Data Intake Report

Name: Explorator Data Analysis of Competing Cab Companies
Report date: March 20th 2023
Internship Batch: LISUM19
Version:1.0
Data intake by: Virginia Mullins
Data intake reviewer: Virginia Mullins
Data storage location: https://github.com/vlmullin/DG_WK2_DataSets

Tabular data details:

Cab_Data.csv

Total number of observations	359393
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	20,663 KB

Customer_ID.csv

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	40631 KB

City.csv

Total number of observations	21
Total number of files	1
Total number of features	3
Base format of the file	.csv
Size of the data	1 KB

Transaction_ID.csv

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	8788 KB

Proposed Approach:
- Secondary Storage Dedup