**Task 1: Answer the following**

1. Which of the following is NOT a common data encoding technique?

a) One-hot encoding

b) Label encoding

c) Transform encoding

d) Ordinal encoding

2. For a dataset with the 'color' column containing three unique values 'red', 'green', and 'blue', how many columns will result from one-hot encoding this column?

a) 1

b) 2

c) 3

d) 4

3. Which feature scaling method is sensitive to outliers?

a) Min-Max scaling

b) Standardization (Z-score normalization)

c) Robust scaling

d) L2 normalization

4. In the standardization method, what are the mean and standard deviation of the standardized feature?

a) mean = 0, standard deviation = 1

b) mean = 1, standard deviation = 0

c) mean = 1, standard deviation = 1

d) mean = 0, standard deviation = 0

5. Which visualization tool is best suited for showing the distribution of a single continuous variable?

a) Bar chart

b) Pie chart

c) Histogram

d) Scatter plot

6. Which chart type is ideal for visualizing hierarchical data structures?

a) Line chart

b) Treemap

c) Area chart

d) Bar chart

7. If a distribution has a long tail on the right side, it is said to be:

a) Negatively skewed

b) Positively skewed

c) Normally distributed

d) Uniformly distributed

8. The process of transforming raw data into a format that is both understandable and usable by machine learning models is known as:

a) Feature selection

b) Feature extraction

c) Feature transformation

d) Feature engineering

9. Which technique can be used to reduce the dimensionality of the dataset while preserving as much variance as possible?

a) One-hot encoding

b) Principal Component Analysis (PCA)

c) SMOTE

d) Min-Max scaling

**Programming Assignment: Exploring and Transforming the Titanic Dataset**

**Dataset**: You can use the famous Titanic dataset available on platforms like Kaggle.

**Data Loading and Exploration:**

* Load the Titanic dataset into a pandas DataFrame.
* Display the first 5 rows to understand the dataset's structure.

**Encoding:**

One-hot Encoding:

* Apply one-hot encoding on the 'Sex' column (which has values 'male' and 'female').
* This should result in two new columns 'male' and 'female' with binary values.

Ordinal Encoding:

* Create a new column 'Pclass\_encoded' by encoding the 'Pclass' column (1 = First Class, 2 = Second Class, 3 = Third Class) using ordinal encoding.

**Standardization**:

Apply standardization on the 'Age' column to create a new column named 'Age\_standardized'.

**Data Visualization using Matplotlib:**

* Histogram: Plot a histogram for the 'Age\_standardized' column to visualize the distribution of ages.
* Bar Plot: Create a bar plot that displays the number of passengers for each 'Pclass'.
* Pie Chart: Create a pie chart representing the male-female distribution on the ship.

**Submission Guidelines:**

Ensure your code is well-commented, explaining each major step, especially during the encoding and standardization process.

Include a brief write-up (around 200 words) explaining your observations from the visualizations you generated.