

# Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 13/2/2023

Internship Batch: LISUM18

Version:<1.0>

Data intake by: Abdelrahman Atef

Data intake reviewer: na

Data storage location: [hemany1/G2M-insight-for-Cab-Investment-firm \(github.com\)](https://github.com/hemany1/G2M-insight-for-Cab-Investment-firm)

## Tabular data details: Cab\_Data

Total number of observations	359392
Total number of files	1
Total number of features	8
Base format of the file	.csv
Size of the data	19.2+ 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	10.1+ 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	608.0+ bytes

## 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	1.5+ MB

## Proposed Approach:

- Apply Exploratory Data Analysis
- Check for duplicates, missing values, and outlier

