Dear Sprocket Central Pty Ltd team,

Thank you for sharing the 3 data sheet (Transaction, Customer Address, Customer Demographic). The

Data quality assessment is shared below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Total Number of columns | Total Number of records | Unique customer\_id |
| Transaction | 13 | 20000 | 3494 |
| Customer Address | 6 | 3999 | 3999 |
| Customer Demographic | 13 | 4000 | 4000 |

* Null values – There are multiple null values in transaction and customer demographic sheet. Reason of null values should be checked weather the customers are not willing to provide or the error is from the companies end.
* Data type not specified – There are multiple columns whose datatype is not specified or incorrect. Make sure while creating a database each column datatype should be specified.
* Uninformative column – There is one column that contains no information just some random scientific notations. It might be due to wrong fetching of data or the database is not created properly, it should be checked.
* Location error – Names of locations are not mentioned properly, if not maintained properly one location will be divided into multiple locations.

This are the first observation we have found the above issues. We will further analyse the data, perform data cleaning, and transform the data for further use. Questions will be raised along the way and assumptions documented.

Yours sincerely,

Piyush Mishra