Dear Kathleen

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.

TABLE NAME	NO. OF RECORDS	DISTINCT RECORDS	DUE DATE RECEIVED
Transactions	20000	20000	28 March 2023
NewCustomerList	1001	1001	28 March 2023
CustomerDemographics	4001	4001	28 March 2023
CustomerAddress	3999	3999	28 March 2023

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 for the underlying data used to drive business decisions.

- Additional customer_ids in "Transactions" and CustomerAddress" table, but not in "CustomerDemographics".
- Various columns, such as the brand of a purchase, or job title, have empty values in certain records.
  1. In Transactions table, a total of 1542 cells are blank, which means 0.60% of data is missing.
  2. In NewCustomerList table, a total of 217 cells are blank, meaning 0.87% of data is missing.
  3. In CustomerDemographics table, a total of 1045 cells are blank, meaning that 2.01% of data is missing.
  These records have been removed from the dataset to ensure data quality.
- Inconsistent values or the same attribute. (ex.: "Victoria" being represented as "V", "Vic")
- Inconsistent data type for same attribute (ex.: numeric values for some values, string for some others)

We are moving forward for further processes, to help you in better and more efficient ways, and its great to spend time analyzing your data.

Thanks & Regards

Soumyadeep Paul
(KPMG Team)