

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 17/7/2022

Internship Batch: LISUM11: 30

Version:

Data intake by: Manar alwsaidi

Data intake reviewer: <intern who reviewed the report>

Data storage location: <location URL eg: github, cloud>

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	.xlsx
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	.xlsx
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	.xlsx
Size of the data	608.0+ bytes

Tabular data details:

Total number of observations	49171
Total number of files	1
Total number of features	4
Base format of the file	.xlsx

Size of the data	1.5+ MB
------------------	---------

Note: Replicate same table with file name if you have more than one file.

Proposed Approach:

- Mention approach of dedup validation (identification)
- Mention your assumptions (if you assume any other thing for data quality analysis)

**Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.
Please do not forget to remove this section while converting the file into pdf.**