# Data Collectionand Preprocessing Phase

| Date | 15 March 2024 |
|---|---|
| Team ID | 740095 |
| Project Title | Smart Lender-Flight Delay Prediction |
| Maximum Marks | 2 Marks |

## Data Quality Report:

A data quality report in flight delay prediction focuses on assessing the accuracy, completeness, and reliability of the data used for modeling. Key metrics include data consistency, checking for missing or duplicated entries, and verifying the timeliness of data updates. The report should also evaluate the source credibility, ensuring data is gathered from reliable and authoritative providers, such as aviation authorities and weather services. Anomalies and outliers should be identified and addressed to prevent skewed predictions. Additionally, the report must highlight any data integration issues, ensuring seamless merging of data from different sources. High-quality data is crucial for building an effective and trustworthy flight delay prediction model.

## Data Quality Report:

| Data Source | Data Quality Issue | Severity | Resolution Plan |
|---|---|---|---|
| | | | |

| Smart Interz | Missing values in the 'FL_NUM', 'MONTH', 'DAY_OF_MONTH', 'DAY_OF_WEEK', 'ORIGIN', 'DEST','DEP_DEL15,'CRS_ARR_TIME'And,'ARR_DEL15',' columns. | Moderate | Use mean/median imputation. |
|---|---|---|---|