The Manager,

Sprocket Central Pty Ltd,

The following are the data quality issues as observed from the data set you gave us for analysis and also the various ways of mitigating the issues.

* **Accuracy:** The accuracy of your data set is a bit okay, but there are some data that are incorrect and are needed to be fixed by filling them with a correct data.
* **Completeness:** The data set given to us fulfils the expectation of comprehensiveness, although there are some missing values, so in order for us to have an optimum result, some of this data will be filled with an average value if they are numeric and they will be dropped if they are not numeric.
* **Consistency:** From the data set given to us, the customer demographic shows that we have 4000 customers, while the customers address shows that there are 4003 customers. This can be solved by removing 3 customers from the customer address data.
* **Timeliness:** Sine the data set represent the transaction that was done for the past three years, the data set is timely and can be used for the analysis.
* **Validity:** The “First sold date” in the transaction data set and “Default” in the customer’s address data set are not in a usable format. To use the data set, there is need to delete or ignore this information since they are not in a usable format.
* **Uniqueness:** This data set is a unique data set since there is no replication of information