mod2_precheck

1. Which of the following describes the mean?

    - ☐ The most frequently occurring number in a set.
    - ☐ The middle value in a set of numbers.
    - ☐ The difference between the highest and lowest numbers in a set.
    - ☑ The average of all numbers in a set.

2. In which of the following scenarios is the median equal to the mode?

    - ☑ ${1,2,3,5,6,6,7}$
    - ☐ ${3,4,4,5,6}$
    - ☐ ${3,4,5,6,7,7}$
    - ☐ ${2,2,4,8,10}$

3. Which of the following best describes the first quartile (Q1) of a dataset?

    - ☐ The value below which 50% of the data falls
    - ☐ The average of all data points
    - ☑ The value below which 25% of the data falls
    - ☐ The value below which 75% of the data falls

4. What does the Interquartile Range (IQR) measure?

    - ☐ The difference between the maximum and minimum data value
    - ☑ The range of the middle 50% of data
    - ☐ The range of the entire dataset
    - ☐ The average distance between Q1 and Q3

5. Which of the following best defines an outlier?

    - ☐ The average of two medians
    - ☐ The median of a dataset
    - ☐ A data point that fits well within the data range
    - ☑ A data point that deviates significantly from other data points in a dataset

6. In a box plot, what does the line inside the box represent?

    - ☐ Maximum value
    - ☐ First quartile (Q1)
    - ☐ Third quartile (Q3)
    - ☑ Median

7. Which of the following defines the 'interquartile range' in the context of a box plot?

    - ☐ The distance between the median and the nearest end of the box
    - ☐ The distance between the two whiskers
    - ☐ The width of the entire plot
    - ☑ The width of the box itself

8. In a survey, if you ask respondents their marital status (single, married, divorced), which kind of data are you collecting?

    - ☐ Continuous
    - ☑ Categorical
    - ☐ Ratio
    - ☐ Interval

9. What is a primary use of a bar plot?

    - ☐ To display the distribution of a continuous variable
    - ☑ To compare the frequency or other measures of multiple categorical variables
    - ☐ To show the relationship between two continuous variables
    - ☐ To depict the progression of data over time

10. Histograms are used primarily for which type of data?

    - ☐ Categorical
    - ☐ Qualitative
    - ☑ Continuous
    - ☐ Binary

11. In a contingency table, what does the intersection of a row and a column represent?

    - ☑ The frequency of occurrence of the combined categories represented by that row and column
    - ☐ The sum of the row and the column
    - ☐ The average of the row and the column
    - ☐ The total number of observations

12. Which of the following is true about variance and standard deviation?

    - ☑ Variance is the average of squared differences from the mean, and standard deviation is its square root
    - ☐ Standard deviation is the average of squared differences from the mean, and variance is its square root
    - ☐ Variance is the square root of the mean, and standard deviation is the square of the mean
    - ☐ Variance and standard deviation are the same

mod2_review

1. Which of the following statistics is considered robust because it is least affected by outliers?

    - ☐ Variance
    - ☐ Mode
    - ☐ Mean
    - ☑ Median

2. In a dataset, if the value of the second quartile (Q2) is 50 and the value of the third quartile (Q3) is 80, which of the following must be true?

- ☐ 75% of the data values are less than 50
- ☐ 50% of the data values are between 50 and 80
- ☐ The maximum value in the dataset is 80
- ☑ 25% of the data values are greater than 80

3. Why is the IQR considered a robust measure of variability?

- ☑ It is unaffected by extreme values or outliers
- ☐ It uses all values in the dataset
- ☐ It is always smaller than the range
- ☐ It equals the range divided by 2

4. How can the IQR be useful in identifying outliers in a dataset?

- ☐ Data points outside Q1 + 1.5(IQR), Q3-1.5(IQR) are considered outliers
- ☑ Data points outside Q1-1.5(IQR), Q3 + 1.5(IQR) are considered outliers
- ☐ Data points outside Q1, Q3 are considered outliers
- ☐ Data points inside Q1, Q3 are considered outliers

5. In a box plot, which of the following statements is true?

- ☐ The line inside the box always represents the mean of the dataset
- ☐ The box represents the range of all data points in the dataset
- ☑ Whiskers extend to cover typical data points, but they may not reach the maximum and minimum values
- ☐ Outliers are always plotted within the box

6. Why might outliers be retained in a dataset for analysis?

- ☐ They help in making the dataset appear larger
- ☐ They enhance the aesthetic appeal of visual representations
- ☐ They are always errors and should never be retained
- ☑ They can provide insight into rare events or phenomena

7. Why is it important to identify categorical data before analysis?

- ☐ Because it follows a bell curve
- ☐ Because it can always be converted to numerical data
- ☑ Because it requires different statistical methods than numerical data
- ☐ Because it always has missing values

8. How does a stacked bar plot represent data?

- ☐ It visualizes the distribution of a single continuous variable
- ☐ It plots the relationship between continuous variables
- ☑ It displays the categories on the x-axis and their frequencies or other measures on the y-axis with each bar divided into subcategories
- ☐ Each bar depicts the median of each category

9. Which statement about histograms is correct?

- ☐ All histograms must have filled bars
- ☐ The width of the bars in a histogram can vary when the bin intervals are the same
- ☐ Hollow histograms can only represent data with negative values
- ☑ The horizontal axis of a histogram represents the bins of data

10. What is typically found in the margins of a contingency table?

- ☑ Marginal totals (or marginal sums)
- ☐ Median Values
- ☐ Zeros
- ☐ Variables

11. Which of the following transformations will not change the value of variance for a dataset?

- ☑ Taking the logarithm of each data point.
- ☐ Squaring each data point
- ☐ Multiplying each data point by a constant
- ☐ Adding a constant to each data point