Dear Kathleen,

Thank you for providing us with the 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 surrounding the issue presented.

Summary Table

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Accuracy | Completeness | Consistency | Currency | Relevancy | Validity | Uniqueness |
| Customer  demographic | DOB: inaccurate  Age:  missing | Job title: blanks  Customer Id :  incomplete | Gender: inconsistency | Deceased customer: Filter Out | Default column: delete |  | There are no duplicates |
| Customer address |  | Customer id: incomplete | States: Inconsistency |  |  |  | There are no duplicates |
| Transactions | Profit: missing | Customer id: incomplete  Online order: blanks  Brand: blanks |  |  | Canceled Status Order: filter Out | List price,  Product first sold date | There are no duplicates |

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. The following recommendations will improve the accuracy of data used to influence the business decisions of Sprocket Central Ltd in the future.

**Accuracy Issues**

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

**Completeness**

Additional customer\_ids were inconsistent among "Customer Demographic," "Customer Address," and "Transactions"

* Blanks in job title for "Customer Demographic," in online order and brand column for "Transactions"

**Consistency**

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

**Currency**

• People that are "Y" in the deceased indicator are not current customers for "Customer Demographic

**Relevancy**

Lack of relevancy or comprehensibility in default\_column for "Customer Demographic" and order status for "Transactions”

**Validity**

Format of list price, product sale date for "Transactions"

That summarises all data quality issues discovered through the first stage 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 company but will increase the level of analysis that can be performed by KPMG and other hired analysis teams

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

Kind regards,

Shreyash Jaiswal