Problem Statement

Problem 1

1 - Define the problem and perform exploratory Data Analysis

* Check shape, Data types, statistical summary
* Univariate analysis
* Multivariate analysis
* Use appropriate visualizations to identify the patterns and insights
* Key meaningful observations on individual variables and the relationship between variables.

2 - Data Pre-processing (Prepare the data for modelling):

* Missing Value Treatment (if needed)
* Outlier Detection (treat, if needed)
* Feature Engineering
* Encode the data
* Train-test split

3 - Model Building

* Linear regression
* Apply linear Regression using Sklearn
* Using Stats models Perform checks for significant variables using the appropriate method
* Create multiple models and check the performance of Predictions on Train and Test sets using R-square, RMSE & Adjust R-square.

=====================================================================

Problem 2

1 - Define the problem and perform exploratory Data Analysis

* Problem definition
* Check shape, Data types, statistical summary
* Univariate analysis
* Multivariate analysis
* Use appropriate visualizations to identify the patterns and insights
* Key meaningful observations on individual variables and the relationship between variables

2 - Data Pre-processing (Prepare the data for modelling):

* Missing value Treatment (if needed)
* Outlier Detection (treat, if needed)
* Feature Engineering (if needed)
* Encode the data
* Train-test split

3 - Model Building and Compare the Performance of the Models

* Build a Logistic Regression model
* Build a Linear Discriminant Analysis model
* Build a CART model
* Prune the CART model by finding the best hyperparameters using GridSearch
* Check the performance of the models across train and test set using different metrics
* Compare the performance of all the models built and choose the best one with proper rationale

4 - Business Insights & Recommendations: