Hello [customer name]

Thank you for providing the databases related to the sales of Sprocket Central Pty Lt. Below, we will elaborate on the problems encountered related to data quality and integrity and some recommendations to fix them.

Worksheet Name	Data quality issue
Transactions	Completeness & Relevancy
New Customer List	Completeness & Consistency
Customer Demographic	Completeness, Relevancy, Accuracy & Consistency
Customer Address	Relevancy

Issues and Recommendations

Below, we will detail different recommendations to improve the quality of the data and facilitate the subsequent analysis process.

- **Precision error**. In the column of the date of birth, there is a precision error, no human being born in 1843 should be alive.

Solution: filter out the dates that are unrealistic, as well as create demographic groups according to age to facilitate the possible analysis. Ex: (-18, 18-24, 24-30...)

- **Null values**. Vital columns for analysis such as brand purchased, job title, last name, or status of online transitions have rows with incomplete information.

Solution: Although cells with null information are a minority, it is important to eliminate them so that they do not lead to erroneous data analysis and improve data quality.

- **Inconsistent data type:** Consistency should be maintained in all the values of each attribute.

Solution: Change all 'F','Femal' to Female. All 'M' to 'Male' for 'gender' column. Change all 'New South Wales' to 'NSW' and 'Victoria' to 'VIC', in state column of Customer Demographics. Change 'Argiculture' to 'Agriculture' to correct the grammatical error

- **Numerous columns contain irrelevant information**. Information that is over-understood or that does not provide any added value or usefulness to future analysis.

Solution: Records with 'Y' in deceased_indicator in "Customer Demographics" or the records related to the country (Australia).

Both my team and I hope that this report will be of value. Once the changes have been made, we will move on to a much more detailed analysis of the situation, as well as an interactive visualization of the data through a dashboard.

Looking forward to hearing from you,

David de la Iglesia

Junior Data Consultant