Data Intake Report

Name: G2M Insights
Report date: 05/14/2023
Internship Batch: LISUM21
Version: 1.0
Data intake by: Shiva Ramezani
Data intake reviewer:
Data storage location: https://github.com/ShivaRamezani/G2M_cap

Cab 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	20.2 MB

City data details:

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

Customer data details:

Total number of observations	49172
Total number of files	1
Total number of features	4
Base format of the file	csv
Size of the data	1 MB

Transaction data details:

Total number of observations	440099
Total number of files	1
Total number of features	3
Base format of the file	csv
Size of the data	8.58 MB

Proposed Approach:

- I use the data in each of the tables to derive results. After that, I tried to merge every two tables that have a feature in common to get valuable investigations. After all, I merged all 4 tables together and could get very interesting result out of it.
- For further investigation and implementing models, I would do scaling, normalization, dimensional reduction, and splitting.