

Data Intake Report

Name: G2M insight for Cab investment firm

Report date: July 14, 2022

Internship Batch:LISUM11

Version:1.0

Data intake by:Shovan Roy

Data intake reviewer:

Data storage location: <https://github.com/shovan20/DataSets>

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_ID

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	1.05 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	759 B

Proposed Approach:

- Forking the dataset repository into your local system.
- Reading the Dataset into the jupyter file
- Exploring the dataset.
- The dataset was clean and there were no missing or duplicate values.