**Dear Client,**

**REPORT ON QUALITY DATA ASSESSMENT**

Our most profound gratitude to our client sprocket central pty Ltd. For trusting us with the task to analyze their data for the purpose of effective management and better decision making by providing the following data on their respective customers, they are as follows;

1. Customer demographic.

2. Customer Address.

3. Transaction Data.

Some errors were identified in the provided data and the methods used to rectify and alleviate the identified data inconsistency from occurring in the future are listed below. Also recommendations are provided to prevent further occurrence of errors in data collections and enhance data quality.

* **Presence of blank/empty values in certain records such as** *online orders, brands etc***.**

*Alleviation: if only few of the rows in the data set are empty, ensure to separate the record entirely from the training set, except if it is a very essential field that cannot be neglected i.e. absence of a customer’s last name from a data set.*

*Recommendation*: Ensure that very essential details that are needed must be completely entered before a users registration can be completed/approved in the course of data collection.

* **The use of inconsistent values for the same attributes** *i.e. customers address: address, customers demographic: gender etc.*

*Alleviation: using a general term that is concise and well defined value in place of the other values used to represent an attribute for consistency.*

*Recommendation*: The use of dropdown list of a specified option for users while entering their data helps in reducing inconsistent data value representation.

* **Inconsistent data types for the same attribute** *i.e. for currency and DOB in customers transaction and demographics respectively.*

*Alleviation: for rows having inconsistent data types, they are filtered out i.e. non-numeric data types are removed from numeric data set*

*Recommendation:* Users should be restricted to using only specific data types in some given field i.e. other characters should not be allowed in those fields.

Furthermore, the team will continue with the data cleaning, standardization and transformation for the purpose of model analysis. Some questions will be raised and recommendations will be made along the way. After this is thoroughly done, a perfect model will be built to suit the need of sprocket pty Ltd. Using the refined data through directives given to us by our client.

Kind regards,

[Junior consultant, Emmanuel Favour]