Data Intake Report

Name: G2M insight for Cab Investment firm
Report date: 13/08/2024
Internship Batch: LISUM36
Version:1.0
Data intake by: Stephanie Dawsonn-Andoh
Data intake reviewer:
Data storage location: https://github.com/stephandoh/Sondoh_Glacier/tree/main/Week_2

Tabular data details:

Total number of observations	1048575
Total number of files	1
Total number of features	14
Base format of the file	csv
Size of the data	115.87 MB

Proposed Approach	Assumptions
Ensured data were in the range of the timelines provided in the problem overview	
Matched weather data to corresponding dates of travel in transaction data.	Assumed that weather patterns were accurately reflected on corresponding dates.
Categorized weather conditions into broader categories (e.g., rain, snow) for analysis.	Assumed that weather categorization did not significantly oversimplify the data.
Corrected any discrepancies in date formats and ensured uniformity across datasets.	Assumed that the date format in weather data was consistent with other datasets.
Additional weather was obtained from here: https://www.kaggle.com/datasets/prasad22/weather-data	
Performed feature engineering	