

Data Intake Report

Name: <G2M insight for Cab Investment firm>

Report date: <18-3-2022>

Internship Batch: <LISUM07>

Version:<1.0>

Data intake by: <Tarek Mohamed Abdullah Amin>

Data intake reviewer:<intern who reviewed the report>

Data storage location: <<https://github.com/DataGlacier/DataSets>>

Tabular data details:

Total number of observations	<359392>
Total number of files	<4>
Total number of features	<14>
Base format of the file	<csv>
Size of the data	<31.1 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)