

Data Collection and Preprocessing Phase

Date	12 March 2024
Team ID	SWTID1720089323
Project Title	Ecommerce Shipping Prediction Using Machine Learning
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
Kaggle Dataset	Missing values in the(shipping mode, customer details, product specifics, and discount offers)	Moderate	Use mean/median imputation.
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.
Kaggle Dataset	Outliers in attributes.	High	Outliers are detected using quartiles and replaced with the standard value(median).