Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: July 14, 2024

Internship Batch: LISUM35

Version: 1.0

Data intake by: Chenxin Shen

Data intake reviewer: None

Data storage location: https://github.com/DataGlacier/DataSets

**Tabular data details:**

Cab\_Data.csv

|  |  |
| --- | --- |
| **Total number of observations** | 359392 |
| **Total number of files** | 4 |
| **Total number of features** | 7 |
| **Base format of the file** | .csv |
| **Size of the data** | 20.2 MB |

City.csv

|  |  |
| --- | --- |
| **Total number of observations** | 20 |
| **Total number of files** | 4 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 759 Bytes |

Customer\_ID.csv

|  |  |
| --- | --- |
| **Total number of observations** | 49171 |
| **Total number of files** | 4 |
| **Total number of features** | 4 |
| **Base format of the file** | .csv |
| **Size of the data** | 1 MB |

Transaction\_ID.csv

|  |  |
| --- | --- |
| **Total number of observations** | 440098 |
| **Total number of files** | 4 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 8.58 MB |

**Proposed Approach:**

There are no duplicates in the data