

Data Intake Report

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

Report date: **13- Mar-2023**

Internship Batch: **LISUM19**

Version: **1.0**

Data intake by: **Yash Jadwani**

Data intake reviewer: **Yash Jadwani**

Data storage location: **GitHub**(<https://github.com/yashjadwani/Week2/tree/master/Datasets>)

Tabular data details:

File Name	Cab_Data.csv
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	21.8 MB

File Name	City.csv
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

File Name	Customer_ID.csv
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	1027 KB

File Name	Transaction_ID.csv
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

File Name	Holiday.csv
Total number of observations	57
Total number of files	1
Total number of features	6
Base format of the file	CSV
Size of the data	3 KB

Data Validation

- There were no duplicate or null records found while analyzing the data.
- There are some trips on which customer has been charged so much so I have made an assumption that 'The data is accurate and free from errors or inconsistencies'