Statistics

**→Question :2.** What type of data does not have a log-normal distribution or a Gaussian distribution? Give proper example

**→Answer:** Categorical data does not typically follow a log-normal distribution or a Gaussian (normal) distribution. Categorical data consists of categories or labels with no inherent order or numerical value. It represents qualitative characteristics rather than quantities.

For example, categorical data includes types of animals (e.g., dog, cat, bird), colors (e.g., red, blue, green), or educational levels (e.g., high school, college, graduate school).

Unlike numerical data, categorical data cannot be meaningfully described using mathematical distributions like the log-normal or Gaussian distribution. These distributions are suited for continuous numerical data, not discrete categories.

**→Question :3.** What is the meaning of the five-number summary in Statistics? Give proper example

**→Answer :** The five-number summary in statistics provides a concise summary of the distribution of a dataset using five key values:

1] The minimum value (the lowest value)

2] 25th Percentile or Q1

3] 50th Percentile or Q2 or Median

4] 75th Percentile or Q3

5] Maximum Value (the highest value)

Example -

Consider the following dataset representing the scores of students in a test:

65,72,78,82,85,88,90,92,95,98

The five-number summary for this dataset would be:

* Minimum: 65
* Q1 (First Quartile): 77.5 (average of 72 and 78)
* Median (Second Quartile): 85
* Q3 (Third Quartile): 91 (average of 90 and 92)
* Maximum: 98

This summary provides a quick overview of the spread of scores, showing that the majority of scores fall between 77.5 and 91, with a median score of 85.It gives insights into the central tendency, spread, and skewness of the data.

**→Question :4.**What is correlation? Give an example with a dataset & graphical representation on jupyter Notebook

**→Answer :** Correlation is a statistical measure that describes the extent to which two variables change together. It indicates the strength and direction of the relationship between the variables. A correlation coefficient value close to 1 indicates a strong positive correlation, close to -1 indicates a strong negative correlation, and close to 0 indicates little to no correlation.

Link → https://colab.research.google.com/drive/1fQEUobAM87uEkFB4OmUm\_7cc5a9BpdrV?usp=sharing