Data Intake Report

Name: Retail Forecasting

Report date: August 18, 2024

Internship Batch: LISUM35

Version:1.0

Data intake by: Chenxin Shen

Data intake reviewer: Nuzat Siddiqua

Data storage location:<https://github.com/nuzats19/Retail-Forecasting-for-a-Large-Beverage-Company>

**Tabular data details:**

| **Total number of observations** | 1218 |
| --- | --- |
| **Total number of files** | 1 |
| **Total number of features** | 12 |
| **Base format of the file** | .xlsx |
| **Size of the data** | 73.8 KB |
|  |  |

**Proposed Approach:**

* We will remove duplicates and keep only one of them if we find any identical rows.
* Perform exploratory data analysis (EDA) to identify patterns, trends, seasonality, and outliers.
* We will remove if any outliers present in the dataset.
* We will group the data by products (which is the first column) and perform the same analysis on each group.
* Build at least 4-5 multivariate forecasting models to test different methodologies.
* Use time series cross-validation methods to evaluate models.
* Optimize models to get better accuracy.