

Quiz

Basic Data Statistics

Question 1:

You're analyzing a dataset and find that the skewness of a particular variable is -1.5. What does this indicate about the distribution of the data?

- A) The distribution is positively skewed.
- B) The distribution is negatively skewed.
- C) The distribution is symmetric.
- D) Skewness cannot be determined from this value.

Question 2:

You're performing data analysis and notice that the correlation coefficient between two variables is 0.87. What does this value imply about the relationship between the variables?

- A) There is a strong positive linear relationship between the variables.
- B) There is a strong negative linear relationship between the variables.
- C) There is a weak positive linear relationship between the variables.
- D) There is no linear relationship between the variables.

Question 3:

You're comparing the spread of two datasets. Dataset A has a standard deviation of 15, and Dataset B has a standard deviation of 5. Which of the following statements is correct?

- A) Dataset A has less spread than Dataset B.
- B) Dataset A has more spread than Dataset B.
- C) Both datasets have the same spread.
- D) The spread cannot be determined from standard deviations alone.

Answers:

1. B) The distribution is negatively skewed.
2. A) There is a strong positive linear relationship between the variables.

3. B) Dataset A has more spread than Dataset B.

Exploratory Data Analysis Quiz Questions

Question 1:

Which of the following visualization techniques is best suited for identifying potential outliers and understanding the spread and distribution of data?

- A) Line Plot
- B) Box Plot
- C) Scatter Plot
- D) Histogram

Question 2:

In the context of Exploratory Data Analysis, what does the term "Skewness" refer to?

- A) The measure of the degree of linearity between two variables.
- B) The measure of the symmetry of a distribution around its mean.
- C) The measure of the spread of data points around the mean.
- D) The measure of the correlation between two categorical variables.

Question 3:

You're analyzing a dataset with several missing values. What would be the most appropriate action during the Exploratory Data Analysis phase?

- A) Remove all rows with missing values to avoid distortion.
- B) Replace missing values with the median of the entire dataset.
- C) Impute missing values based on domain knowledge or statistical methods.
- D) Fill in missing values with zero to maintain the overall structure.

Answers:

1. B) Box Plot
2. B) The measure of the symmetry of a distribution around its mean.
3. C) Impute missing values based on domain knowledge or statistical methods.

