## **Market Basket Analysis**

**Problem Statement**

* Although individual customer transaction data is not available, infer potential product associations within departments using sales data.
* Develop cross-selling strategies based on these inferences.

## **What Is Market Basket Analysis?**

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 involves examining substantial datasets, such as historical purchase records, to unveil inherent product groupings and identify items that customers tend to buy 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 likely 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.

## **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, expressing 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 indicate the likelihood of purchasing one item given the purchase of another item.
6. Interpret the results of the market basket analysis, identifying frequent purchases, assessing the strength of the association between items, and uncovering other relevant insights into customer behavior and preferences.

## **Types of Market Basket Analysis**

1. Predictive Market Basket Analysis employs supervised learning to forecast future customer behavior. By recognizing cross-selling opportunities through purchase patterns, it enables applications like tailored product recommendations, personalized promotions, and effective demand forecasting. Additionally, it proves valuable in fraud detection.
2. Differential Market Basket Analysis compares purchase histories across diverse segments, unveiling trends and pinpointing buying habits unique to specific customer groups. Its applications extend to competitor analysis, identification of seasonal trends, customer segmentation, and insights into regional market dynamics.

## 

## 

## **What Is Association Rule for Market Basket Analysis?**

Let I = {I1, I2,…, Im} be an itemset. These item sets are called antecedents. Let D, the data, be a set of database transactions where each transaction T is a nonempty itemset such that T ⊆ I. Each transaction is associated with an identifier called a TID(or Tid). Let A be a set of items(itemset). T is the Transaction that is said to contain A if A ⊆ T. An Association Rule is an implication of form A ⇒ B, where A ⊂ I, B ⊂ I, and A ∩B = φ.

The rule A ⇒ B holds in the data set(transactions) D with supports, where ‘s’ is the percentage of transactions in D that contain A ∪ B (i.e., the union of set A and set B, or both A and B). This is taken as the probability, P(A ∪ B). Rule A ⇒ B has confidence c in the transaction set D, where c is the percentage of transactions in D containing A that also contains B. This is taken to be the conditional probability, like P(B|A). That is,

* *support(A⇒ B) =P(A ∪ B)*
* *confidence(A⇒ B) =P(B|A)*

Rules that meet both a minimum support threshold (called min sup) and a minimum confidence threshold (called min conf) are termed as ‘Strong’

* *Confidence(A⇒ B) = P(B|A) =*
* *support(A ∪ B) /support(A) =*
* *support count(A ∪ B) / support count(A)*

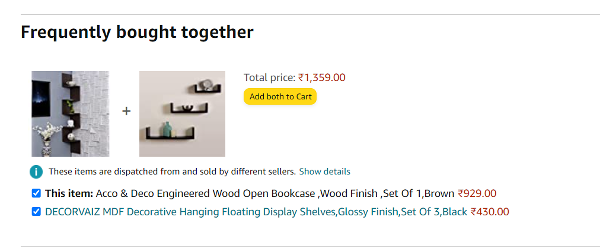
Generally, Association Rule Mining can be viewed in a two-step process:

1. Find all Frequent itemsets: *By definition, each of these itemsets will occur at least as  
   frequently as a pre-established minimum support count, min sup*.
2. Generate Association Rules from the Frequent itemsets: *By definition, these  
   rules must satisfy minimum support and minimum confidence.*

### **Association Rule Mining**

Association Rule Mining is primarily used when you want to identify an association between different items in a set and then find frequent patterns in a transactional database or relational database.

The best example of the association is as you can see in the following image.



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

### **Apriori Algorithm**

The Apriori Algorithm widely uses and is well-known for Association Rule mining, making it a popular choice in market basket analysis. AI and SETM algorithms consider it more accurate. 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.

## 

## 

## **Advantages of Market Basket Analysis**

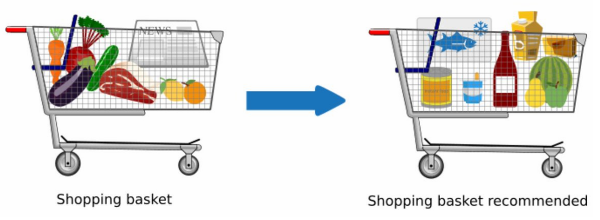
There are many advantages to implementing Market Basket Analysis in marketing. Market Basket Analysis (MBA) applies to customer data from point of sale (PoS) systems.

It helps retailers in the following ways:

* Increases customer engagement
* Boosts sales and increases RoI
* Improves customer experience
* Optimizes marketing strategies and campaigns
* Helps in demographic data analysis
* Identifies customer behavior and pattern

## **Market Basket Analysis From the Customers’ Perspective**

Let us take an example of market basket analysis from Amazon, the world’s largest eCommerce platform. From a customer’s perspective, Market Basket Analysis is like *shopping at a supermarket*. Generally, it observes all items bought by customers together in a single purchase. Then it shows the most related products together that customers will tend to buy in one purchase.



**Inferences**

In market basket analysis we have tried to find out the association within different departments with the different stores in the given sales data

**Market Basket Analysis Objective:**

The goal is to uncover associations within different departments across different stores using sales data. This entails identifying which departments tend to have items purchased together across stores.

**Features in Sales Data:**

The sales data includes features such as department number, store number, date, weekly sales, and an "is holiday" feature. These features are crucial for analyzing transactional patterns and identifying associations between departments.

**Output:**

The output of the analysis consists of association rules that reveal relationships between departments. Each rule specifies a set of items (antecedent) that are associated with another set of items (consequent). For example, a rule might indicate that when department 29 is present, items in stores 1, 2, 3, ..., 45 tend to be purchased together.

**Implementation:**

The Apriori algorithm is applied to the sales data with specific parameters such as minimum support, confidence, lift, and length. This algorithm efficiently discovers frequent itemsets and generates association rules based on these itemsets.

**Association Results:**

The association results are printed out, showing each association rule along with its support, confidence, and lift. For example, a rule might indicate that when department 29 is present, there is a high confidence (e.g., 95%) that stores 1, 2, 3, ..., 45 will also be present in the same transaction, with a lift value indicating the strength of the association.

**Grouping by Department Number and Date:**

Additionally, grouping the data by department number and date, potentially to analyze sales trends and associations within each department over time.

Overall, our approach encompasses standard practices in market basket analysis and leverages the Apriori algorithm to uncover meaningful associations within the provided sales data, which can lead to actionable insights for optimizing product placement, cross-selling strategies, and marketing efforts across different departments and stores.

## **Develop cross-selling strategies based on these inferences**

Based on the associations identified among different departments and stores, we can tailor cross-selling strategies to optimize sales and enhance customer experience.

Here's how we can interpret and utilize the inferred associations:

**Optimize Store Layout:**

Utilize the associations between departments and stores to optimize the layout of each store. Place related departments in close proximity to each other within the store to encourage cross-store purchases. For example, if departments 29 and 30 are frequently purchased together across various stores, consider placing them adjacent to each other to facilitate cross-department sales.

**Create Bundled Offers:**

Identify departments that exhibit strong associations across multiple stores, such as departments 35 and 36. Create bundled offers or package deals that include products from both departments to incentivize customers to make cross-department purchases.

**Targeted Promotions:**

Develop targeted promotions and marketing campaigns based on the associations between departments and stores. For instance, if department 37 frequently co-occurs with departments 1, 2, 4, and 40 across different stores, create promotions that span these departments to capitalize on customer purchase patterns.

**Cross-Store Loyalty Programs:**

Implement cross-store loyalty programs that reward customers for making purchases across different departments and stores. Encourage customers to explore a wider range of products by offering incentives, discounts, or loyalty points for cross-store purchases.

**Dynamic Pricing Strategies:**

Adjust pricing strategies based on the associations between departments and stores. Offer discounts or special pricing for products that are frequently purchased together across different stores to stimulate sales and increase customer satisfaction.

**Inventory Management:**

Optimize inventory management by stocking related products from associated departments across different stores. Ensure sufficient stock levels for items that are commonly purchased together to meet customer demand and prevent stockouts.

**Data-Driven Decision Making:**

Continuously analyze sales data and customer purchase patterns to refine cross-selling strategies. Leverage insights from association analysis to make data-driven decisions regarding product assortment, store layout, promotions, and pricing strategies.

By leveraging the associations identified among different departments and stores, we can develop targeted cross-selling strategies that enhance the shopping experience, increase sales, and drive customer loyalty across the entire retail network.

## **Conclusion**

In this tutorial, we discussed Market Basket Analysis and learned the steps to implement it from scratch using Python. We then implemented Market Basket Analysis using Apriori Algorithm. We also looked into the various uses and advantages of this algorithm and learned that we could also use FP Growth and AIS algorithms to implement Market Basket Analysis.

**Key Takeaways**

* Businesses use Market Basket Analysis as a strategy to design store layouts based on customers’ shopping behavior and purchase histories.
* This idea is also applicable to machine learning algorithms to teach machines to help businesses, especially in the e-commerce sector.
* In this article, we have gone through a step-by-step guide to implementing the apriori algorithm in Python and also looked into the math behind the association rules.

**Reference**

<https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/>