Data Intake Report

Name: G2M insight for Cab investment firm

Report date: 14-05-2023

Internship Batch: LIMSUM21

Version:<1.0>

Data intake by: Akhil Abraham

Data intake reviewer:

Data storage location: https://github.com/Akhilabraham9/G2M-Cab-insghts

**Tabular data details: Cab\_Data**

|  |  |
| --- | --- |
| **Total number of observations** | 59392 |
| **Total number of files** | 1 |
| **Total number of features** | 7 |
| **Base format of the file** | .csv |
| **Size of the data** | 20663kb |

**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** | 1kb |

**Tabular data details: Customer\_ID**

|  |  |
| --- | --- |
| **Total number of observations** | 491711 |
| **Total number of files** | 1 |
| **Total number of features** | 4 |
| **Base format of the file** | .csv |
| **Size of the data** | 1027kb |

**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** | 8788kb |

**Tabular data details: US\_holiday**

|  |  |
| --- | --- |
| **Total number of observations** | 342 |
| **Total number of files** | 1 |
| **Total number of features** | 5 |
| **Base format of the file** | .csv |
| **Size of the data** | 16kb |

**2Proposed Approach:**

* Convert ‘Date of Travel’ to date time format in Cab Data.
* Change all float values to max 2 decimal places.
* Replace unwanted commas from Population and Users in City Data.
* Download Holiday data for future analysis.
* Create a profit column for future analysis.
* Merge all data to one dataset.