Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 08/19/2023

Internship Batch: LISUM24

Version:<1.0>

Data intake by: Gideon Osei Bonsu

Data intake reviewer: Data Glacier

Data storage location:

**Tabular data details:**

**File Name:** Cab\_Data

|  |  |
| --- | --- |
| **Total number of observations** | 359392 |
| **Total number of files** |  |
| **Total number of features** | 7 |
| **Base format of the file** | .csc |
| **Size of the data** | 20.1 MB |

**File Name:** City

|  |  |
| --- | --- |
| **Total number of observations** | 20 |
| **Total number of files** |  |
| **Total number of features** | 3 |
| **Base format of the file** | .csc |
| **Size of the data** | 759 B |

**File Name:** Customer\_ID

|  |  |
| --- | --- |
| **Total number of observations** | 49171 |
| **Total number of files** |  |
| **Total number of features** | 4 |
| **Base format of the file** | .csc |
| **Size of the data** | 1.00 MB |

**File Name:** Transaction\_ID

|  |  |
| --- | --- |
| **Total number of observations** | 440098 |
| **Total number of files** |  |
| **Total number of features** | 3 |
| **Base format of the file** | .csc |
| **Size of the data** | 8.58 MB |

* I used the following steps to analyze the data:
* Ask: At this stage, I asked pressing questions regarding the purpose of the analysis and what the customer is hoping to get from the analysis.
* Prepare: this is where I imported the required libraries and read datasets.
* Process: This is where I cleaned the data, checked for null values, blank cells, cleaned column names, and grouped datasets to create a master dataset.
* Analyze: At this stage, I analyzed the data to draw insights from it.
* Share: This stage ran through “Analyze” stage as I shared visualizations of various sections of analysis of the data.
* Act: Here, I made recommendations.