**Lab Guide:**

**Lab Exercise 1: Data Analysis**

Objective: generate insights across different types of analytics.

Step 1: Descriptive Analytics – Understanding the Data

Prompt:

* List all columns with their data types.
* Identify missing values in the dataset.
* Show summary statistics (mean, median, min, max) for numerical columns.
* Count the number of approved (Loan\_Status=Y) and rejected (Loan\_Status=N) loans.
* Create a bar chart for the distribution of Property\_Area (Urban, Rural, Semiurban).

Step 2: Diagnostic Analytics – Identifying Patterns & Relationships

Prompt:

* Compare Applicant Income, Coapplicant Income, and Credit\_History for both groups.
* Is income a bigger factor than education in determining loan amount?
* Do graduates apply for specific loan types (home loans vs. personal loans) that affect amounts?
* Compare **credit history distribution** for both groups.
* Are self-employed applicants more likely to have a low Credit\_History score?
* Do self-employed applicants request **higher loan amounts** with lower declared incomes?
* Is there a strict **approval threshold** based on Credit\_History scores?
* Does income level compensate for poor credit history in approvals?
* Are there **exceptions** where people with poor credit history still get approved?
* Does the loan amount depend more on **property area** (urban vs. rural) rather than income?
* Are **loan term restrictions** (shorter terms for high-income applicants) affecting loan amounts?
* Do some high-income applicants have **co-applicants with low incomes**, reducing eligibility?
* Do married applicants apply with a co-applicant, increasing approval chances?
* Are married applicants more likely to have stable income sources?
* Do banks prefer married applicants due to lower default risks?

**Step 3: Predictive Analytics – Forecasting Loan Approval**

Goal: Use historical data to predict outcomes.

* Predict loan approval (Loan\_Status) based on income, credit history, and dependents.
* What factors most influence loan approval?
* Train a logistic regression model to predict Loan\_Status and show accuracy.
* Create a decision tree to classify whether a loan will be approved or not.
* Find the probability of approval for an applicant with ₹50,000 income, 2 dependents, and a credit score of 1.

**Step 4: Prescriptive Analytics – Recommending Actions**

Goal: Suggest actionable strategies.

* Recommend strategies to increase loan approval rates.
* What steps can applicants take to improve their approval chances?
* Suggest an automated rule to approve loans based on income and credit history.
* Identify the optimal loan amount for different income groups to minimize defaults.
* What changes in lending policy could improve loan acceptance rates by 10%?

**Explore Further and Share your insights!!!**