

## Data Collection and Preprocessing Phase

Date	18 June 2024
Team ID	739634
Project Title	Flight Delays Prediction Using Machine Learning
Maximum Marks	2 Marks

### Data Quality Report:

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

### Data Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle ,UCI Dataset	Missing values in the dataset 'DEP_TIME', 'DEP_DELAY', 'DEP_DEL15', 'ARR_TIME', 'ARR_DELAY', 'ARR_DEL15', and 'ACTUAL_ELAPSED_TIME' columns.	Moderate	Use mean/median imputation.
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.