a) Dimension or Size of the Dataset:

1. Number of Rows:
it contains 500 rows (based on the dataset description).

2. Number of Columns:

The dataset has 14 columns.

3. Missing Values:

- **Total Missing Values: 0**
- **Rows with Missing Values: 0**
- **Columns with Missing Values: 0**

4.Target Variable: loan_status

b) Methods Applied in the Code:

1. **Preprocessing Methods:**
    - **Column Transformer:**
      Applied to handle numerical and categorical features separately.
        - **Numerical Features:**
            - **Standardized using StandardScaler.**
        - **Categorical Features:**
            - **One-hot encoded using OneHotEncoder(handle_unknown='ignore').**
    - **Imputation: (although it isn't explicitly shown in the provided code, this is commonly included in ColumnTransformer pipelines).**
    - **Feature Selection: Dropped the target variable (loan_status) from the feature set.**
2. **Machine Learning Methods:**
    - **Random Forest Classifier:**
        - **Used as the first model to predict loan status.**
    - **Decision Tree Classifier:**

        - **Applied as a simpler tree-based model.**
    - **Logistic Regression:**
        - **Used for linear classification.**
    - **Gradient Boosting Classifier:**
        - **Applied for boosting-based predictions.**
    - **Model Evaluation:**
        - **Accuracy: Computed using accuracy_score.**
        - **Confusion Matrix: Visualized using sns.heatmap.**
        - **Classification Report: Includes**
          precision,
          recall,
          F1-score.

3. **Additional Details from the code:**

   a)dataset loading

   b)Exploratory Data Analysis(EDA)

   c)Feature Engineering
   #Feature cleanup
   #Target Feature Splitting

   d)Train-Test Spilt:
   # 80%training data
   # 20%testing data

   e)Evaluation Metrics:
   #Accuary score
   #confusion matrix
   #classification report

   f)Heat map visualation of confusion matrix

   g)Classifier pipelines

#column transformer
#RandomForestClassifier
#DecisionTreeClassifier