FINDINGS

1. **Data Loading and Initial Exploration**
   * The Data set had 614 Rows and 13 columns
   * **The columns that contain missing/null values are:**

Gender: 13 missing values

Married: 3 missing values

Dependents: 15 missing values

Self\_Employed: 32 missing values

LoanAmount: 22 missing values

Loan\_Amount\_Term: 14 missing values

Credit\_History: 50 missing values

**Data Cleaning Action**

* + Dropped columns with missing values less than 5%.
  + Filled remaining missing values with appropriate strategies (mean, median, or mode).

1. **Data Preprocessing**
   * Converted categorical variables(e Gender, Married,Self\_Employed, Education, and Property Area) into numerical values using label encoding.

1. **Model Performance:**
   * Random Forest Classifier provided the highest accuracy after hyperparameter tuning, making it the best model for predicting loan status.
   * The accuracy of the model improved from 77.76% before tuning to a higher score after tuning to 80.66

**Loan Approval Predictions:**

* + Predictions on new data were successfully generated, and the loan status for each application was determined.