Dear Client,

This is Shreyansh Vaishnav from KGMP Data Analytics Team. While going through the Data Quality assessment of the Dataset provided by your company I came through several issues and missing values within the dataset.

While going through the dataset the following insights are drawn and I have also came up with several techniques to improve the data quality and the performance of your business and it will eventually help you to bring out better analytics and results for your company.

**1. Transactions Section (Total records: 20000 past 3 Months) errors encountered:**

|  |  |
| --- | --- |
| **Missing Values/Field Names** | **Count of Missing Values** |
| Online Orders | 360 |
| Brand | 197 |
| Product Line | 197 |
| Product Class | 197 |
| Product Size | 197 |
| Standard Cost | 197 |
| Product First Sold Date | 197 |

**Improvement Required:**

1. Replacing missing Product first Sold Dates with mode of Product first Sold Date.

2. Removing blank transactions as fake orders.

3. We can consider the missing Standard Costs as the mean of Standard Cost.

**2. New Customer List Section errors encountered:**

|  |  |
| --- | --- |
| **Missing Fields** | **Count of Missing Values** |
| Last Name | 29 |
| Date Of Birth | 17 |
| Job Tittle | 106 |
| Job Industry Category | 165 |

**Improvement Required:**

1. Replacing missing Last Names with a fixed/Uniform Last Name.

2. Replace missing DOB with a specific Date of Birth.

3. Replacing gender ‘U’ as ‘Unidentified’.

**3. Customer Demographic Section errors encountered:**

|  |  |
| --- | --- |
| **Missing Fields** | **Count of Missing Values** |
| Last Name | 125 |
| DOB | 87 |
| Job Tittle | 506 |
| Job Industry Category | 656 |
| Tenure | 87 |

**Improvements Required:**

1. For tenure values to keep the consistency of data, take the mean of available values and replace the missing values with that mean.

2. DOB should be replaced with a fixed DOB value.

3. Gender values with ‘F’ and ‘Femal’, should be merged with ‘Female’ and value ‘M’ should be merged with ‘Male’.

Regards,

KPMG (Data Analytics Team)