**Task – 1**

**(KPMG Virtual Internship Task -1 : Mail the Client regarding the data analysis on the data they provided)**

Hello [client point-of-contact]

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd.

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 reoccurrence

of data quality issues and improve the accuracy of the underlying data used to drive business

decisions.

* **Additional customer\_ids in the ‘Transactions table’ and ‘Customer Address table’ but not in ‘Customer Master (Customer Demographic)’**
* **Various columns, such as the brand of a purchase, or job title, have empty values in certain records**
* **Inconsistent values for the same attribute (e.g. Victoria being represented as “V”, “Vic” and “Victoria”)**
* **Inconsistent data type for the same attribute (e.g. numeric values for some fields and strings for others)**

Moving forward, the team will continue with the data cleaning, standardisation 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,

Bideepto Bhattacharjee