Data Intake Report

Name: G2M Cab DataSets
Report date: 06/12/2022
Internship Batch: LISUM 10: 30 May – 30 Aug 2022
Version: 1.0
Data intake by: Jose Luis Castañeda Terrones
Data intake reviewer:
Data storage location: https://github.com/DataGlacier/DataSets

Tabular data details:

Total number of observations	359 392
Total number of files	
Total number of features	7
Base format of the file	csv
Size of the data	20.663 KB

Total number of observations	49 171
Total number of files	
Total number of features	4
Base format of the file	csv
Size of the data	1.027 KB

Total number of observations	440 098
Total number of files	
Total number of features	3
Base format of the file	csv
Size of the data	8.788 KB

Total number of observations	20
Total number of files	
Total number of features	3
Base format of the file	csv
Size of the data	1KB

Proposed Approach:
- Cab_Data and Transaction_ID could be merged using the Transaction_ID column.
- I assumed that the integers in the Cab_Data file for the Date of Travel column are excel times and should be transformed to a suited format for analysis.