

Data Collection and Preprocessing Phase

Date	25 march 2025
Team ID	LTVIP2025TMID28445
Project Title	Cosmetic insights
Maximum Marks	3 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 ingredient lists	high	Use imputation techniques (e.g., replacing with "Unknown" or using ingredient databases to fill missing entries)
Cosmetic Product Dataset	Duplicate product entries	Moderate	Remove duplicates based on product ID and brand name