

DataCollectionandPreprocessingPhase

Date	20 June2025
TeamID	LTVIP2025TMID51415
ProjectTitle	CosmeticInsights:NavigatingCosmeticsTrendsandConsumerInsightswithTableau
MaximumMarks	3 Marks

DataQualityReportTemplate:

DataSource	DataQuality Issue	Severity	ResolutionPlan
Cosmetics datasets	InconsistentData Types	Moderate	<ul style="list-style-type: none"><li>• <b>Price:</b>Shouldideallybenumeric(e.g.,floatorinteger). While it appears to be numeric, it's worth confirming that there are no non-numeric characters present.</li><li>• <b>Rank:</b>Shouldbenumeric(floatordecimal)torepresent rating values. Verify that all values are within a reasonable range (e.g., 1-5 or 0-5).</li><li>• <b>Combination,Dry,Normal,Oily,Sensitive:</b>Theseare likely intended to be boolean (True/False or 1/0). The data appears consistent with 1/0, but it should be explicitly verified.</li></ul>
Cosmetics datasets	MissingValues andIncomplete Data	Low	<ul style="list-style-type: none"><li>• <b>Ingredients:</b>The"Ingredients"columnistruncatedin many entries (e.g., "Algae (Sea...", "Galactomy...", "Water,Dic...").Thismakesitdifficulttofully understand the product's composition.</li><li>• <b>Brand Names:</b> Somebrand names areabbreviated(e.g., "DRUNK EL" instead of "Drunk Elephant"). Consistency isneeded.</li></ul>
Cosmetics datasets	DataInconsistency andTypos	High	<ul style="list-style-type: none"><li>• <b>Inconsistent Capitalization:</b> Brand and product names showinconsistentcapitalization(e.g.,"LAMER"vs."la mer" or "Truth Seru" vs. "Truth Serum").</li></ul>

Cosmetics datasets	Data Validation	High	<ul style="list-style-type: none"><li>• <b>Verify Rank Values:</b> Ensure all ranks are within the expected range.</li><li>• <b>Confirm Data Types:</b> Explicitly define the data types for each column (numeric, boolean, string).</li></ul>
Cosmetics datasets	Data Cleaning	High	<ul style="list-style-type: none"><li>• <b>Standardize Brand and Product Names:</b> Ensure consistent capitalization and full names.</li><li>• <b>Correct Typos:</b> Fix errors like "Bt" and ensure consistent spelling.</li><li>• <b>Complete Ingredient Lists:</b> If possible, obtain the full ingredient lists for each product.</li><li>• <b>Remove Special Characters:</b> Replace or remove special characters as needed.</li></ul>

