Data Intake Report

Name: <G2M insight for Cab Investment firm>
Report date: <03/07/2023>
Internship Batch:<June 30 – Sept 30>
Version:<1.0>
Data intake by:<Yunxin Gan>
Data intake reviewer:<intern who reviewed the report>
Data storage location: < https://github.com/YunxinG107112/VC>

Tabular data details:

Cab data (cab.csv)

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	<20.1 MB>

City data (city.csv)

Total number of observations	<19>
Total number of files	<1>
Total number of features	<3>
Base format of the file	<.csv>
Size of the data	<759 bytes>

Customer_id data (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	<1.0 MB>

Transaction_id data (transaction_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	<1.0 MB>

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

Proposed Approach:
- Explanatory data analysis