**KPMG – Sprocket Central Pty Ltd**

Dear [Client point-of-contact],

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. The below table highlights the summary statistics from the three datasets received. Please let us know if the figures are not aligned with your understanding.

Notable data quality issues that were encountered and the methods used to mitigate the identified data inconsistencies are as follows. Furthermore, recommendations have been provided to avoid the re-occurrence of data quality issues and improve the accuracy of the underlying data used to drive business decisions.

*Data Quality Issues in the datasets:*

1. Transactions Worksheet:

* It can be observed that there are 360 orders whose online\_order values are blank.
* There are 197 orders whose brand, product\_line, product\_class, product\_size standard cost and product first sold date is blank/null.
* 1378 products have the same id that is 0 which I guess is not possible.
* Certain product have their first sold date before 2000 which might be wrong.

1. New Customer List Worksheet:

* 17 Customers data does not have their Date of Birth and their Gender also is defined by the letter U which does not signify anything.

1. Customer Demographics Worksheet:

* DOB contains one row with value 1843-12-21 which is not practically possible.
* 87 customers do not have their DOB assigned or filled as well as their Gender is also undefined.
* Cust\_ids: {1, 54, 57} have improper gender value.
* There are 87 customers who don’t have a tenure mentioned.

Recommendations:

* For key datasets, such as transactions, Less than 1% of transactions (totaling less than 0.1% of revenue) have missing fields. These records have been removed from the training dataset.
* Data Type for certain columns have been transformed as that will make data interpretation process a bit difficult.

Moving forward, the team will continue with the data cleaning, standardization and transformation process for the purpose of model analysis. Questions will be raised along the way and assumptions documented. After we have completed this, it would be great to spend some time with your data SME to ensure that all assumptions are aligned with Sprocket Central’s understanding.

Kind regards,

[Junior Consultant Name]