Hello,

Thank you for providing us with datasets from Sprocket Central Pty Ltd. Here in the following is some quality issues that we discovered within the three datasets during our analysis. Please let us briefly describe the matters.

**Accuracy issues:**

* DOB was inaccurate for "Customer Demographic" and missing an age column and missing a profit column for "Transactions"

**Completeness:**

* Additional Customer IDs were inconsistent among "Customer Demographic," and "Customer Address" and "transactions". The data received may not be synced across all spreadsheets in the workbook.

• Blanks in job title for "Customer Demographic and for online orders and brand column for

"Transactions".

**Consistency:**

* Inconsistency in gender for "Customer Demographic" and "Customer Address respectively"

**Currency:**

* People that are "y" in deceased indicator are not current customers for "Customers demographic"

**Relevancy:**

* Lack of relevancy in default column for "Customer Demographic" and order status for

"Transactions"

**Validity:**

* Format of list price and product-sale date for "Transactions"

**Here are some Suggestions for the reduction of quality issues:**

* You can take a mode year value for the missing records of customers DOB.
* You can assign a uniform last name of customers; which values are missing.
* Replace gender 'U' with reference to the customer name and make a consistency.
* You can take a mean of rest of the values and assign the mean value to the missing fields in order to maintain the consistency of data.
* Eliminate the blank orders considering fake orders.

For your convenience, we all also provided the data after cleaning all the quality issues.

Best Regards

Ammar Jubril (Virtual Intern)

on Behalf of KPMG Data Analytics Team