Data Intake Report

**Name: G2M insight for Cab Investment firm**

**Report date: 14 – 09 – 2022**

**Internship Batch: LISUM13:30**

**Version: 1.0**

**Data intake by: Kavinilavan Muthukumar**

**Data intake reviewer: Data Glacier**

**Data storage location: https://github.com/kavinilavanM/Data-Glacier-**

**internship/tree/main/Week-2/Dataset**

**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: 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** | 8.58 MB |

**Tabular data details: Customer\_Data**

|  |  |
| --- | --- |
| **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: 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** | 4 KB |

**Proposed Approach:**

* No Null values found in all four datasets.
* Data format is changed to yy/mm/dd.