**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 11 July 2024 |
| Team ID | 740023 |
| Project Title | Smart Lender - Applicant Credibility Prediction For Loan Approval |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing Template**

Data exploration and preprocessing for loan approval prediction begins with thorough exploratory data analysis (EDA) to understand distributions and relationships among features like credit scores, income levels, and loan amounts. Addressing missing data is critical; imputation methods such as mean, median, or predictive modeling are applied. Numerical features are scaled to ensure uniformity in influence, while categorical variables are encoded (e.g., one-hot encoding) for compatibility with machine learning models. These steps ensure the dataset is ready for training robust models that predict loan approval outcomes accurately based on processed data.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Dimensions: 614 rows × 13 columns  Descriptive statistics: |
| Univariate Analysis |  |
| Bivariate Analysis |  |
| Multivariate Analysis |  |
| Outliers and Anomalies |  |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Handling Missing Data |  |
| Data Transformation |  |
| Feature Engineering | Attached the codes in final submission. |
| Save Processed Data | - |