

Subject: Data quality assessment results for Sprocket Central Pty Ltd

Dear Sir/Madam,

I hope this email finds you well. Thank you for reaching out to us regarding your data quality concerns. As promised, we have completed a comprehensive assessment of your data and would like to share our findings with you.

We have reviewed the datasets you provided and have identified data quality issues that have been discovered.

Summary table:

	Accuracy	Completeness	Consistency	Currency	Relevance	Validity	Uniqueness
Customer Demographic	<ul style="list-style-type: none">• Inaccurate DOB• Missing age	<ul style="list-style-type: none">• Blank job titles• Blank DOB	<ul style="list-style-type: none">• Inconsistent gender		<ul style="list-style-type: none">• Delete irrelevant data	<ul style="list-style-type: none">• allowable values in:<ul style="list-style-type: none">• Gender have (Male/Female)• Job industry• Wealth segment• Deceased_indicator• owns_car	<ul style="list-style-type: none">• No duplicated records
Customer Address			<ul style="list-style-type: none">• Inconsistent states			<ul style="list-style-type: none">• State• Country	<ul style="list-style-type: none">• No duplicated records
Transactions	<ul style="list-style-type: none">• Missing profit column	<ul style="list-style-type: none">• Blank online orders• Blank brands		<ul style="list-style-type: none">• Outdated transaction dates	<ul style="list-style-type: none">• delete cancelled orders	<ul style="list-style-type: none">• Online order• Order_status• Product_line• Product_class• product_size	<ul style="list-style-type: none">• No duplicated records
New Customer List		<ul style="list-style-type: none">• Blank job titles				<ul style="list-style-type: none">• format of list_price and product_sale_date	<ul style="list-style-type: none">• No duplicated records

Here is a detailed description of the data quality issues and strategies to mitigate these issues:

Accuracy issues:

- Inaccurate values **DOB** and a missing **Age** column in the “**CustomerDemographic**” sheet, along with missing **profit** column in the “**Transactions**” sheet
- Solution:
 - filtering out inaccurate **DOB** and add an **Age** column to extract age categories later for the “**CustomerDemographic**” sheet.
 - Add a **profit** column for the “**Transactions**” sheet to facilitate data analysis later.

Completeness issues:

- Blanks in **job_title**, **DOB** for “**CustomerDemographic**” sheet, **online_order** and **brand** in “**Transactions**” sheet, and **job_title** in “**NewCustomerList**” sheet.
- Solution:
 - We can delete the rows that contain the blanks if that won't affect our analysis later or ask the client for more details.

Consistency issues:

- The gender column contains inconsistency data such as the words “Female, Female”, and “F” for the female gender and “Male, Mal” for the male gender.
- Solution:
 - Replace all words that refer to female gender with one word “Female” and “Male” for all words that refer to the male gender.

Currency issues:

- The transactions dates are all outdated (from the year 2017) and currently we are on 2023.
- Solution:
 - Try to provide recent data if it is possible.

Relevancy issues:

- The presence of irrelevant data like default columns, canceled orders and deceased customers.
- Solution:

- Delete all the irrelevant data and filter out canceled orders and deceased customers.

Validity issues:

- Adding dropdown lists for columns data such as gender, wealth_segment, online_orders, product_class...., and changing format of list_price and product_sale_date

Uniqueness issues:

- There is no uniqueness issues as all ID values are unique.

Improving your data quality can have a significant impact on the success of your business, from enhancing your decision-making capabilities to increasing efficiency and productivity. We believe that by addressing these issues, you will be able to unlock new opportunities and achieve even greater success.

Thank you for entrusting us with this important assessment. If you have any questions or would like to discuss the results further, please do not hesitate to contact us.

Best regards,

Anouar