**Loan Approval Prediction Analysis**

**Project Objective** The aim of this project is to analyze loan application data, visualize approval trends, and build predictive models to determine loan approval status effectively.

**Data Cleaning and Preprocessing**

1. **Handling Missing Values**:
   * Imputed categorical variables like Gender and Married with the mode.
   * Filled numerical columns like Loan Amount with the mean to address skewness.
2. **Feature Engineering**:
   * Created Total Income and its logarithmic transformation for better modeling.
   * Applied log transformation to Loan Amount for normalization.

**Exploratory Data Analysis**

* Explored relationships between loan approval (Loan Status) and factors like:
  + **Income**: Higher combined incomes correlated with increased approval rates.
  + **Education**: Graduate applicants had better chances of approval.
  + **Property Area**: Urban areas showed higher approval rates compared to rural ones.

**Model Training and Results**

1. Models Used:
   * **Logistic Regression**: Best performance with 80% accuracy.
   * **Support Vector Machine**: Achieved 68% accuracy.
   * **Decision Tree** and **Random Forest**: Lower accuracies (40% and 32%, respectively).
2. Evaluation:
   * Logistic Regression emerged as the most reliable model for predicting loan approval.

**Conclusion**

This analysis highlights key factors influencing loan approval and the effectiveness of different predictive models. Logistic Regression is recommended for its balance of simplicity and accuracy. Future efforts could focus on improving predictions through hyperparameter tuning and adding advanced features.