**Market Basket Analysis**

**Phase 3**

**Problem Statement:**

The problem is to perform market basket analysis on a provided dataset to unveil hidden patterns and associations between products. The goal is to understand customer purchasing behavior and identify potential cross-selling opportunities for a retail business. This project involves using association analysis techniques, such as Apriori algorithm, to find frequently co-occurring products and generate insights for business optimization.

**Market basket analysis with Apriori algorithm**

The retailer wants to target customers with suggestions on itemset that a customer is most likely to purchase .I was given dataset contains data of a retailer; the transaction data provides data around all the transactions that have happened over a period of time. Retailer will use result to grove in his industry and provide for customer suggestions on itemset, we be able increase customer engagement and improve customer experience and identify customer behavior. I will solve this problem with use Association Rules type of unsupervised learning technique that checks for the dependency of one data item on another data item.

Association Rule is most used when you are planning to build association in different objects in a set. It works when you are planning to find frequent patterns in a transaction database. It can tell you what items do customers frequently buy together and it allows retailer to identify relationships between the items.

**Data Source:**

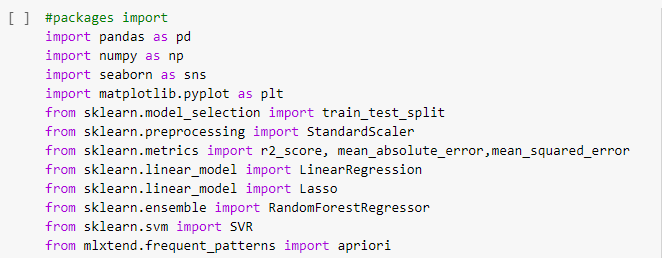
* **Google Collab:** Google Colab, a cloud-based Jupyter notebook environment, serves as our primary coding platform.
* **Python and other libraries** for association analysis and machine learning.

**Implementation Steps:**

1. **Import all necessary libraries:**

Importing all the necessary python libraries like numpy, pandas, seaborn, matplotlib, sklearn and mlxtend for performing association analysis techniques and machine learning.

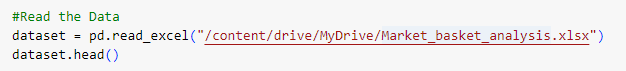
**Code:**

****

1. **Load and Explore the dataset:**

Load the excel file “Market\_basket\_analysis.xlsx” using pandas and explore the dataset to understand its structure.

**Code:**

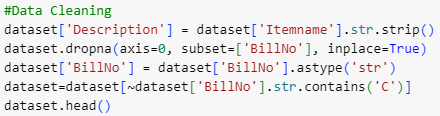


1. **Data Preparation:**

For preparing the data, Data cleaning is done. The following data cleaning steps are performed:

* **Removing duplicate data:** Removes any rows in the dataset using the dropna() function.
* **Correcting data errors:** Converts the column to a string type, which is necessary for performing string operations on the data.
* **Removing invalid data**

**Code:**

****

1. **Training Model:**

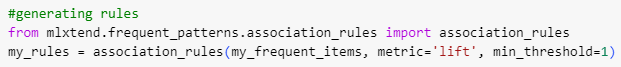
To train an association rule mining model, first identify the frequent itemsets in your data. This can be done using the Apriori algorithm. Once identified the frequent itemsets, generate association rules from them using the association\_rules() function.

Steps involved in training an association rule mining model:

* Identify frequent itemsets.
* Generate association rules.

**Code:**

****

****

1. **Marketing Recommendation:**

Now the condition based rules fileration using the the previously generated rules my\_rules is done.

**Code:**

****

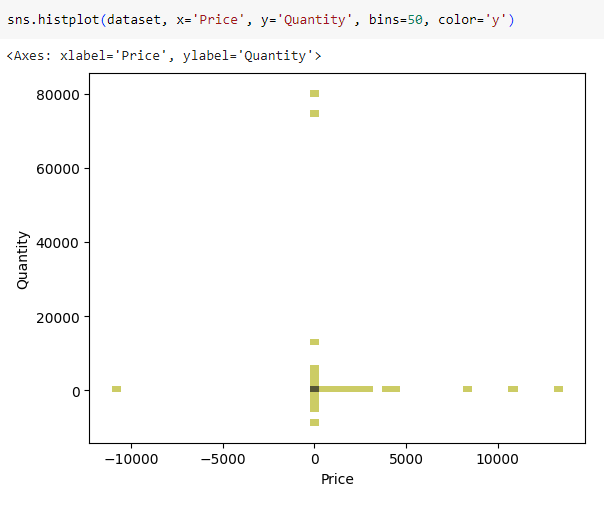
****

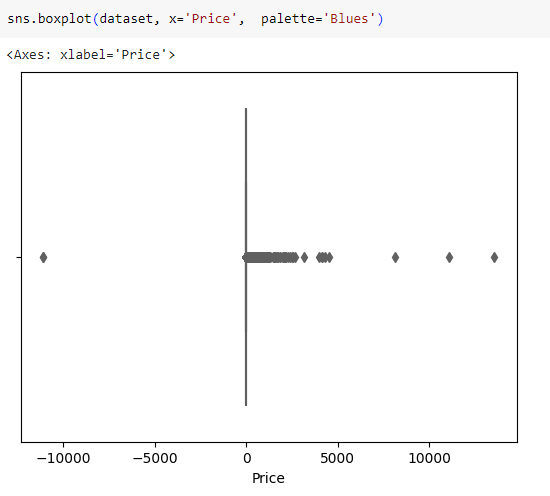
****

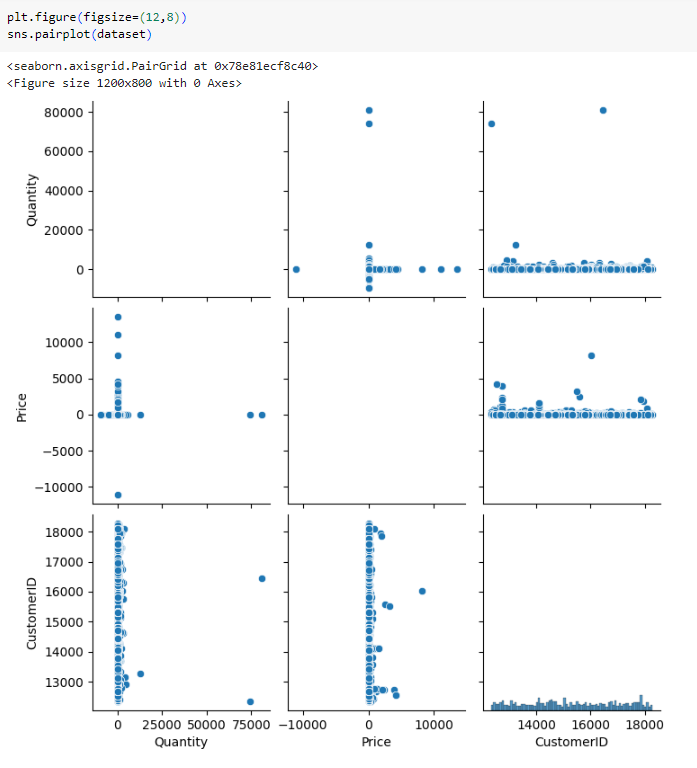
1. **Data Visualization:**

For visulizing the data, various plots such as 2D Histogram, boxplot, and pairplot using seaborn python library.

**Code:**

****

****

****

****

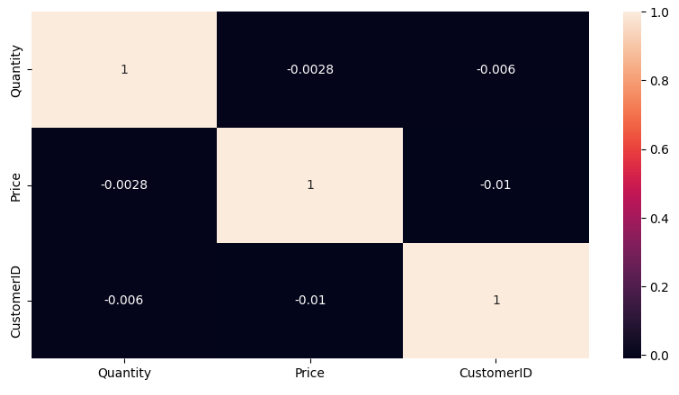
1. **Visualization Correlation:**

In this step, generating a heatmap in seaborn to visualize the correlation between variables in the dataset.

**Code:**

****

**Output:**

****