**Loan Approval Analysis**

## **Introduction:**

The loan approval process involves evaluating multiple factors to determine whether an applicant qualifies for a loan. This project analyzes a dataset containing demographic, income, loan amount, and credit history information to identify patterns influencing loan approval decisions. The goal is to provide financial institutions with a predictive model for data-driven decision-making.

## **Dataset Overview:**

The dataset contains the following key features:

* Numerical: ApplicantIncome, CoapplicantIncome, LoanAmount, Loan\_Amount\_Term, Credit\_History
* Categorical: Gender, Married, Dependents, Education, Self\_Employed, Property\_Area, Loan\_Status

### **Initial Data Exploration**

* Descriptive Statistics:
  + Summary of numerical features such as income and loan amounts.
  + Frequency distribution of categorical variables.
* Missing Values:
  + Identified missing values in columns such as Gender, Dependents, Self\_Employed, LoanAmount, Loan\_Amount\_Term, and Credit\_History.

## 

## 

## **Data Cleaning**

### **Handling Missing Values**

* Categorical Columns:
  + Gender, Married, and Self-Employed were filled with "Unknown".
  + Numerical Columns:
  + LoanAmount: Filled with the median value.
  + Loan\_Amount\_Term and Credit\_History: Filled with 0.
  + Dependents:

-Replaced symbols (e.g., +) and filled missing values with 0.

### **Result:**

* Missing values were successfully addressed, ensuring data consistency for further analysis.

**Initial Setup**:

* Importing essential libraries like pandas, numpy, seaborn, and matplotlib.
* Warning filters to suppress unnecessary warnings.

**Loading the Dataset**:

* Loading the dataset using pd.read\_csv().

**Data Exploration**:

* A preview of the dataset using df and statistical summary using df.describe().
* Analysis of categorical columns by calculating like Gender, Married, Dependents, Education, Self\_Employed, and Property\_Area

## **Analysis:**

### **1. Demographic Analysis**

-Visualizing the following column by loan status using a bar plot

* Gender
* Marital Status
* Education
* Self-Employment
* Credit history
* Loan\_Amount\_Term

-Analyzing Dependents with Loan\_Status using groupby

**2. Income and Loan Amount Analysis**

* Visualizing the relationship between applicant income and loan approval using boxplots
* Visualizing the relationship between applicant income and loan approval using boxplots
* correlation analysis between applicant income, co-applicant income and loan amount using Heat Map
* Analyzing gender, education, and marital status with loan status using groupby.

### **3. Credit History and Loan Term Analysis**

* + Interaction between credit history and loan term revealed complementary trends using pivot table

### **4. Property Area Analysis**

* Analyze the distribution of loan approvals across different property areas (Urban, Semiurban, Rural) using groupby with aggregate function mean() a and then plotting its bar graph.

## **Conclusion**

Key factors influencing loan approvals include credit history, applicant income, and property area. Demographics such as gender, marital status, and education also play important roles. Financial aspects like loan amount and term significantly impact approval rates, with higher amounts and longer terms generally leading to fewer approvals.