

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 14/03/2024

Internship Batch: LISUM31

Version:<1.0>

Data intake by: Jeffrey Joseph Joseph Madhava Rao

Data intake reviewer:<intern who reviewed the report>

Data storage location:

Tabular data details:

Cab_Data.csv

Total number of observations	359392
Total number of files	1/4
Total number of features	7
Base format of the file	.csv
Size of the data	21.2 MB

City.csv

Total number of observations	20
Total number of files	2/4
Total number of features	3
Base format of the file	.csv
Size of the data	759 bytes

Customer_ID.csv

Total number of observations	49171
Total number of files	3/4
Total number of features	4
Base format of the file	.csv
Size of the data	10,51,215 bytes

Transaction_ID.csv

Total number of observations	440098
Total number of files	4/4
Total number of features	3
Base format of the file	.csv
Size of the data	89,98,194 bytes

Proposed Approach:

- We have been provided with four data sets (Cab_Data, Customer_ID, Transaction_ID, and City)
- We have performed EDA to understand each data set and understand the relations between them.
- We have merged them into a master data set and have performed few analysis for invesment options between the two Cabs in the data set.

Areas Investigated are

- Payment modes of each company
- Gender wise income
- Cab usage in each state based on cabs
- Profit analysis