Introduction to Statistics

Statistics/Application

Descriptive
→ Central tendency
→ Dispersion
→ Symmetricity

Inferential

Probability distribution
→ pmf
→ pdf
→ cdf
× CTL
Statistical test
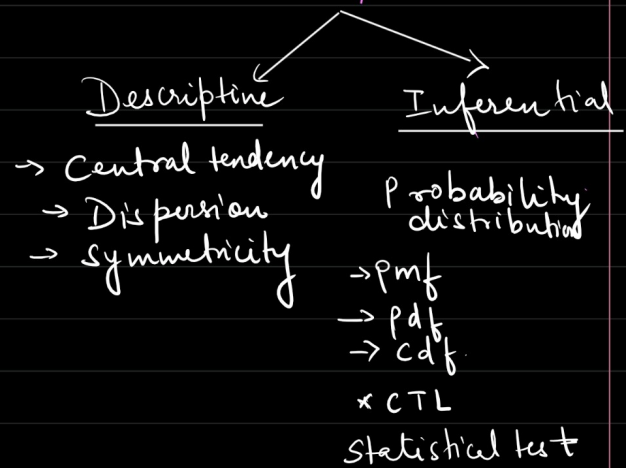
Definition : Statistics is a mathematical science including methods of Collecting, Organizing and analysing data in such a way that meaningful Conclusions can be drawn from them.

Why statistics in data Science?

DS/ML → We try to learn pattern in data

* **data :→** facts or pieces of information that can be stored, measured and re-accessed.

* data is used to bring insights to increase a company's revenue.

Source of data ↑   reaccess the data conveniently ↑

* Collecting, Organising & analysing.

↳ Portal
↳ manually
↳ surveys
↳ Internet clicks
↳ experiments
↳ directly from customers through rating & reviews.

↳ database
   eg SQL
   NoSQL.

↳ finding insights from data using tools like pandas, Numpy, matplotlib, seaborn and so on.

Examples → scores made by Virat Kohli in T-20 world cup
$$[0,1,9,10,\ldots\ldots 74]$$
→ Bumrah Economy rate
→ Rohit Sharma in t-20.
→ Demand of a product for an ecommerce site.
→ Height of students in the class.

Motivation / uses of statistics

① Weather forecast.
② Sports Analytics ⟶ HS, SR, MIN.
③ Election Campaign
④ FMCG | e-commerce.
⑤ medical / genetics / pharmaceutial.