



Loan_Approval_Analysis

Data Analyst Project

By Laxmi Chhawari

June 2025

Data Set

Column Name	Description
Loan_ID	Unique ID for each loan application
Gender	Applicant's gender (Male/Female)
Married	Marital status (Yes/No)
Dependents	Number of dependents (0, 1, 2, 3+)
Education	Education level (Graduate/Not Graduate)
Self_Employed	Employment type (Yes/No)
ApplicantIncome	Applicant's monthly income
CoapplicantIncome	Co-applicant's monthly income
LoanAmount	Loan amount (in thousands)
Loan_Amount_Term	Loan term (in months)
Credit_History	Credit history (1 = Good, 0 = Bad, sometimes NaN)
Property_Area	Area type (Urban/Semiurban/Rural)
Loan_Status	Whether the loan was approved or not (Y/N)

 **Records:** 614 entries (loan applications)

 **DATASET LINK:** <https://www.kaggle.com/datasets/altruistdelhite04/loan-prediction-problem-dataset>

 **SOURCE:** The data is extracted from Kaggle, which is popular for its public datasets.

PROJECT OVERVIEW:

This project involved analyzing loan application data using Excel, SQL, and Power BI to identify approval trends and optimize decision-making criteria. The dataset included 300+ records with variables like Gender, Income, Education, Self-Employment, etc.

PROBLEM STATEMENT:

To uncover patterns in loan approval decisions and recommend data-driven strategies that reduce loan default risk while maintaining fair approval practices.

OBJECTIVE:

- Identify factors that affect loan approval
- Analyze applicant characteristics (income, education, employment, etc.)
- Use SQL and Power BI to find patterns and communicate insights visually
- Recommend data-driven solutions to improve loan disbursal strategy

TOOLS USED:

Tool	Purpose
Excel	Data cleaning, formatting, CSV conversion
SQL (MySQL)	Data analysis using queries
Power BI	Interactive dashboards and KPIs
PDF Writer	Final case study summary export

Step-by-Step Process

Phase 1: Data Cleaning (Excel)

- Removed blank or null values from Gender, Married, Self_Employed, Dependents, Credit_History, and LoanAmount
- Converted data to proper formats (dates, income as numbers)
- Saved the cleaned data in **CSV UTF-8** format for import

Phase 2: Data Analysis (SQL)

Analyzed the data using MySQL:

1. **Overall approval rate**
2. **Approval rate by gender**
3. **Approval rate by education level**
4. **Approval by self-employment status**
5. **Impact of credit history on approval**
6. **Income level vs. loan amount ratio**
7. **Approval percentage for different property areas**

 **Total SQL Queries Used: 10+**

Phase 3: Dashboard Building (Power BI)






Key Visuals

1. Bar chart: Loan Status by Education
2. Stacked bar: Approval by Gender
3. Pie chart: Property Area distribution
4. KPI Cards:
 - Total Applications
 - Approval Rate
 - Avg. Loan Amount
 - Self-Employed Approval Rate

Filter

- Gender

Key Insights (Quantified)

Insight	Value
 Overall approval rate	~69%
 Approval rate for Graduates	~80%
 Approval rate for Self-Employed	<50%
 Strong Credit History → Approval rate	~85-90%
 Risk Reduction Potential	Estimated 30–40%

Conclusion

This project demonstrates the power of combining **Excel, SQL, and Power BI** to solve a real-world financial problem. By identifying risk indicators (like low income, poor credit history, or inconsistent employment), banks can make **smarter decisions**, reduce **default risks**, and improve the **approval process**.