

Quiz 2 - Data Exploration

1. Which of these statements is true about samples and variables?

- A sample is an instance or example of an entity in your data.
- **All of these statements are true.**
- A sample can have many variables to describe it.
- A variable describes a specific characteristic of an entity in your data.

2. Other names for 'variable' are

- categorical, nominal
- **feature, column, attribute**
- sample, row, observation
- numerical, quantitative

3. What is the purpose of exploring data?

- **To gain a better understanding of your data.**
- To gather your data into one repository.
- To digitize your data.
- To generate labels for your data.

4. What are the two main categories of techniques for exploring data?

Choose two.

- Histogram
- Outliers
- **Visualization**
- Trends
- Correlations
- **Summary statistics**

5. Which of the following are NOT examples of summary statistics?

- mean, median, mode
- **data sources, data locations**
- standard deviation, range, variation
- skewness, kurtosis

6. What are the two measures for measuring shape as mentioned in the lecture? Choose two.

- **Kurtosis**
- **Skewness**
- Contingency Table
- Range
- Mode

7. Which of the following would NOT be a good reason to use a box plot?

- To show and compare distribution values
- To show data distribution shapes such as asymmetry and skewness.
- **To show correlations between two variables.**

8. All of the following are true about data visualization EXCEPT

- **Is more important than summary statistics for data exploration**
- Should be used with summary statistics for data exploration.
- Is useful for communicating results.
- Provides an intuitive way to look at data.