Dear Manager,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd.

|  |  |
| --- | --- |
| **Sheet name** | **Data Quality Issue** |
| Transactions | Accuracy, Completeness, Relevancy, Validity |
| Customer Demographic | Accuracy, Completeness, Consistency, Currency, Relevancy |
| Customer Address | Completeness, Consistency |

Notable data quality issues that were encountered and the methods used to mitigate the identified data inconsistencies are as follows.

# **Transactions Data:**

* Null values removed
* Cancelled ordered filtered out
* List price was changed from Int format to Currency ($) format
* Product first sold date was changed from Int format to date format

# **Customer Demographic Data:**

* Null values removed
* DOB is inaccurate so some ages are missing
* Gender had spelling errors which were fixed. Some were also recorded as ‘U’
* Deceased customers are filtered out
* Default did not make sense so it was deleted

# **Customer Address Data:**

* State values are in both shortform and longform. Changed all to their shortforms

Warm Regards,

T.G.Tarun Kishore.