Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | 739674 |
| Project Title | Smart Lender- Flight delay Prediction |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing Report**

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employedforpreprocessingtaskslikenormalizationandfeatureengineering.Datacleaningwill address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Dimension:  11231rows×26columns Descriptive statistics:  C:\Users\DELL\Desktop\data.PNG |
| Univariate Analysis |  |

|  |  |
| --- | --- |
|  |  |
| Bivariate Analysis | C:\Users\DELL\Desktop\biagarph.PNG  **C:\Users\DELL\Desktop\arrivalgraph.PNG** |
| Heat map Analysis | C:\Users\DELL\Desktop\heatmap.PNG |

|  |  |
| --- | --- |
| Loading Data | C:\Users\DELL\Desktop\loadingdata.PNG |
| Handling Missing Data &  Replacing null Values | C:\Users\DELL\Desktop\hiding data.PNG |
| Handling outliers | C:\Users\DELL\Desktop\handlingoutlo.PNG |
| Handling Categorical values | C:\Users\DELL\Desktop\capet.PNG  C:\Users\DELL\Desktop\1.PNG |
| Splitting data into independent and dependent Variables | C:\Users\DELL\Desktop\depe.PNG |