Sprocket Central Pty Ltd

28th Nov, 2023

Hi there,

I have gone through the datasets that you have shared with us. There are several problems with the dataset regarding the data quality. During processing on datasets, we get the following over view.

**No of records** **Unique Id’s** **Receiving date**

|  |  |  |  |
| --- | --- | --- | --- |
| Transaction dataset | 20002 | 3495 | 28/11/2023 |
| Customers Demographic | 4002 | 4002 | 28/11/2023 |
| Customers address | 4001 | 4001 | 28/11/2023 |

Issues related with datasets:

1- Additional customer’s id’s in Transaction datasets, may be there is a problem of data loss.

2- Various columns such as brand purchase have empty values.

3- Inconsistent data type with same attributes.

4-Inconsistent data values with same attribute.

We are moving forward for further process like data cleaning, standardization etc and it will be great to spend some time with your data.

Regards,

Ganesh Dalal

KPMG