**Steps**

**Step 1: Load and Inspect the Data**

1. Import the necessary libraries.
2. Load the Titanic dataset.
3. Display:
   * The first few rows of the dataset.
   * Basic information about the dataset (e.g., columns, data types, and null values).
   * Summary statistics for numerical columns.

**Step 2: Handle Missing Values**

1. Identify columns with missing values.
2. Handle missing values:
   * For numerical columns.
   * For categorical columns.
3. Provide a brief explanation of your chosen method for handling missing data.

**Step 3: Check for Duplicates**

1. Check if there are duplicate rows in the dataset.
2. If duplicates exist, remove them.
3. Provide a summary of the cleaned dataset.

**Step 4: Feature Extraction**

1. Focus on the Cabin column:
   * Extract the first letter of the cabin (e.g., C85 → C).
   * Treat missing cabins as a separate category.
2. Create a new column for the extracted cabin letter.

**Step 5: Encode Categorical Variables**

1. Identify columns with categorical data.
2. Convert categorical data into a format suitable for analysis:
   * Encode categorical features into numerical values.

**Step 6: Visualize the Data**

Use **matplotlib** to create the following visualizations. Ensure that each plot is clear and well-labeled, with a title, axis labels, and legends where necessary.

1. **Line Plot**:
   * Plot the number of passengers across different ages.
   * X-axis: Age
   * Y-axis: Count of passengers
2. **Bar Plot**:
   * Compare the number of passengers for each Pclass (Passenger Class).
   * X-axis: Passenger Class
   * Y-axis: Count of passengers
3. **Scatter Plot**:
   * Show the relationship between Age and Fare.
   * X-axis: Age
   * Y-axis: Fare
   * Color-code points based on whether the passenger survived (Survived).
4. **Histogram**:
   * Plot the distribution of ages.
   * Use appropriate bins (e.g., 10 or 20).
   * X-axis: Age
   * Y-axis: Count of passengers
5. **Bar Plot for Extracted Cabin Features**:
   * Show the distribution of passengers across the extracted cabin letters.
   * X-axis: Cabin Letter
   * Y-axis: Count of passengers
   * Add a bar for missing values as a separate category.

**Step 7: Draw Insights**

1. Summarize your observations based on the visualizations.
   * Example: "Most passengers belonged to the third class, and the majority of passengers with missing cabin data did not survive."
2. Highlight any trends or correlations in the data.