Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 2022-07-14
Data intake by: Seoyoung Cho
Data intake reviewer:
Data storage location:

Tabular data details:

Cab_Data.csv

Total number of observations	359492
Total number of features	7
Base format of the file	.csv
Size of the data	62.8 MB

City.csv

Total number of observations	20
Total number of features	3
Base format of the file	.csv
Size of the data	4.2 KB

Customer_ID.csv

Total number of observations	49171
Total number of features	4
Base format of the file	.csv
Size of the data	4.2 MB

Transaction_ID.csv

Total number of observations	440098
Total number of features	3
Base format of the file	.csv
Size of the data	33.9 MB

oversea.csv

Total number of observations	20
Total number of features	2
Base format of the file	.csv
Size of the data	1.7 KB

Proposed Approach:

Oversea.csv file is extracted from the information provided by
https://www.bankrate.com/finance/credit-cards/merchants-guide-to-credit-card-processing-fees/

Assumptions:
1. Assumed net yearly profit is determined by the difference of charged price and cost of trip.
2. Assumed the payment card type is Visa and the average processing fee is 2.4%
3. Neglected all NA values
4. Categorized income group by having top 66th percentile (2/3rd) being higher income group. Between 33rd percentile and 66th percentile being middle class group and below 33rd percentile being the lower income group.