**Perform following task on your own selected dataset:**

1. **Data Summarization**:
   * Calculate basic summary statistics (mean, median, standard deviation, etc.) for each numerical variable.
   * Count the frequency of unique values for categorical variables.
   * Calculate the number of missing values for each variable.
2. **Data Visualization**:
   * Create histograms or density plots to visualize the distribution of numerical variables.
   * Generate bar plots or pie charts to visualize the distribution of categorical variables.
   * Create box plots to identify outliers and understand the spread of data.
   * Construct scatter plots to explore relationships between pairs of variables.
   * Use heatmaps to visualize correlations between variables.
3. **Handling Missing Data**:
   * Explore the patterns of missing data across variables.
   * Decide on an appropriate strategy for handling missing values (imputation, removal, etc.).
4. **Outlier Detection and Treatment**:
   * Identify and visualize outliers in numerical variables.
   * Decide whether to remove, transform, or treat outliers based on domain knowledge and analysis goals.
5. **Feature Engineering**:
   * Create new features or transformations of existing features that might be more informative.
   * Consider scaling or normalizing features if necessary.
6. **Data Distribution Analysis**:
   * Test the normality of numerical variables using statistical tests (e.g., Shapiro-Wilk test).
   * Visualize the data distribution and assess skewness and kurtosis.
7. **Bivariate Analysis**:
   * Analyze relationships between pairs of variables through scatter plots, regression analysis, or correlation matrices.
   * Investigate how one variable changes with respect to another (e.g., time series analysis).
8. **Grouping and Aggregation**:
   * Group data by categorical variables and calculate summary statistics within each group.
   * Explore differences or patterns between different groups.
9. **Time Series Analysis**:
   * Decompose time series data into trend, seasonality, and residual components.
   * Plot time series data over time and look for trends and patterns.
10. **Dimensionality Reduction**:
    * Perform dimensionality reduction techniques like Principal Component Analysis (PCA) to visualize high-dimensional data.
    * Select relevant features for modeling using feature selection methods.
11. **Data Transformation**:
    * Apply mathematical transformations (e.g., logarithmic or exponential transformations) to normalize data.
    * Convert categorical variables to numerical format using encoding techniques.