To,

Sprocket Central Pty Ltd

Subject - Data Quality issues

Respected Sir/Madam,

I hope this email finds you well. As per your request, I have thoroughly reviewed the datasets provided by Sprocket Central Pty Ltd and have identified several data quality issues that require attention. Below is a comprehensive list of the issues found in each dataset:

|  |  |  |
| --- | --- | --- |
| **Sheet Name** | **No. of Rows (Before Cleaning)** | **No. of Rows (After Cleaning)** |
| Customer Demographic | 4000 | 2863 |
| Customer Address | 4000 | 4000 |
| Transactions | 20000 | 19946 |
| New Customer List | 1000 | 736 |

**Customer Demographic Dataset: -**

In this Dataset, the Default column has many garbage values that don’t mean anything to the dataset.

The columns like last\_name, job\_title, job\_industry\_category, and DOB have so many null values.

Also, In the Gender column have some unknown value ‘U’. Change the value of “F” to “Female” and “M” to “Male”.

**Customer Address: -**

In this Dataset, state columns have some values like "New South Wales" and "Victoria" Change them to “NSW” and “VIC” respectively.

**Transactions: -**

Missing Values in Columns like brand, Product\_line, Product\_class, Product\_size, list\_price,

Standatd\_cost and Product\_first\_sold\_date contain missing values.

In the Product\_first\_sold\_date column Data type is Number so change it to Date Data type, and check that data type.

**New Customer List: -**

Columns like DOB, job\_title, and job\_industry\_category contain null values in this dataset. Need to change the tenure, postcode, property valuation, and past\_3\_year\_bike\_related\_purchase Datatype

Kind Regards,

Klinton Raval