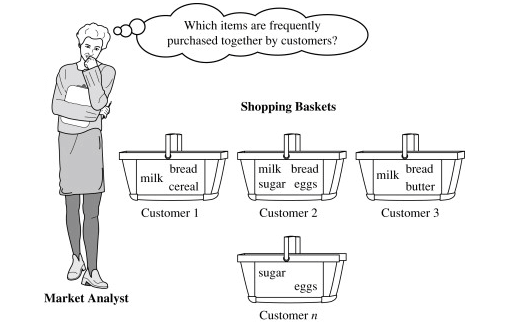
**MARKET BASKET INSIGHTS**

Market basket analysis is a strategic data mining technique used by retailers to enhance sales by gaining a deeper understanding of customer purchasing patterns. This method entails the examination of substantial datasets, such as historical purchase records, in order to unveil inherent product groupings and identify items that tend to be bought together.

By recognizing these patterns of co-occurrence, retailers can make informed decisions to optimize inventory management, devise effective marketing strategies, employ cross-selling tactics, and even refine store layout for improved customer engagement.

For example, if customers are buying milk, how probably are they to also buy bread (and which kind of bread) on the same trip to the supermarket? This information may lead to an increase in sales by helping retailers to do selective marketing based on predictions, cross-selling, and planning their ledge space for optimal product placement.

Now, just think of the universe as the set of items available at the store, then each item has a Boolean variable that represents the presence or absence of that item. Now each basket can then be represented by a Boolean vector of values that are assigned to these variables. The Boolean vectors can be analyzed for purchase patterns that reflect items that are frequently associated or bought together. Such patterns will be represented in the form of association rules.



Source: Sciencedirect

## **How Does Market Basket Analysis Work?**

1. Collect data on customer transactions, such as the items purchased in each transaction, the time and date of the transaction, and any other relevant information.
2. Clean and preprocess the data, removing any irrelevant information, handling missing values, and converting the data into a suitable format for analysis.
3. Use association rules mining algorithms such as Apriori or FP-Growth to identify frequent item sets, sets of items often appearing together in a transaction.
4. Calculate the support and confidence for each frequent itemset, which expresses the likelihood of one item being purchased given the purchase of another item.
5. Generate association rules based on the frequent itemsets and their corresponding support and confidence values. Association rules express the likelihood of one item being purchased given the purchase of another item.
6. Interpret the results of the market basket analysis, identifying which items are frequently purchased together, the strength of the association between items, and any other relevant insights into customer behavior and preferences.
7. Use the insights from the market basket analysis to inform business decisions such as product recommendations, store layout optimization, and targeted marketing campaigns.

## **3 Types of Market Basket Analysis**

1. It involves identifying frequent item sets and generating association rules that express the likelihood of one item being purchased with the purchase of another item. It is used to identify the relationships or associations between items in a transactional dataset.
2. This type of market basket analysis focuses on the order in which items are purchased in a transaction. It identifies frequent item sequences and generates sequential association rules describing the likelihood of one item sequence being followed by another.
3. This type of market basket analysis involves grouping similar items or transactions into clusters or segments based on their attributes. It helps to identify customer segments with similar purchasing behaviors, which can inform product recommendations and marketing strategies.

## **Algorithms Used in Market Basket Analysis**

There are multiple data mining techniques and algorithms used in Market Basket Analysis. One of the important objectives is “to predict the probability of items that are being bought together by customers.”

1. Apriori Algorithm
2. AIS
3. SETM Algorithm
4. FP Growth

### 1. **Apriori Algorithm**

Apriori Algorithm is a widely-used and well-known Association Rule algorithm and is a popular algorithm used in market basket analysis. It is also considered accurate and overtop AIS and SETM algorithms. It helps to find frequent itemsets in transactions and identifies association rules between these items. The limitation of the Apriori Algorithm is frequent itemset generation. It needs to scan the database many times, leading to increased time and reduced performance as a computationally costly step because of a large dataset. It uses the concepts of Confidence and Support.

### 2. **AIS Algorithm**

The AIS algorithm creates multiple passes on the entire database or transactional data. During every pass, it scans all transactions. As you can see, in the first pass, it counts the support of separate items and determines then which of them are frequent in the database. Huge itemsets of every pass are enlarged to generate candidate itemsets. After each scanning of a transaction, the common itemsets between the itemsets of the previous pass and the items of this transaction are determined. This algorithm was the first published algorithm which is developed to generate all large itemsets in a transactional database. It focused on the enhancement of databases with the necessary performance to process decision support. This technique is bounded to only one item in the consequent.

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

PRADEEPA.K 11.10.2023

NANDHA COLLEGE OF TECHNOLOGY