Data Intake Report

Name: G2M Cab Data
Report date: 09/14/2022
Internship Batch: LISUM 13:30
Version: 1.0
Data intake by: XYZ
Data intake reviewer: Alyssa Hubiak
Data storage location: GitHub - https://github.com/DataGlacier/DataSets

Tabular data details:

Cab_Data.csv

Total number of observations	359,392
Total number of files	4 total
Total number of features	7
Base format of the file	.csv
Size of the data	20,663 KB
City.csv	
Total number of observations	20
Total number of files	4 total
Total number of features	3
Base format of the file	.csv
Size of the data	1 KB
Customer_ID.csv	
Total number of observations	49,171
Total number of files	4 total
Total number of features	4
Base format of the file	.csv
Size of the data	1,027KB
Transaction_ID.csv	
Total number of observations	440,098
Total number of files	4 total
Total number of features	3
Base format of the file	.csv
Size of the data	8,788KB

Note: Replicate same table with file name if you have more than one file.

Proposed Approach:
- Can see that there are multiple unique identifiers to join this data on, including Transaction ID and Customer ID. Can combine on City for City data.
- All are CSV formatted to read into python
- No assumptions to be made at this time

Note: Convert this doc in pdf and provide the link of pdf file in your dashboard. Please do not forget to remove this section while converting the file into pdf.