

Data Intake Report

Name: Cab industry

Report date: 14.08.2024

Internship Batch:LISUM36

Version:<1.0>

Data intake by: Fatmatüzzehra Öztürk

Data intake reviewer:

Data storage location: https://github.com/zehraozturkk/Cab_Industry

Tabular data details:

Total number of observations	359392
Total number of files	1
Total number of features	2
Base format of the file	csv
Size of the data	55,5 MB

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

Proposed Approach:

- Mention approach of dedup validation (identification)
- Mention your assumptions (if you assume any other thing for data quality analysis)

Data Quality and Processing Update

- **Duplicate Records:** No duplicate records were found in the dataset.
- **Datetime Conversion:** The `Date` column has been successfully converted to `datetime` format.
- **Additional Columns:** New columns have been added to the dataset to enrich the data.
- **Holiday Integration:** Holiday data has been integrated into the dataset, providing additional context for analysis.