Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 13th August, 2022

Internship Batch: LISUM12: 30 July - 30 October 2022

Version: 1.0

Data intake by: Joseph Nnodim

Data intake reviewer:

Data storage location:

**DATA INTAKE REPORT**

**Tabular data details:**

**Transaction\_ID data:** The file containing this data is in csv format

|  |  |
| --- | --- |
| **Total number of observations** | 440,098 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 8,788kb |

**Cab\_Data**

|  |  |
| --- | --- |
| **Total number of observations** | 359,392 |
| **Total number of files** | 1 |
| **Total number of features** | 7 |
| **Base format of the file** | .csv |
| **Size of the data** | 20,663 |

**Customer\_ID data:** The file containing this data is in csv format

|  |  |
| --- | --- |
| **Total number of observations** | 49,171 |
| **Total number of files** | 1 |
| **Total number of features** | 4 |
| **Base format of the file** | .csv |
| **Size of the data** | 1,027kb |

**City data:** The file containing this data is in csv format

|  |  |
| --- | --- |
| **Total number of observations** | 19 |
| **Total number of files** | 1 |
| **Total number of features** | 3 |
| **Base format of the file** | .csv |
| **Size of the data** | 1kb |

**Proposed Approach:**

* Mention approach of dedup validation (identification)
* Mention your assumptions (if you assume any other thing for data quality analysis)

**Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.**

**Please do not forget to remove this section while converting the file into pdf.**