6 march assignment

Q1. Statistics is the sense of collecting, organizing, and analyzing data.

Q2. There are two types of statistics:-

- Descriptive statistics: These are used to describe and summarize data. They include measures such as mean, median, mode, and standard deviation.

- Inferential statistics: These are used to make inferences or predictions about a larger population based on a sample of data. They include techniques such as hypothesis testing and regression analysis.

Q3.There are two types of statistical data:-

- Quantitative data: This type of data is numerical and can be measured. It includes variables such as height, weight, and temperature.

- Qualitative data: This type of data is non-numerical and descriptive. It includes variables such as gender, race, and occupation.

Q4. 1.Qualitative data

2.Qualitative data

3.Quantitative data

4.Quantitative data

Q5.

Level of measurement, also known as scale of measurement, is a concept in statistics and data analysis that categorizes variables into different levels based on the nature of the data and the mathematical operations that can be performed on them. There are

four commonly recognized levels of measurement: nominal, ordinal, interval, and ratio. Let's explore each level with examples:

1. **Nominal Level of Measurement:**

   - At this level, data is categorized into distinct categories or groups with no inherent order or ranking.

   - Nominal data can only be compared for equality or inequality.

   - Examples:

     - **Gender:** Categorizing individuals as "Male," "Female," or "Other."

     - **Eye Color:** Grouping people into categories like "Blue," "Brown," "Green," etc.

2. **Ordinal Level of Measurement:**

   - Ordinal data represents categories that have a meaningful order or ranking but have undefined or unequal intervals between them.

   - You can compare ordinal data for relative order or rank, but not for precise differences.

   - Examples:

     - **Education Level:** Ordering educational levels as "High School," "Associate's Degree," "Bachelor's Degree," "Master's Degree," "Doctorate," etc.

     - **Socioeconomic Status:** Categorizing individuals as "Lower Class," "Middle Class," "Upper Class."

3. **Interval Level of Measurement:**

   - Interval data has ordered categories with equal intervals between them, but there is no true zero point.

   - You can perform addition and subtraction on interval data, but meaningful multiplication and division are not applicable.

   - Examples:

     - **Temperature (in Celsius or Fahrenheit):** The difference between 20°C and 30°C is the same as the difference between 30°C and 40°C. However,

there is no absolute zero temperature.

- **IQ Scores:** IQ tests are designed so that the intervals between scores are equal, but there is no true "zero" IQ.

4. **Ratio Level of Measurement:**

- Ratio data has all the properties of interval data, but it also has a true zero point, meaning the absence of the characteristic being measured.

- You can perform all arithmetic operations on ratio data.

- Examples:

    - **Age:** A person's age is ratio data because it has an absolute zero (birth) and you can perform operations like addition, subtraction, multiplication, and division on ages.

    - **Height:** Measuring height in centimeters or inches is ratio data because it has an absolute zero (no height).

Q6.

Imagine you're asking people how much they like your company's service on a scale of 1 to 5, with 1 being really unhappy and 5 being really happy.

Now, if you make a mistake and treat these scores like regular numbers (1, 2, 3, 4, 5) instead of thinking of them as categories (like "not happy," "a bit happy," "moderately happy," "pretty happy," "really happy"), you might add them up and divide by the number of people to find an average. But that average won't really mean much because you're treating these categories like numbers.

For example, if you get an average of 3, you might wrongly think that most people are kind of okay with your service.

In reality, it's better to look at what most people picked as their level of happiness (like "moderately happy," which is 3), and what level makes up the middle (that's the median). This way, you'll get a better picture of how people actually feel about your service.

Understanding this helps you make better decisions for your company. If most people are just "moderately happy," you might want to figure out how to make them "really

happy."

So, knowing the right way to deal with this kind of data is super important to make sense of what people are saying.

Q7. How nominal data types are different from ordinal data type

Nominal data types are different from ordinal data types because nominal data has no inherent order or ranking among categories, while ordinal data does have an order or ranking. Nominal data can only be compared for equality or inequality, while ordinal data can be compared for relative order or rank but not for precise differences. For example, gender is a nominal data type because there is no inherent order or ranking among the categories of "male," "female," and "other," while education level is an ordinal data type because there is a meaningful order or ranking among the categories of "high school," "associate's degree," "bachelor's degree," "master's degree," and "doctorate."

Q8. which type of plots can be used to display data in terms of range

Box plots can be used to display data in terms of range, as they show the minimum value, maximum value, median, and quartiles of a dataset. Another type of plot that can be used is the range plot, which displays the range of values for each group or category in a dataset. Scatter plots can also be used to display the range of values for two variables and show the relationship between them.

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

Ans. Descriptive statistics are used to describe and summarize data, while inferential statistics are used to make inferences or predictions about a larger population based on a sample of data. Descriptive statistics include measures such as mean, median, mode, and standard deviation, and are used to provide a summary of the characteristics of a dataset. Inferential statistics include techniques such as hypothesis testing and regression analysis, and are used to make predictions or draw conclusions about a larger population based on a sample of data.

Example of descriptive statistics: A researcher collects data on the heights of students in a class and calculates the mean, median, and standard deviation of the dataset to describe the characteristics of the data.

Example of inferential statistics: A pollster conducts a survey of 1000 voters to predict the outcome of an upcoming election. They use the data from the sample to make inferences about the preferences of the larger population of voters.

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.

Ans. Some common measures of central tendency include the mean, median, and mode. The mean is the arithmetic average of a dataset and is used to describe the center of the dataset. The median is the middle value of a dataset and is used to describe the typical value of the dataset. The mode is the most common value in a dataset and is used to describe the most frequently occurring value in the dataset.

Some common measures of variability include the range, variance, and standard deviation. The range is the difference between the maximum and minimum values in a dataset and is used to describe the spread of the dataset. The variance is a measure of how much the data deviates from the mean, and the standard deviation is the square root of the variance. Both the variance and standard deviation are used to describe the spread of the dataset, with the standard deviation being the more commonly used measure as it is easier to interpret and is in the same units as the data.