

Dear Sprocket Central Pty Ltd,

Hope this email finds you well. I am writing to provide you with an important update regarding the data quality of the three datasets we've received from you. After a thorough access of the datasets, some key areas where require attention have been highlighted to enhance the data quality. Details of the key findings and the actions we recommend addressing each issue have been outlined below:

Dataset Transactions:

Quality Concerns	Field(s)	Action(s) Recommended
Data Completeness	online_order; brand; product_line; product_class; product_size; standard_cost; product_first_sold_date	Excluding missing value records from dataset or set default values for those missing values per business guideline
Inappropriate/ inconsistent Value Format	list_price; product_first_sold_date; standard_cost	Format list_price values to currency; Format product_first_sold_date values to date; Unify all values of standard_cost to currency

Dataset CustomerDemographic:

Quality Concerns	Field(s)	Action(s) Recommended
Data Completeness	last_name; DOB; job_title	Determine if missing values of these fields are crucial for making business decision and whether take action to exclude those records
Inconsistent Values	gender	Correct and replace inconsistent values with unified values; Look into the value of U for its meaning
Inappropriate/ Inconsistent Value Format	past_3_years_bike_related_purchased; tenure; DOB	Format their values to number; One of DOB values is in text format and need to be converted it to date
Data Accuracy	DOB	DOB contains birthdays that result with 123 years old and 128 years old

		should be excluded from the dataset
Data Relevancy	default; deceased_indicator	default filed values have no meaning and need to be deleted from the dataset; dataset should only include customers who are still alive, so records with Y as deceased_indicator need to be excluded from the dataset

Dataset CustomerAddress:

Quality Concerns	Field(s)	Action(s) Recommended
Inconsistent Value	state	Correct and replace inconsistent values with unified values;

If you have any questions or concerns regarding these findings and recommendations, please don't hesitate to reach out to me directly.

Looking forward to our continued collaboration.

Best Regard,
Data Analyst Intern
KPMG