Dear Client (Sprocket Central Pty Ltd),

I'm excited about the opportunity to collaborate with you on your data base and by processing the information we received.

The result is as follows

|  |  |  |  |
| --- | --- | --- | --- |
|  | No of records | Unique Id’s | Receiving date |
| Transaction dataset | 20001 | 3495 | 24/06/2023 |
| Customers Demographic | 4001 | 4001 | 24/06/2023 |
| Customer Address | 4000 | 4000 | 24/06/2023 |

Issues found in the datasets:

1.Additional in **Transactions,** customer\_id column in the datasets as it occur in the in data loss.

2.Various columns like brand, order online, order online, product size have empty values (blanks).

3. many data types with the same properties.

4. In **CustomerDemographic,** DOB wrong data where added.

5. In **CustomerDemographic,** the gender section needed to be replaced as (Female being represented as “F”, “FEMAL”, ”FEMALE”) as well as for Male represented as “M” or “Male” are for the same gender, and in **CustomerAddress**, the state data like NSW represented as and “New south Wales”, “NWS”, and Victoria being represented as “Vic” and “Victoria.

6. AS moving forward the further process of cleaning, standardization and will be great to spend time evaluating the data.

7. In **CustomerDemographic**, The default column data are unreadable or unidentified.

Nirmal Raj