Data Intake Report

Name: Retail Forecasting Group Project for an Australian Beverage Business
Report date: 10/19/2024
Internship Batch: LISUM37
Version:1.0
Data intake by: Data Forecasting Team
Data intake reviewer: Charles Ibitamuno
Data storage location: https://github.com/Impact026/Data-Forecasting-project/upload/main

Tabular data details:

Customer_ID

Total number of observations	1218
Total number of files	1
Total number of features	12
Base format of the file	csv
Size of the data	51 KB

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

Proposed Approach:
- Converting the xlsx file to csv file
- Using pandas to convert the csv file into a data frame.
- Checking the features and number of rows on the file
- Determining whether there are any unique rows in the data frame

After doing a preliminary analysis of the data, I found out that there are no duplicates in the file. Also, after checking for missing data, there wasn't any missing data in the given file. There are six unique rows under the "Product" column, "SKU1", "SKU2", "SKU3", "SKU4", "SKU5", and "SKU6", all sold from February 2017 to December 2020. I haven't made a EDA check but will in future.