

Data Intake Report

Name: <Cab Investment (week2)>

Report date: <July 13, 2023>

Internship Batch:<LISUM23:30>

Version:<1.0>

Data intake by:<Dingyun Hu>

Data intake reviewer:<Dingyun Hu>

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	<30.84 M >

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

Proposed Approach:

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