Dear XYZ,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. The recommendations have been provided to avoid the reoccurrence of data quality issues and improve the accuracy of the underlying data used to drive business decisions.

● In Transaction dataset, there are many duplicate values present in product\_id and customer\_id column.

● In Customer demographic dataset, there is some age issue with some customer. As the age of some customers are showing more than 120 years. Some names in customer\_name column is also not present.

● Job titles are also missing at many places in Customer demographic dataset

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,

Shivam Kumar