

Dear Sprocket Central Pty Ltd,

Thank you for providing us dataset from Sprocket Central Pty Ltd. We have reviewed the dataset and summarized the following data quality issues that we discovered within the dataset. The table below is a summary of a few data quality issues with Sprocket central Pty Ltd dataset.

Table: Summary Data Quality Issues

Worksheet Name	Data Quality Issue					
	Accuracy	Completeness	Consistency	Currency	Relevancy	Validity
Transaction		online order & brand: blanks				list price & product first sold date: format
New Customer List		job title: blanks				
Customer Demographic	DOB: Inaccurate	DOB & job title: blanks	gender: inconsistency	deceased customer: filter out	default: column delete	DOB: format
Customer Address			state: inconsistency			

We have taken relevant steps to identify these issues and given recommendations below to avoid further data quality issues in the future. Following recommendations will improve accuracy of data used to influence business decisions of Sprocket Central Pty Ltd in the future.

- 1. Worksheet name Transaction where we identified blank values for column “online\_order” and “brand”. The column for “list\_price” was converted into a currency format and “product\_first\_sold\_date” was converted onto a date/time format.**
  - a. We identified various blank values in the columns mentioned above, it is important to remove or filter out blank values from the dataset as the raise the data quality issue for completeness and can skew further analysis results.
  - b. The column for “list\_price” was converted into currency and “product\_first\_sold\_date” was converted into a date/time format which make it more easily to interpret. Changing the format according to the data type will be easier to read and analyze at the later stage.
- 2. Worksheet name new Customer List where we identified blank values for column “job\_title”.**

- a. As mentioned above, blank values were discovered in the sheet for the column "job\_title". Blanks are treated as incomplete data and can skew further analysis results. So, we filter out the blank values from the worksheet.

**3. Worksheet name Customer Demographic which has inaccurate and validity for "DOB", blank values for "job\_title", inconsistency for "gender", currency for "deceased\_customer" and irrelevant field called "default".**

- a. For "DOB" there is inaccurate data, where there is a customer's date of birth that is deviating so we need to filter it out and the data format was converted to date format. This column also contains blank values, these blank values correspond to the gender column which contains the 'U' value. When we filter out blank values, gender with 'U' value is also filtered out.
- b. Blank values were filtered out from "job\_title"
- c. The column for "gender" which is categorical variable has inconsistency, there are spelling errors for female, and some rows had abbreviations. This was replaced from M to Male and from F and Femal to Female. The column also consisted an irrelevant variable 'U' which was filtered out from the column. Enforce a drop-down list for the user entering the data rather than a free text field.
- d. For "deceased\_customer" we only use customer data that is still alive, so we filter out data Y or customers who have died. When data is current, it is more relevant to the analysis to be performed and less likely to lead to inaccurate conclusions.
- e. Irrelevant field called "default" contains obscure data, so it was removed as it had no relationship to the data.

**4. Worksheet name Customer Address which has inconsistency for state.**

- a. There are different state spellings but have the same meaning. Replace all 'New South Wales' to 'NSW' and 'Victoria' to 'VIC' for states.

That's all the summary of the analysis of data quality issues that found in the dataset provided by Sprocket Central Pty Ltd. The recommendations given are a simple and effective way to improve data quality for further analysis.

Please let us know if you have any questions regarding the data quality issues identified.

Kind Regards,

Sri Hartina Dj