Logistic Regression Practice: Loan Prediction Dataset

This exercise is based on a Loan Prediction dataset. The goal is to build and evaluate a Logistic Regression model to predict loan approval status. Students will perform data preprocessing, exploratory analysis, model training, and evaluation step by step.

# Part A: Questions

* 1. Load the dataset and inspect its structure (shape, columns, datatypes, first 5 rows).
* 2. Perform data cleaning: handle missing values, encode categorical variables, drop irrelevant features.
* 3. Perform Exploratory Data Analysis (EDA): summary statistics, countplots/histograms, correlation heatmap.
* 4. Define the target variable (Loan\_Status) and independent variables.
* 5. Split the dataset into training (70%) and testing (30%) sets.
* 6. Train a Logistic Regression model on the training set.
* 7. Make predictions on the test set.
* 8. Evaluate the model using Accuracy, Precision, Recall, F1 Score, and Confusion Matrix.
* 9. Interpret coefficients: which features influence loan approval most strongly?

# Part B: Answers with Code & Explanation