IBM Naan Mudhalavan Project

Group-1 Artificial Intelligence

Team Member -1

311521106070

Parthiban M

Project name

Market Basket Analysis

Phase-1

Document Submission

Market Basket Analysis



Definition:

**Market basket analysis is a powerful technique used to understand customer behavior by examining the relationships between items purchased in a transaction. It helps businesses gain insights into product associations, customer preferences, and can be used for various purposes such as cross-selling, inventory management, and personalized recommendations.**

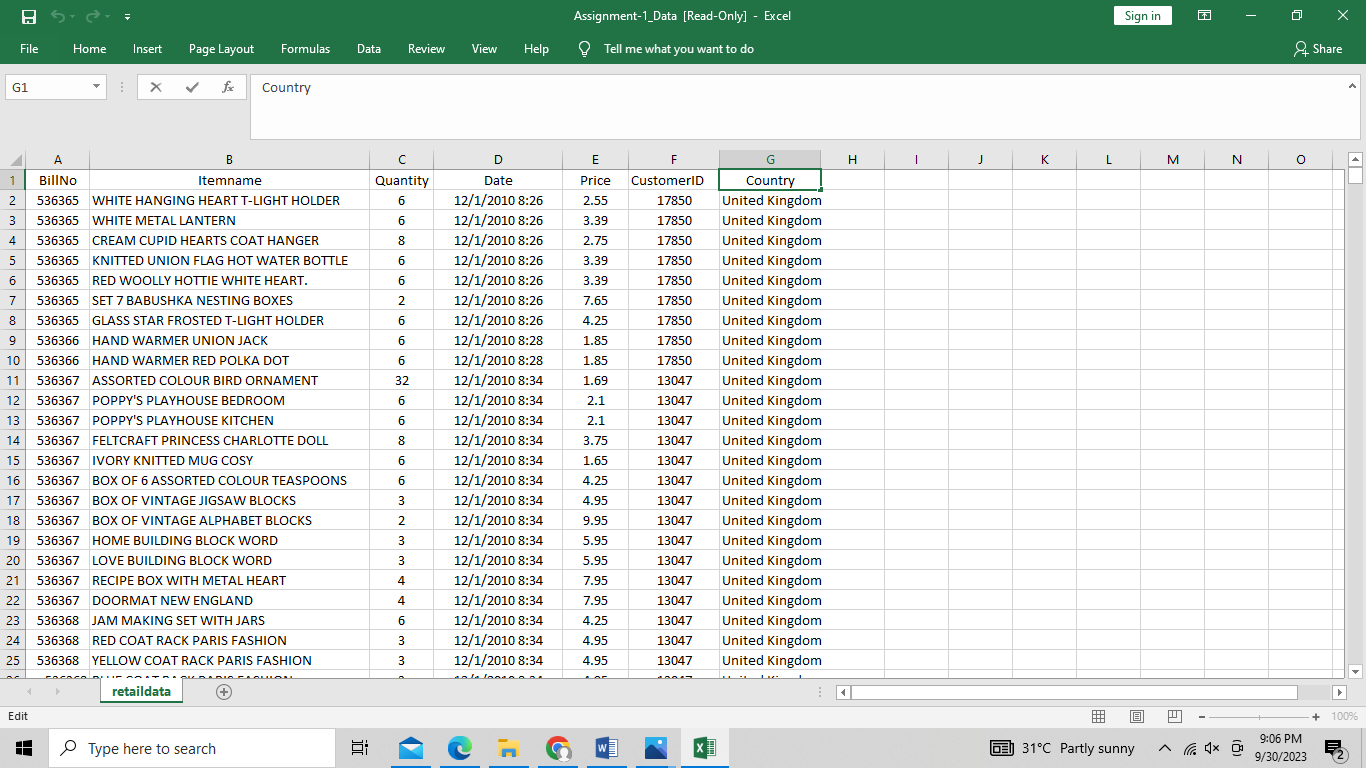
**Aim:**

**To Perform Market Basket Analysis on Given Dataset to Unveil Hidden Patterns and association Between the Products**

**Given Set Of DataBase:**

**DataSet Link:**

**https://www.kaggle.com/datasets/aslanahmedov/market-basket-analysis**



**Example Code:**

from mlxtend.preprocessing import TransactionEncoder

from mlxtend.frequent\_patterns import apriori, association\_rules

transactions = [Given data set]

te = TransactionEncoder()

te\_ary = te.fit\_transform(transactions)

df = pd.DataFrame(te\_ary, columns=te.columns\_)

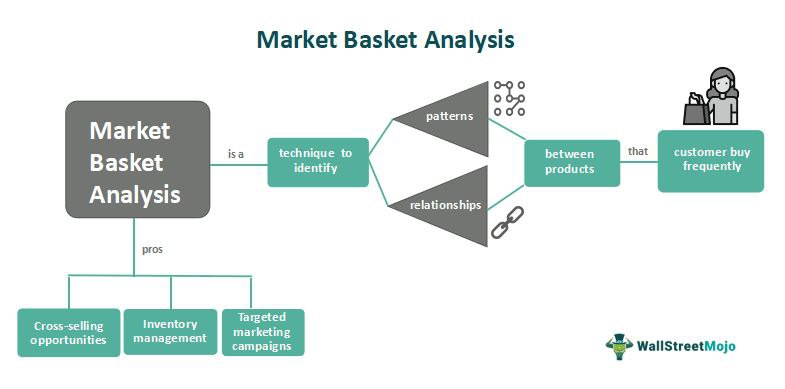
frequent\_itemsets = apriori(df, min\_support=0.2, use\_colnames=True)

rules = association\_rules(frequent\_itemsets, metric="lift", min\_threshold=1)

print(rules)

**Procedure:**

1. Market basket analysis is a technique used in data mining and retail analytics to uncover patterns of co-occurrence among items in a transactional dataset.
2. By analyzing customer purchase data, market basket analysis aims to identify associations or relationships between products that are frequently purchased together.
3. For Example Amazon uses Market Basket Analysis



The main steps involved in market basket analysis include data collection, data preparation, itemset generation, rule generation, and rule evaluation.

1. Data Collection:

First, transactional data is collected, which typically consists of customer purchases in a retail setting. Each transaction represents a collection of items purchased together.

1. Data Preparation:

The data is preprocessed to remove noise and outliers, and formatting the data into a suitable input format for analysis.

1. Itemset Generation:

Next, frequent itemsets are generated. Itemsets refer to groups of items that often appear together in transactions. The most common algorithm used for this task is the Apriori algorithm.

1. Rule Generation:

From the frequent itemsets, association rules are generated. Association rules describe the relationships between items based on their co-occurrence in transactions. These rules are typically in the form of "If item A is purchased, then item B is also likely to be purchased."

1. Rule Evaluation:

Finally, the generated rules are evaluated based on various metrics such as support, confidence, and lift. These metrics help determine the strength and significance of the associations between items.

The insights gained from market basket analysis can be used by businesses in several ways. For example, retailers can use the knowledge of item associations to optimize store layouts, plan product placements, and improve cross-selling opportunities. Online businesses can use these insights to personalize recommendations and enhance the customer shopping experience.

