

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 08/14/2022

Internship Batch:LISUM12

Version:<1.0>

Data intake by: Sanjukta Choudhury

Data intake reviewer:

Data storage location: https://github.com/SanjuktaCh/DG_G2M-CaseStudy

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	21.5MB

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 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	1MB

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	8.58MB

Proposed Approach:

- There are no null values in the 4 dataset
- Price charged has outliers while the rest of the attributes has no outliers