

Data Intake Report

Project Title: G2M insight for Cab Investment Firm

Report Date: 09-02-2022

Internship Batch: LISUM06

Intern Trainee Name: Vidya Ganesan

Project Reviewer: Data Glacier

Project URL: <https://github.com/vidya-ganesan/G2M-insight-for-cab-investment-firm>

Details of given Datasets:

Cab Data

Given file format	.csv
Total Number of data's	359392
Given File	1
Number of features	8
Size of the data	19.2 MB

Customer_ID

Given file format	.csv
Total Number of data's	49171
Given File	1
Number of features	4
Size of the data	1.5MB

Transaction_ID

Given file format	.csv
Total Number of data's	440098
Given File	1
Number of features	3
Size of the data	10.1MB

City

Given file format	.csv
Total Number of data's	20
Given File	1
Number of features	3
Size of the data	608 bytes

There are no null values in the 4 datasets

Merged all the 4 datasets as a master dataset.

Studied the statistical info for all the data sets.

Performed various EDA for all the datasets .