Data Intake Report

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

**Report date:** June 30, 2022

**Internship Batch:** LISUM11

**Version:**1.0

**Data intake by:** Yanjun Lin

**Data intake reviewer:** Data Glacier

**Data storage location:** https://github.com/

**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** | 21.2 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** | 9 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** | 1.1 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:**

• There is no missing values in all four datasets that mentioned above.