

Data Intake Report

Name: Cab investment firm

Report date: 20/06/2021

Internship Batch: LISUM01

Version: 1.0

Data intake by: ALIREZA EHIAEI

Data intake reviewer –

Data storage location: https://github.com/Alireza-Ehiaei/Data_Sciences/tree/main/Cab_industry/code%26data

Tabular data details for four following data files:

Cab_data

Total number of observations	359392
Total number of features	7
Base format of the file	csv
Size of the data	20 MB

City

Total number of observations	19
Total number of features	3
Base format of the file	csv
Size of the data	1 KB

Customer_ID

Total number of observations	49171
Total number of features	4
Base format of the file	csv
Size of the data	1 KB

Transaction_ID

Total number of observations	440098
Total number of features	3
Base format of the file	csv
Size of the data	9 MB

Proposed Approach:

- **Modifying and Cleaning data**
- **Adding new variables using existing data to have some better insights to data**
- **Proposing some hypothesizes and validating them**
- **Mapping data based on initial analysis and the results of hypothesis**