Dear Manager,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. Regarding the review of the dataset, here are the summary statistics and recommendations to help improve data quality that would help in driving business decisions. Please let us know if the figures are not aligned with your understanding.

1. In the DOB column, the date 1843-12-21 shows the person is 175 years old which is an error.

**Recommendation:** during data collection, there should be an error message that pops up if the date of birth is not within a certain time period.

1. Certain columns such as the genders are not consistent. For example Female is represented as both “Female and F”

**Recommendation:** a drop down list should be used in data collection rather than a free text field to maintain consistency.

1. There is a default column in the customer demographic table that does not contain allowable values.

**Recommendation:** Ensure data entered contain allowable values else an error message should pop up.

1. The customer address table and transactions table have additional customer id’s that are not in the customer demographic table which is the customer master table that would be used as a training set for our model.

**Recommendation:** Ensure all tables are from the same time period as only the customer demographic table would be used as our master table

1. Various columns in the customer demographic, transactions table and new customer list table have empty values.

**Recommendation:** During data collection, ensure that all fields are required to be filled and not skipped.

Moving forward, the team will continue with the data cleaning, standardisation and transformation process for the purpose of model analysis. Questions and assumptions documented. That would be all for now.

Kind regards,

Folashade