

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 16-03-2022

Internship Batch: LISUM07

Version:<1.0>

Data intake by: Maaz Jafrani

Data intake reviewer: Maaz Jafrani

Data storage location: <https://github.com/maazjafrani/Data-Glacier-Internship-Work>

Tabular data details:

Total number of observations	359392
Total number of files	4
Total number of features	19
Base format of the file	.csv
Size of the data	46.1 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)

Proposed Approach:

- No duplicate rows in the global data
- No missing data in the data
- Column Date Travelled is in excel serial number but transformed it into real date