Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 13th April 2022

Internship Batch: LISUM08

Version:

Data intake by: Sheereen Kallivalappil

Data intake reviewer:

Data storage location: https://github.com/SheereenHamza/Data-Glacier/tree/main/Week%202

**Tabular data details: Cab\_Data**

|  |  |
| --- | --- |
| **Total number of observations** | 359392 |
| **Total number of files** | 1 |
| **Total number of features** | 7 |
| **Base format of the file** | .csv |
| **Size of the data** | 20.1 MB |

**Tabular data details: Customer\_ID**

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

**Tabular data details: Transaction\_ID**

|  |  |
| --- | --- |
| **Total number of observations** | 440098 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 8.7 MB |

**Tabular data details: City**

|  |  |
| --- | --- |
| **Total number of observations** | 20 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 1 KB |

**Proposed Approach:**

* No null values in any of the data sets
* No duplicate entries in any of the data sets