To,

Sprocket Central Pty Ltd

Subject - Data Quality issues and strategies to mitigate these issues.

Respected Sir/Madam,

Below are the identified data quality issues and strategies to mitigate these issues.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sheet Name** | **Unique column** | **No. of Rows (Before cleaning)** | **No. of Rows (After cleaning)** |
| Transactions | Transaction\_id | 20001 | 3495 |
| New Customer List | NA (*since it is the target sheet*) | 1001 | 941 |
| Customer Demographic | customer\_id | 4001 | 4001 |
| Customer Address | customer\_id | 4000 | 4000 |

1. **Completeness (Fields with missing values):**

Transactions sheet → *online order, brand, product line, product class, product size, standard cost, product first sold date* contain null values.

New Customer List → *DOB, job title, job industry* contains null values.

Customer Demographic → *DOB, job title, job industry* contains null values.

**Mitigate Steps -**

The Data Handling team can try to tally the old data or relevant records to fill some null values.

1. **Relevancy (Irrelevant values)**

New Customer List → *gender* ⇒ Deleted value = "U" because don’t know whether it is male or female

Customer Demographics → *gender* ⇒ Deleted value = "U" because don’t know whether it is male or female, deceased indicator ⇒ Deleted value = "Y" because want only alive customers, *default* ⇒ Deleted column since it is not decodable

**Mitigate Steps -**

The Data Handling team should tally the record of "U" in gender to find out the correct values.

1. **Consistency (Up to date values)**

Customer Demographics → *gender* ⇒ Changed value = "F" to "Female" and "M" to "Male"

Customer Address → *state* ⇒ Changed value = "New South Wales" to "NSW", "Victoria" to "VIC"

**Mitigate Steps -**

The Data Handling team should check the latest values used for fields so as to keep up to date database

1. **Accuracy (Values not in the correct format)**

Transactions → *product first sold date* ⇒ Changed data type from number to date

**Mitigate Steps -**

The Data Handling team should Check the data type of columns to mitigate this type of issue.

Thank you for your trust in our expertise, and we look forward to a productive partnership with Sprocket Central Pty Ltd.

Best regards,

Kosisochukwu .L. Mbachu-Igwe.