

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 14-03-23

Internship Batch:LISUM19

Version:

Data intake by: Anuj Singh

Data intake reviewer:

Data storage location: <location URL eg: github, cloud>

Tabular data details:

Total number of observations	359392
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	29.7 MB

Tabular data details: City_ID.csv

Total number of observations	20
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	759 bytes

Tabular data details: Customer_ID.csv

Total number of observations	49171
Total number of files	1
Total number of features	4
Base format of the file	.csv

Size of the data	1.00 MB
-------------------------	----------------

Tabular data details: Transaction_ID.csv

Total number of observations	440098
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	8.58MB

Proposed Approach:

- There were no null or duplicate values in the four datasets. Analyzed each dataset as well as merged and Mastered dataset using different analyzing and visualizing techniques in python. The notebook is well documented.