

To Dear Client,

Thank you so much for providing the dataset of Sprocket Central Pty Ltd. including dataset of "Customer Demographic", Customer Address" and Transaction data in past 3 months".

I was going through all these datasets and while analyzing it, I came across some data inconsistency which will hamper our analysis and can be easily eliminated in future collection process.

Below are few anomalies which I have figured out during my data quality review process:-

1 Transition Dataset:

- There are several columns with null values such as brand, product line, product class, product size, standard cost and product first sold date which are unique and very important to understand in terms on business. For a complete clean dataset we have to first fill these null values by average or remove those null values.

2. Customer Demographic Dataset:

- This dataset also contains multiple columns with null values. Also, the data is inconsistent such as in columns gender with "Male" and "Female" there are multiple other vague attributes such as "M", "Femal" and "U". We have to correct these vague terms by replacing it with Male or Female.
- There are customers with DOB from 1843 to 1944 where there is just single customer entry and is redundant for any kind of analysis

3. Customer Address Dataset:

- In this dataset the State column also contains data inconsistency where apart from attributes like New South Wales and Victoria there are also other vague attributes such as "QLD", "VIC" which we have to take care of by either eliminating or replacing with actual correct terms.

4. Customer ID inconsistency:

- Also, In addition to these 3 dataset while comparing Customer Address to Customer Demographic dataset there are 4 id's from Customer Address which are absent in Customer Demographic and 5 id's are absent in Transaction dataset.

These are some minor issues which are hampering data accuracy and quality of the dataset and it is important to fix them to ease out the analysis process and maintained data integrity.

Hope you find some valuable insights in this email.

Thank You so much.

Best Regards,

Shashank Sahu

6/16/2020