

Q1. What is Statistics?

Definition -->

Statistics is a science of collecting,organizing and analyzing the data.

Statistics is used to organize,summarize the data and make inferences about the population.

There are two types of Statistics -->

1. Descriptive Statistics --> It is a type of data which is used to organize and summarize the data.
2. Inferential Statistics --> It is a type in which conclusions are being drawn from the data we have collected, it is mainly used to draw the conclusions.

Statistics is used in a wide variety of fields, including business, economics, medicine, and social sciences.

Most common statistical methods include --> Mean(Avg.),Median(Middle value),Mode(Most occurred value or highest value),Variance(shows the spread of data),Standard deviation(shows that how many data is there in 1 std. or the spread of data about median)

Q2. Define the different types of statistics and give an example of when each type might be used.

1. Statistics is a science of collecting,organizing and summarizing the data.
2. There are two types of Statistics ->
 - A. Descriptive Statistics
 - B. Inferential Statistics
 - a. Descriptive Statistics --> It consists of organizing and summarizing data which contains the methods of Measures of Central Tendency and Measures of Dispersion.
 - i. Measures of Central Tendency --> It contains various methods like finding mean,median,mode.
 - ii. Measures of Dispersion --> It contains various methods like variance,standard deviation.
 - iii. Different types of Graphs --> Histogram,Bar Graph, Pie chart
 - b. Inferential Statistics --> It consists of using data you've measured to form conclusions. It contains the methods like t-test,Hypothesis testing,z-test,P-

value,etc.

When it is used -->

1. Descriptive Statistics is used to organize and summarize data to get results.
2. Inferential Statistics is used to draw conclusions from data measured.

Q3. What are the different types of data and how do they differ from each other? Provide an example of each type of data.

1. There are two types of data -->

A. Quantitative data

B. Qualitative data

- a. Quantitative data --> This is the numerical type of data which is again divided into two types, discrete and continuous data. Quantitative data can be measured.
 - i. Discrete data --> This type of data contains only whole number. For example, no. of children in a family, no. of faculties in a specific college
 - ii. Continuous data --> This type of data can contain any type of value. For example, inches of rainfall, weight of 50 students in a class.
- b. Qualitative data --> This is also called as categorical data. It is divided into two types : Nominal and Ordinal data
 - i. Nominal data --> This type of data is divided into fix categories and divided into fix groups. For example : Gender (Male and Female), Colour of Eyes (Brown, Grey, Black, Green).
 - ii. Ordinal data --> This type of data can be ranked in groups but the difference between them cannot be measured

Q4. Categorise the following datasets with respect to quantitative and qualitative data types: (i) Grading in exam: A+, A, B+, B, C+, C, D, E (ii) Colour of mangoes: yellow, green, orange, red (iii) Height data of a class: [178.9, 179, 179.5, 176, 177.2, 178.3, 175.8,...] (iv) Number of mangoes exported by a farm: [500, 600, 478, 672, ...]

1. Grading in exam: A+, A, B+, B, C+, C, D, E --> Qualitative data (Ordinal data)
2. Colour of mangoes: yellow, green, orange, red --> Qualitative data (Nominal data)
3. Height data of a class: [178.9, 179, 179.5, 176, 177.2, 178.3, 175.8,...] --> Quantitative data (Continuous data)
4. Number of mangoes exported by a farm: [500, 600, 478, 672, ...] --> Quantitative data (Discrete data)

Q5. Explain the concept of levels of measurement and give an example of a variable for each level.

1. There are four scale of measurement of data :

A. Nominal Scale data --> This type of data involves Qualitative/Categorical

data. Here, order does not matter. For example : Survey of 10 people to choose colour :

1. Red -> 5 2. Blue -> 3 3. Yellow -> 2

- B. Ordinal Scale data --> In this type of data, Ranking and Order matters but the difference cannot be measured. For example : Qualification dataset
- P.H.D. - 1st
 - B.E. - 3rd
 - Masters - 2nd
 - B.Com - 4th
 - B.Sc - 5th
- C. Interval Scale data --> In this type of data, Ranking and Order matters as well as the difference can also be measured(excluding Ratio) but the starting point cannot be taken into consideration. For example : Temperature data
- 30 Fahrenheit
 - 45 Fahrenheit
 - 60 Fahrenheit
 - 90 Fahrenheit
- D. Ratio Scale data --> In this type of data, Ranking and Order matters as well as the difference can also be measured and starting point is always zero and Ratio will also be taken into consideration. For example : Grades
- 100,90,60,45,30,15,0

Q6. Why is it important to understand the level of measurement when analyzing data? Provide an example to illustrate your answer.

- There are four types of level of Measurement :
 - Nominal Scale data
 - Ordinal Scale data
 - Interval Scale data
 - Ratio Scale data
- It is important to understand the level of measurement when analyzing data because according to level of measurement it is easy to categorize data and perform operations on them according to its type.
- If we have given a data for example : About Marital status and asked to perform operations on this data, then how could we recognize which operations should be performed on this data, but it is very easy to interpret with the help of level of measurements and we get to know that this is Qualitative(Nominal) type of data, because it is divided into fixed categories as 'Married' and 'Not Married'.
- We know that, Nominal data cannot be ranked and ordered but can be categorised so that, it is easy to perform operations on data depending upon its type and it also saves our time.

Q7. How nominal data type is different from ordinal data type.

- Nominal and Ordinal data are the two types of Qualitative data.
- Qualitative data is the data which can not be measured but can be categorised.
- Nominal data can be categorised into fix no. of categories but can't be ranked or ordered. For example : Marital Status can be categorised into two categories as 'Married' and 'Not Married', etc. Gender can be categorise into categories as 'Male' and 'Female'.
- Ordinal data can be categorised and ranked as well as ordered but the difference between the ranks cannot be measured. For example, Qualification can be ranked based on it's

level in education system, Toppers can be listed according to the percentage they got or CGPA they got in Exams.

Q8. Which type of plot can be used to display data in terms of range?

1. Histogram is the plot that can be used to display data in terms of range. It shows the distribution of data in the set of intervals.
2. The bin size is basically a categorical data.
3. The height of the bars is no. of data points.

Q9. Describe the difference between descriptive and inferential statistics. Give an example of each type of statistics and explain how they are used.

1. There are two types of statistics :
 - A. Descriptive statistics
 - B. Inferential statistics
2. The main difference between Descriptive stats and Inferential stats is descriptive stats is used to organize and summarize data and inferential stats is used to make inferences about population or find conclusions based on sample data.
3. Descriptive statistics contains the methods of measuring Measures of Central Tendency and Measures of Dispersion and Inferential stats contains the methods of measuring like z-test, t-test, Hypothesis testing, Anova test.
4. Descriptive statistics analyzes and interprets the characteristics of a dataset and Inferential statistics uses sample data to make generalizations or predictions about a larger population.
5. Descriptive statistics focuses on entire dataset and Inferential statistics focuses on sample set of data to draw conclusions about entire population.
6. Descriptive statistics is not estimated using sample statistics but Inferential statistics estimated using sample statistics (e.g., sample mean as an estimate of population mean.)
7. Examples of Descriptive statistics are Mean, Mode, Median, Standard deviation, Range, Frequency tables and of Inferential statistics are Hypothesis testing, confidence intervals, regression analysis, ANOVA (analysis of variance) test, chi-square tests, t-tests.

Q10. What are some common measures of central tendency and variability used in statistics? Explain how each measure can be used to describe a dataset.

1. Measures of Central tendency :
 - A. Mean
 - B. Median
 - C. Mode
2. Measures of Variability :
 - A. Range
 - B. Variance
 - C. Standard deviation
3. How each measure can be used to describe a dataset :

A. MEASURES OF CENTRAL TENDENCY

- a. Mean -> The mean is used to find the average of all the values in a dataset.
- b. Median -> The median is used to find the middle value in a dataset.
- c. Mode -> The mode is used to find the most frequent value in the dataset.

B. MEASURES OF VARIABILITY

- a. Range -> Range is used to find the difference between the highest and lowest values in a dataset. It is a simple measure of how spread out the data is.
- b. Variance -> The variance is used to find the measure of how far the values in a dataset are spread out from the mean. Basically, it is used to find out the spread of data.
- c. Standard deviation -> The standard deviation also tells about how the data is spreaded along with median, it is more interpretable measure of variability than the variance, because it is in the same units as the data.

In []:

