**Dear Client**,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. The summary table below highlights key quality issues that we discovered within the three datasets. Please let us know if you have any queries issues surrounding the issues presented.

**Summary Table**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Accuracy** | **Completeness** | **Consistency** | **Currency** | **Relevancy** | **Validity** |
| Customer  Demographic | DOB:  Inaccurate  Age:  missing | Job Title:  Blanks  Customer id:  incomplete | Gender:  inconsistency | Decreased customers:  Filter out | Default column:  delete |  |
| Customer  Address |  | Customer id:  incomplete | States:  inconsistency |  |  |  |
| Transactions | Profit:  missing | Customer id:  Incomplete  Online Order:  Blanks  Brand:blanks |  |  | Cancelled  Status order:  Filter out | List price:  Format  Product sold date:  format |
|  | | | | | | |

Below are more in depth descriptions of data quality issues discovered and methods of mitigation used. Recommendations and explanations have also been included 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.

**Accuracy Issues**

**DOB was inaccurate for “Customer Demographic” and missing an Age\_column; missing a profit column for ”Transactions”.**

Mitigation: Filter out outlier in DOB.

Recommendation: Create an age column,create a profit column in “Transactions” to check accuracy of sales.

**Completeness**

**Additional customer\_ids were inconsistent among “customer demographic”, ”customer address”, and “transactions”**

Mitigation: Filter all customer\_ids from 1 to 3500.

Recommendation:Ensure tables are up to date(from the same time period).

Blanks in job\_title for “customer Demographic”,in online order and brand\_column for “transactions”

Mitigation:filter out the blanks.

Recommendation:Simply job title to another category such as industry or prodive dropdown options for job title.

**Consistency**

**Inconsistency in gender for ‘’customer demographic’’ and ‘’customer address”respectively**

Mitigation: Filter all M under category of Male and femal and F under category of Female gor gender.Filter all New Siuth wales to NSW and Victoria to VIC for states.

Recommendation:Create dropdown options for Male,Female,U in gender. Create dropdown options for all states abbrevations.

**Currency**

**People that are “Y” in decresed\_indicator are not current customers for “Customer demographic” and order\_status for “Transactions”**

Mitigation:Deleted metdata in default column.filter out cancelled order\_status.

Recommendation:check for incomprehensible metadata and delete or format to make comprehensible.

**Validity**

**Format of list\_price,product\_sale\_date for “transaction”.**

Mitigation: filterproduct sale date to short date and list price to currency.

Recommendation:setup columns so that formats such as price and decimals are already in place when entering new data.

**Relevancy**

**Lack of relevancy or comprehensibility in default \_column for “customer demographic “ and order status for ‘’transactions”**

Mitigation:Deleted Metadata in default column.filter out cancelled order status.

Recommendation:Checks for incomprehensible Metadata and delete or format to make comprehensible.

That summarises all data quality issues disovered through the first stages of the data quality analysis.The mitigation strategies suggested are simple and effective ways of improving data quality for future analysis.They will not only improve the analysis output that one can perform within the country but will increase the level of analysis that ca be performed by KPMG and other hired analysis teams,

Please let me know if you have questions regarding mitigation or any data quality issues identified.

Kind regards,

Kaviya.