

|  |  |  |
| --- | --- | --- |
|  | **Data Collection and Preprocessing Phase** | |
|  |  |  |
| Date |  | 15 March 2024 |
|  |  |  |
| Team ID |  | PNT2022TMID124356 |
|  |  |  |
| Project Title |  | SmartLender - Applicant Credibility |
|  |  | Prediction for Loan Approval |
|  |  |  |
| Maximum Marks |  | 6 Marks |
|  |  |  |

**Data Exploration and Preprocessing Report**

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

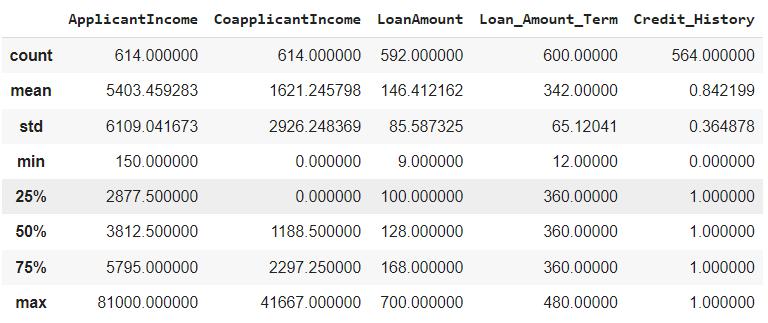
**Section**

**Description**

Dimension:

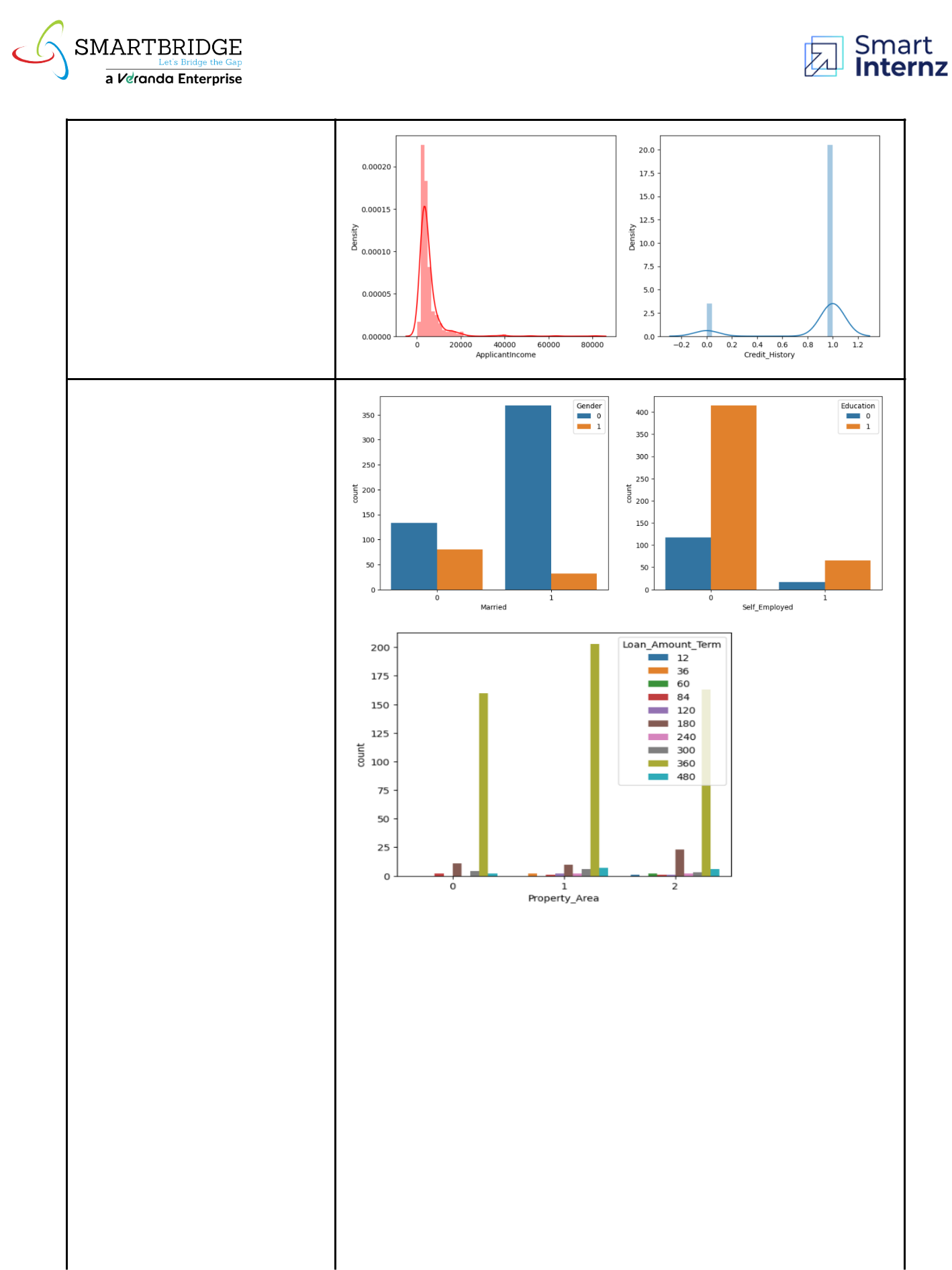
614 rows × 13 columns

Descriptive statistics:

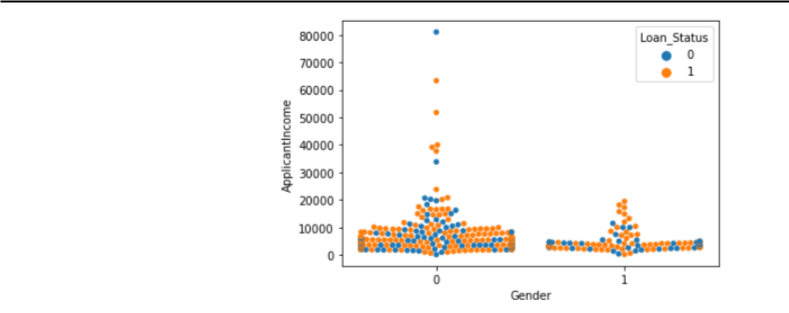


Data Overview

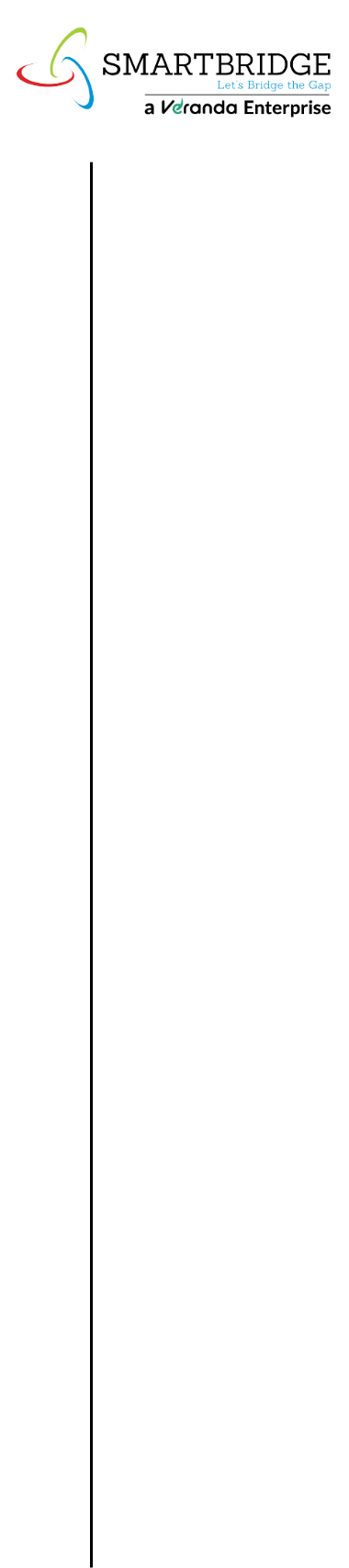
Univariate Analysis



Bivariate Analysis

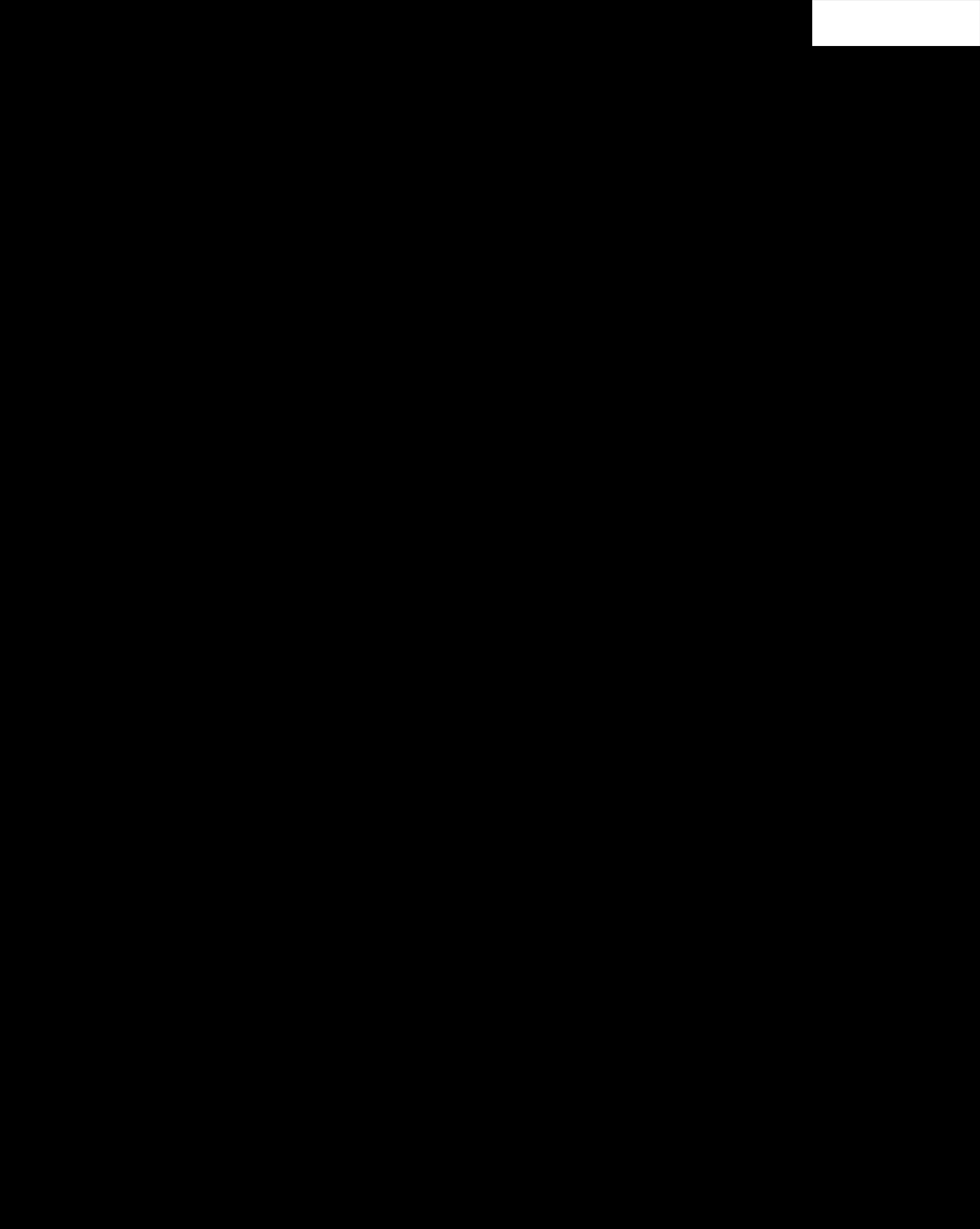


Multivariate Analysis

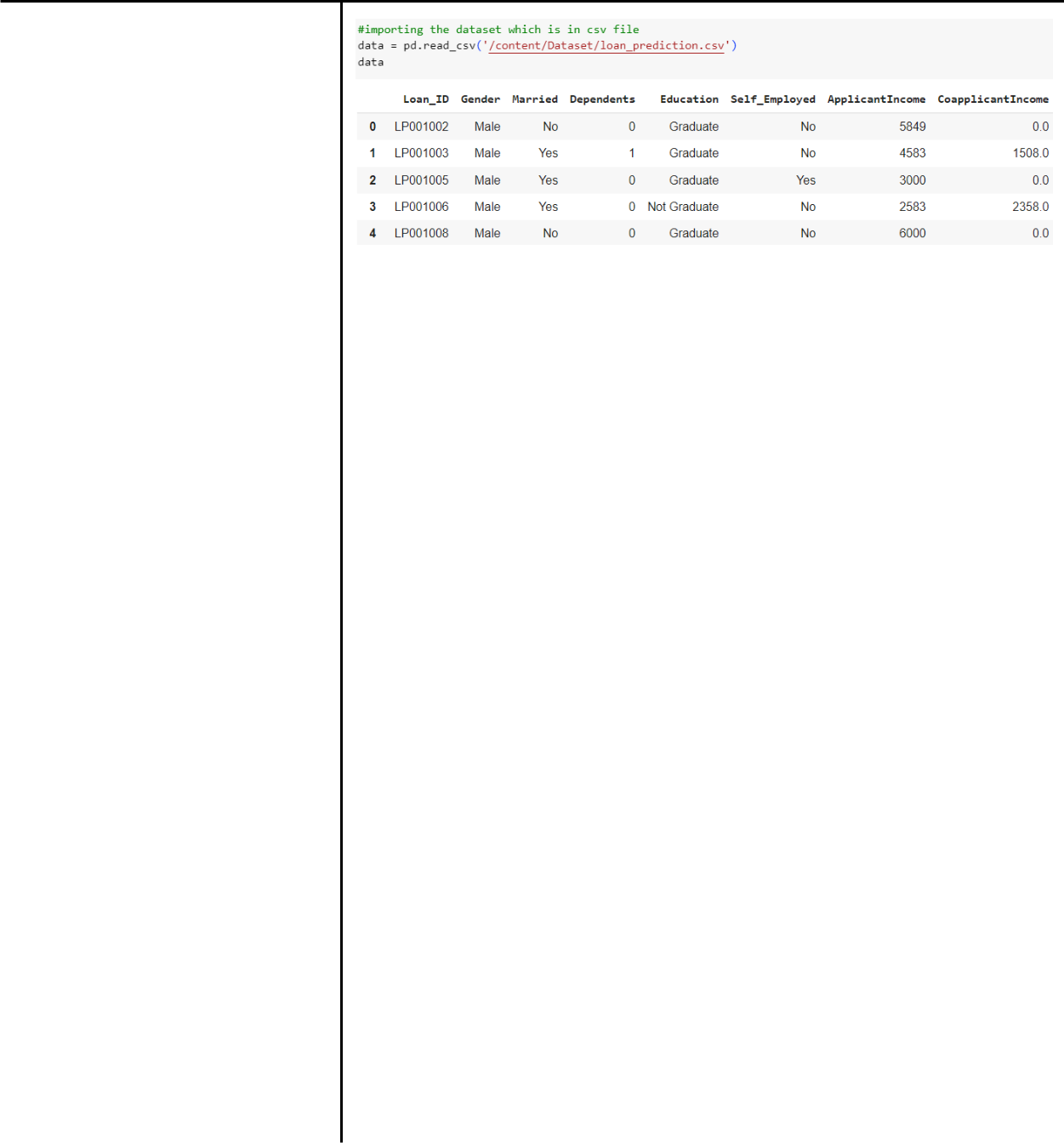


Outliers and Anomalies

-



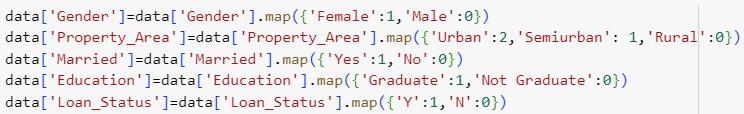
**Data Preprocessing Code Screenshots**



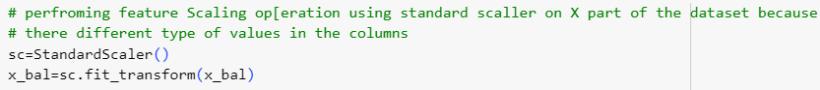
Loading Data



Handling Missing Data



Data Transformation



|  |  |
| --- | --- |
| Feature Engineering | Attached the codes in final submission. |
|  |  |
| Save Processed Data | - |