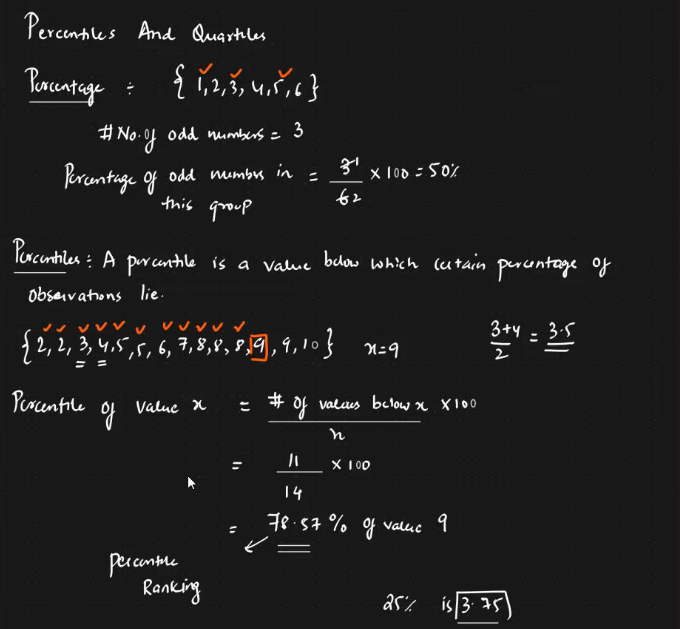
**How it works:**

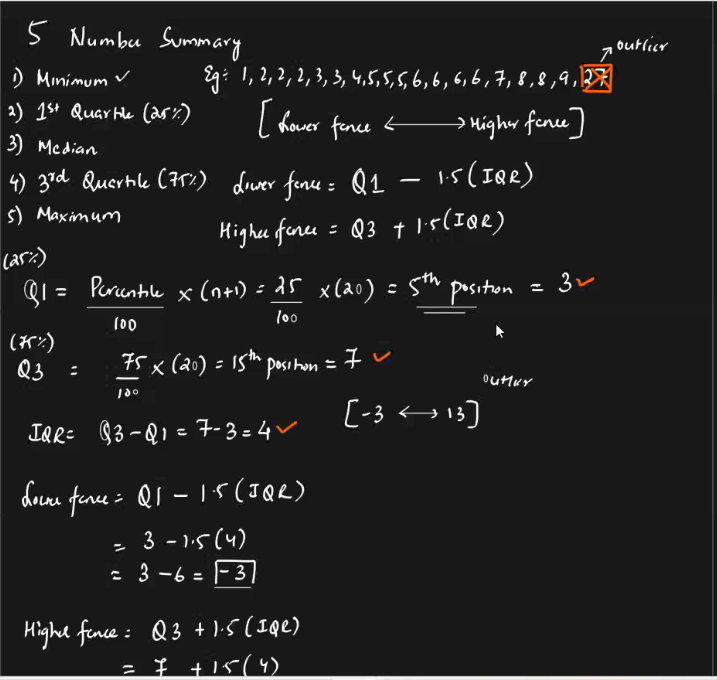
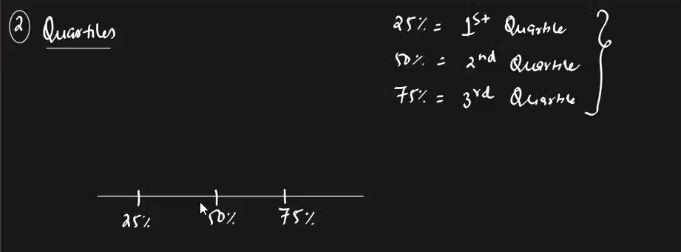
1. KDE uses a "bump" (called a kernel) for each data point. These bumps are smooth and overlap with each other.
2. The size of the bump is controlled by something called **bandwidth**. A smaller bandwidth makes sharp bumps, while a larger bandwidth makes wide, smooth bumps.
3. After placing a bump for each data point, you combine them to get a smooth curve that shows the overall pattern of the data.

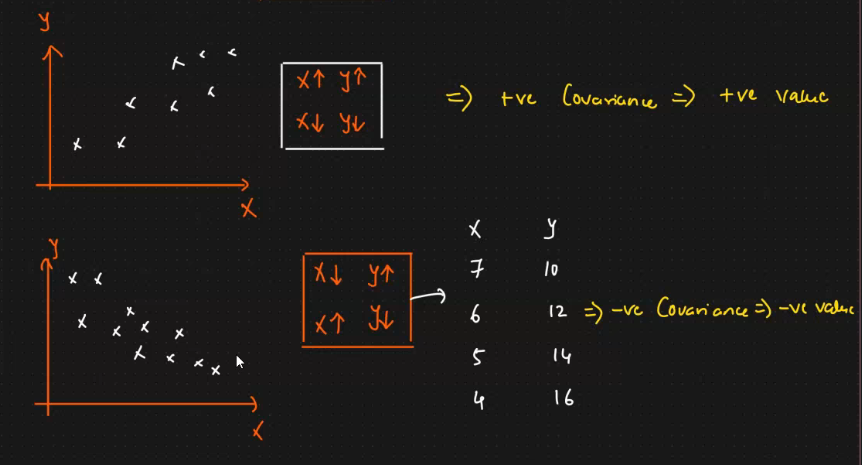
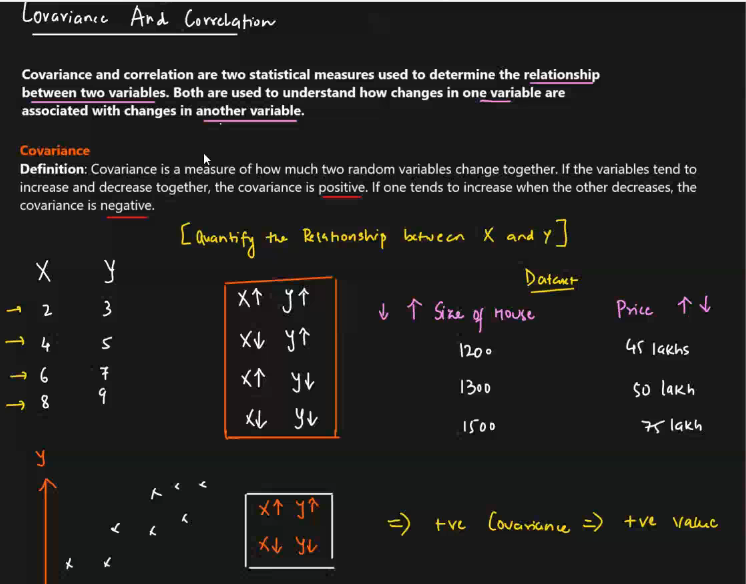
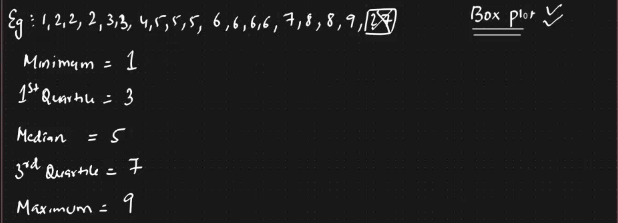
**Why use KDE?**

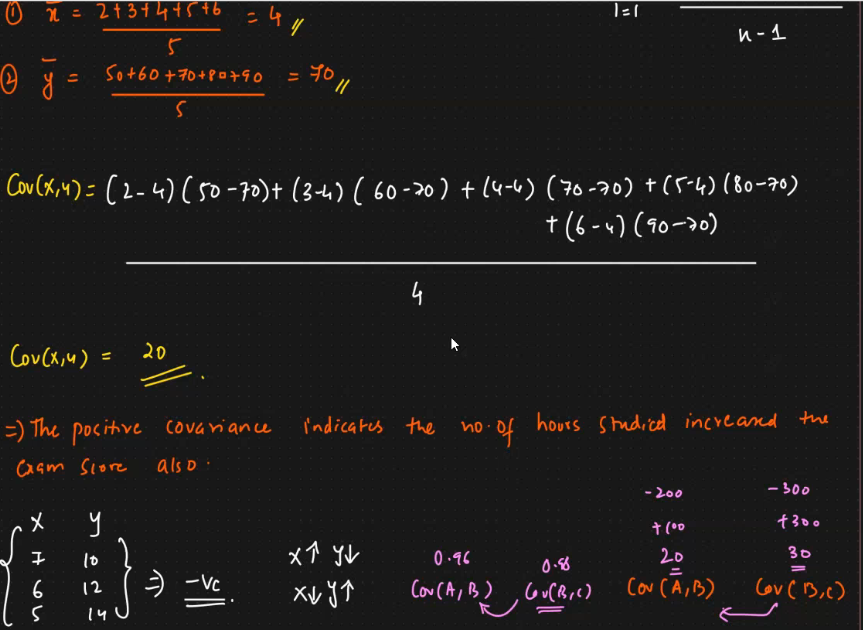
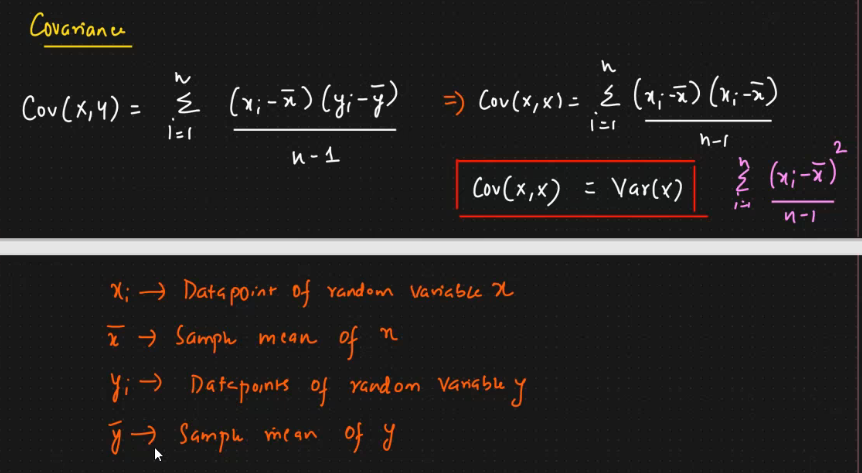
* It gives you a smooth curve instead of a jagged bar chart.
* It’s helpful when you want to understand how data is spread out, especially if the data doesn't follow a simple pattern like a normal distribution.

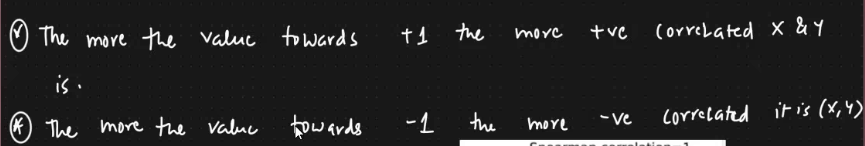
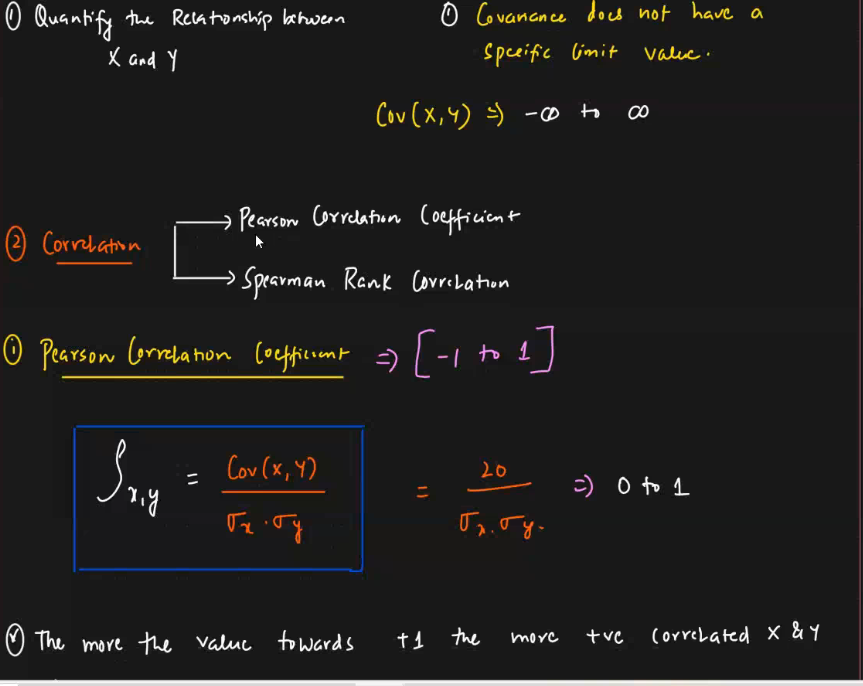
->PERCENTILE AND QUARTILE:





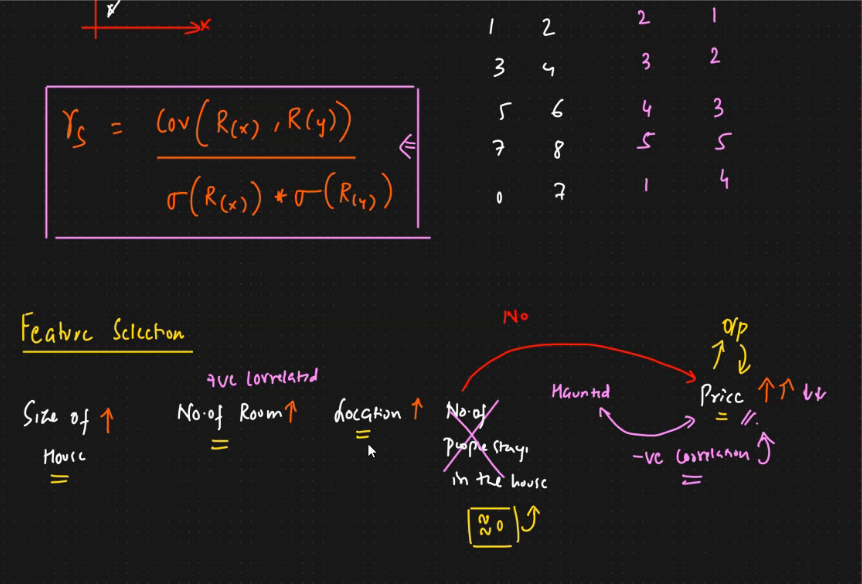


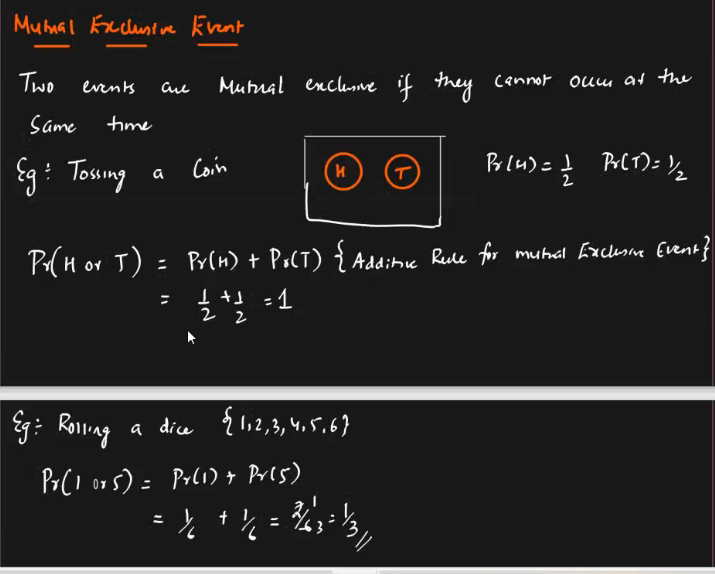
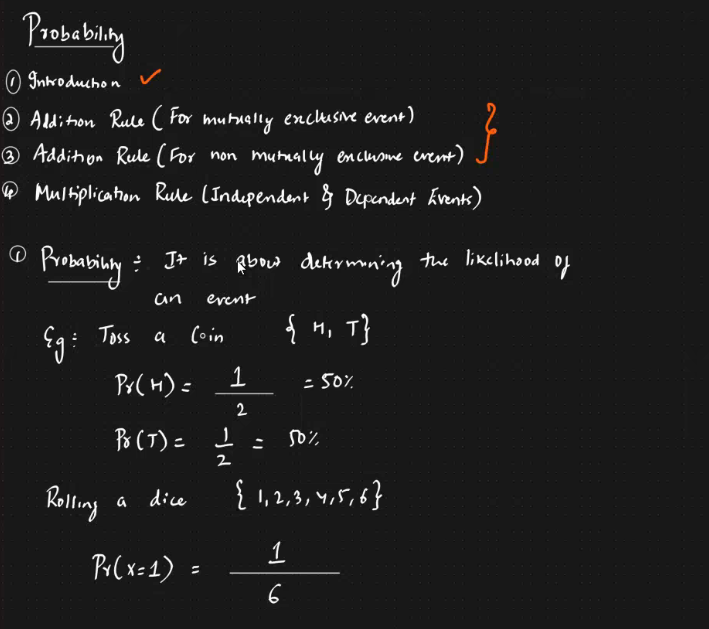


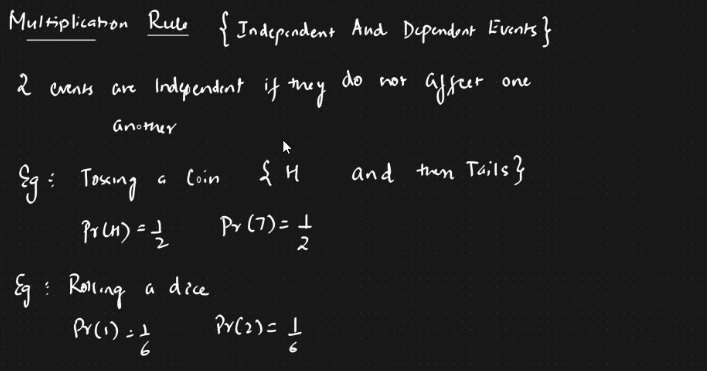
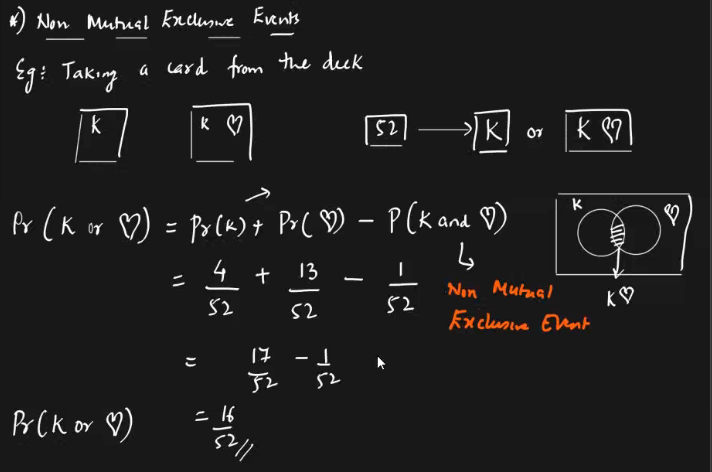


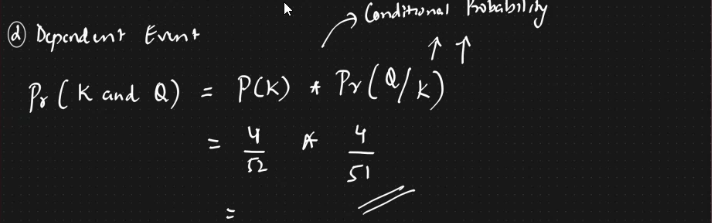
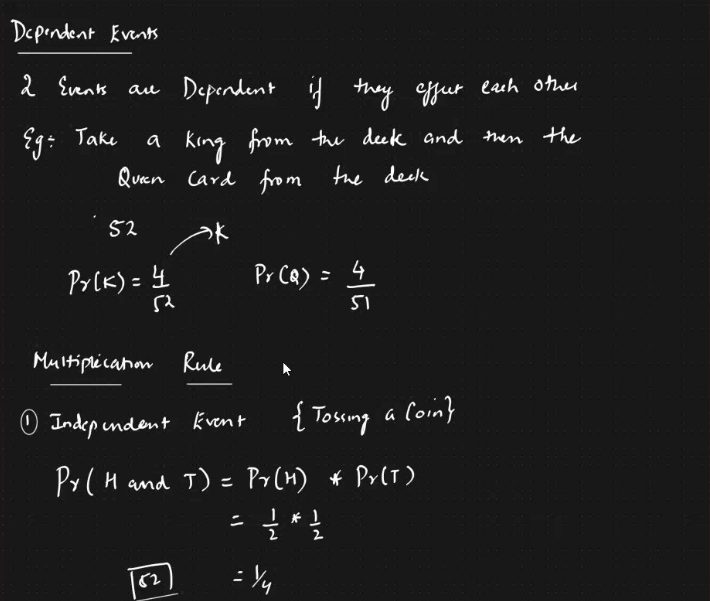
->pearson – linear

->spearman – non-linear









Project:

Apis : flaskapi , streamlit , **FastAPI**