Data Collection and Preprocessing Phase

Date	24 April 2024
Team ID	team-739848
Project Title	Identifying Airline Passenger Satisfaction Using Machine Learning
8Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template 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 Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Some unwanted columns are present in the dataset	Low	Dropped unwanted columns using drop() function.
Kaggle Dataset	Null values are present in the dataset	Medium	Removed the null values from dataset by using the following code: data.dropna(inplace=True)
Kaggle Dataset	Outliers are present in some of the columns of dataset	High	Identified outliers by drawing boxplot and removed them from upper bound and lower bound.

Kaggle Dataset	Some categorical values consisting columns are present in the dataset	High	Converted the string form of data into float by applying Label Encoding on columns consisting of categorical values.