Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	739682
Project Title	SpaceX Falcon 9 First Stage Landing Success Predictor
Maximum 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
Dataset	Missing values in the 'LandingPad' column	Moderate	Drop the 'LandingPad' column as it contains 40% missing values.
Dataset	Inconsistent format in 'Date' column	Low	Standardize the 'Date' format to YYYY-MM-DD.
Dataset	Outliers in 'PayloadMass' column	Moderate	Remove outliers based on domain knowledge or replace with mean/median values.

Dataset	Categorical features not encoded	Low	Apply label encoding to categorical features like 'Orbit', 'LaunchSite', etc.
Dataset	Duplicate entries	Low	Remove duplicate rows from the dataset.