Hello,

Thank you for providing the datasets from Sprocket Central Limited.

We have reviewed the data sets in detail. The below table summarises the key data quality issues discovered through the first data quality analysis stage. Along with it, mitigations/ recommendation are also included to optimize the data collection quality which would eventually improve the quality of analysis & insights.

Kindly go through each of the points & please do let us know in case you have any queries. I’d be happy to discuss to ensure that our assumptions and understanding are aligned.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dataset** | **Transactions** | **New Customer List** | **Customer Demographics** | **Customer Address** | **Mitigation / Recommendations** |
| **Accuracy** |  |  | Error in DOB |  | 1. Include an age column to verify the validity of the DOB entered OR restrict format and plausible timeline to input DOB |
| **Completeness** | Missing: customer ID, online order, product line, product class, product size, standard cost, first sold date | Missing: last name, DOB, job title, job industry | Missing: last name, job title, job industry, tenure |  | 1. Ensure data collected is complete in all aspects.  2. Provide predefined options for job title, job industry, online order, and tenure. |
| **Consistency** | Currency Format: List price, Standard Cost |  | Format: gender | Format: State Name | 1. Standardize format for currency 2. Use Drop down to input Gender, State Name |
| **Timeliness** |  |  | Deceased Customer Data |  | 1. Deceased customers are not current customers, update information whenever possible |
| **Relevancy** |  | Irrelevant: unnamed columns | Irrelevant: default column deleted |  | 1. Remove any incomprehensible or irrelevant data collection. |
| **Uniqueness** |  |  |  | State Name |  |
| **Validity** | Product\_first\_sold\_date incorrect | Gender: U for Unspecified |  |  | 1. Standardize format for gender 2. Product sold date received as integer, input format to be refined. |

Regards,

KPMG (Data Analytics Team)