

# Ep 1 - Introduction to Statistics

Nov 26, 2023

## What is Statistics?

Statistics involves methods and techniques for working with data to gain insights, make predictions, and support decision-making. It encompasses both **descriptive statistics** (summarizing and visualizing data) and **inferential statistics** (making inferences or predictions based on data).

## What is data?

Data refers to a collection of facts, information, or values that are often represented in a structured or unstructured form. It serves as the raw material for various processes, including analysis, modelling, and decision-making.

## Where is Statistics Used?

Statistics is widely used in various fields, including:

- **Science and Research:** It's used to analyse experimental data, test hypotheses, and draw scientific conclusions.
- **Business and Economics:** Businesses use statistics for market research, financial analysis, and forecasting.
- **Healthcare:** Statistics is essential for clinical trials, epidemiology, and health data analysis.
- **Quality Control:** Industries use statistics to maintain product quality and process improvement.
- **Social Sciences:** It's used to study human behaviour, demographics, and societal trends.

## What is the use of Statistics in Data Science and Machine Learning?

Statistics is the backbone of data science and machine learning. Here's how it's crucial in these fields:

- **Data Pre-processing:** Before applying machine learning algorithms, you often need to clean and pre-process data. Statistics helps in handling missing values, outliers, and normalizing data.

- **Descriptive Statistics:** You use statistical metrics like mean, median, and standard deviation to understand the characteristics of your data.
- **Inferential Statistics:** In machine learning, you make predictions or draw inferences about a population from a sample. Statistical methods like hypothesis testing and confidence intervals are applied here.
- **Model Evaluation:** To assess the performance of machine learning models, you use statistics, such as accuracy, precision, recall, and F1-score.
- **Feature Selection:** Statistical techniques like correlation analysis help in selecting relevant features for your models.

## What are the types of Statistics?

- Descriptive Statistics
- Inferential Statistics

### Descriptive Statistics

It involves methods and techniques for summarizing, organizing, and presenting data. Their primary purpose is to provide a clear and concise overview of a dataset. It provides a quick snapshot of a dataset, helping you understand its central tendencies and how spread out the values are.

[Descriptive Statistics in Short](#)

### Inferential Statistics

It goes beyond describing data; they help you make inferences, predictions, or decisions about populations based on sample data. It is essential for making data-driven decisions, drawing conclusions from research, and testing hypotheses. They are particularly important in your goal to bridge the gap between AI and healthcare, where making accurate predictions and inferences based on data is critical for improving patient outcomes and medical research.

[Inferential Statistics in Short](#)