

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	440098
Total number of files	4
Total number of features	3
Base format of the file	csv
Size of the data	8.58 MB

Proposed Approach

Proposed Approach	Assumptions
Linked transaction data with customer and cab company information.	Assumed that transaction records were complete and accurately recorded.
Merged transaction data with weather and holiday data for deeper analysis.	Assumed that transaction dates matched correctly with weather and holiday data.
Removed duplicates and corrected inconsistencies in transaction IDs.	Assumed that duplicate transactions were minimal and identifiable.