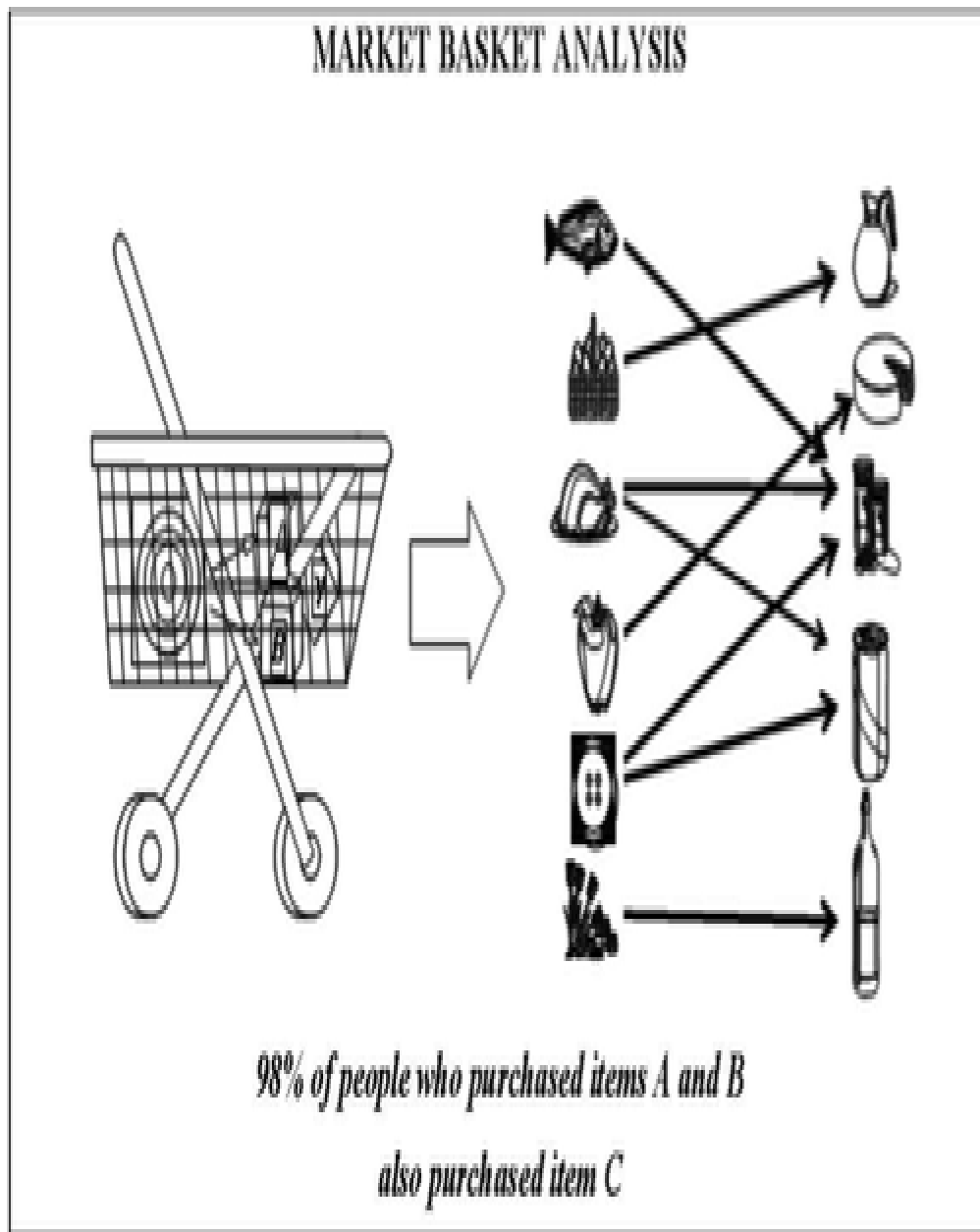


# MARKET BASKET ANALYSIS:



# MARKET BASKET ANALYSIS

## Abstract:

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.

## Introduction:

Nowadays, Machine Learning is helping the Retail Industry in many different ways. You can imagine that from forecasting the performance of sales to identifying the buyers, there are many applications of AI and ML in the retail industry. Market basket analysis is a data mining technique retailers use to increase sales by better understanding customer purchasing patterns. It involves analyzing large data sets, such as purchase history, to reveal product groupings and products likely to be purchased together. In this article, we will comprehensively cover the topic of Market Basket Analysis and its various components and then dive deep into the ways of implementing it in machine learning, including how to perform it in Python on a real-world dataset.

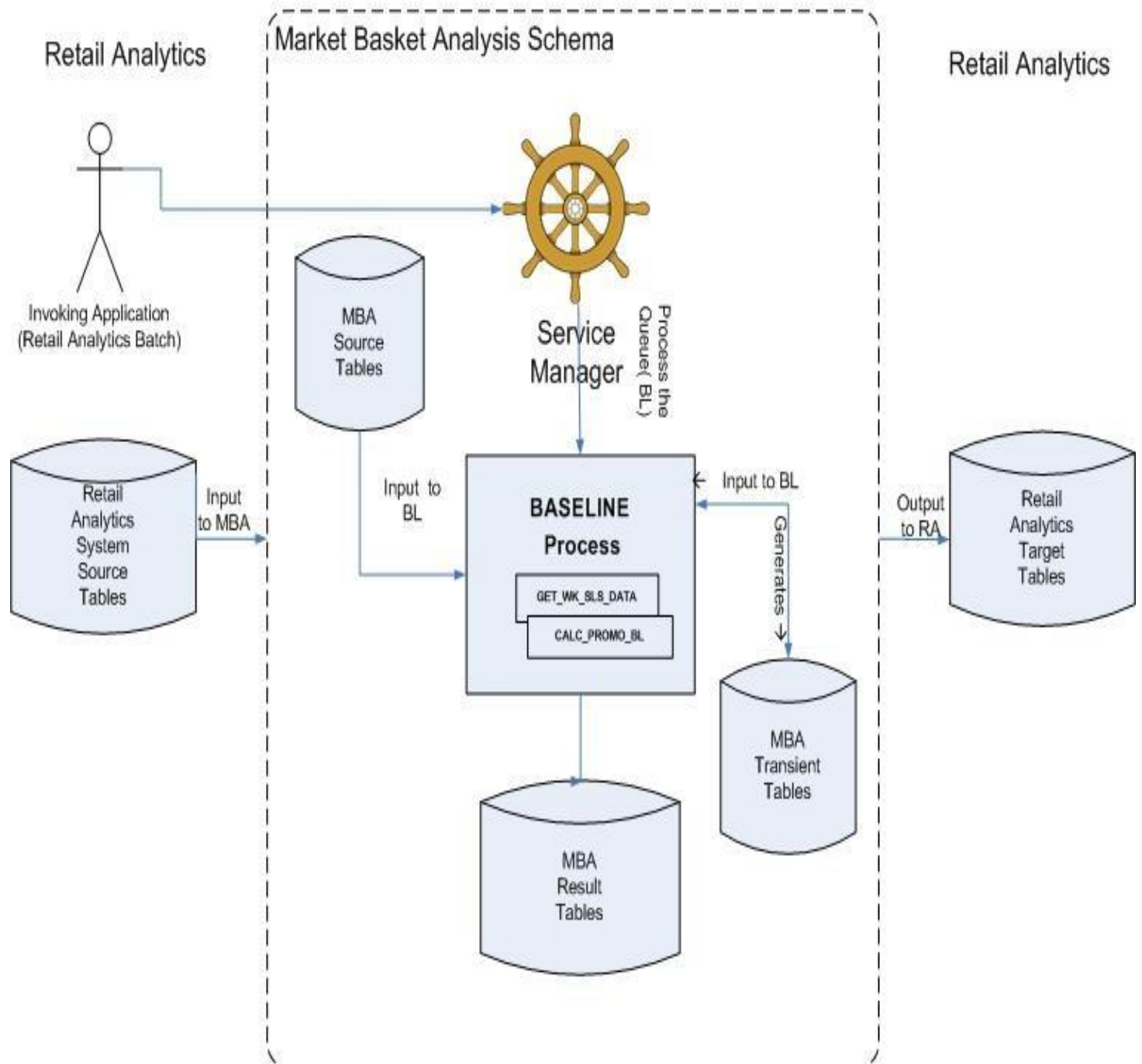
# About market basket analysis:

There are three domains where market basket analysis is used for.

- ❖ The first domain is the creation of personalized recommendations . This methodology is well known nowadays. During the explosion of the e-commerce, personalized recommendations has appeared as a part of the marketing process. Basically, the idea consist in suggesting items to customer based on its preferences. There are two basic ways to do it. The first one, is suggesting items similars the ones the customer has purchased in the past. The second one, is looking for similar customers and recommending items that had purchased the others. Both strategies are often used for companies in order to realize cross-selling and upselling strategies.
- ❖ The second domain where market basket analysis is used is in the analysis of spatial distribution in chain stores . Due the increasing number of products that nowadays exist, physical space in stores has started to be a problem. More and more, stores invest money and time trying to find which distribution of items can lead them to obtain more sells. Due that, knowing in advance which items are commonly purchased together, the distribution of the store can be changed in order to sell more products.
- ❖ The last domain is in the creation of discounts and promotions. Based in customers behaviour, special sales can be performed. For example, if the client knows which items are often purchased together, he can create new offers based in order to increase the sells of those items. As it can be seen, market basket analysis can be used to help retail business in many fields. That's why the client contacted with CleverData, to help him to discover, which associations rules were in its stores.

# Flowchart:

MBA Baseline Process Flow Diagram



## Performance:

Market basket analysis is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analyzing large data sets, such as purchase history, to reveal product groupings, as well as products that are likely to be purchased together.

## Machine learning algorithm:

- The algorithm starts by **generating an itemset through the Join Step**, that is to generate  $(K+1)$  itemset from  $K$ -itemsets. For example, the algorithm generates Cookie, Chocolate and Cake in the first iteration.
- Immediately after that, the algorithm proceeds with the Prune Step, that is to **remove any candidate item set that does not meet the minimum support requirement**. For example, the algorithm will remove Cake if  $\text{Support}(\text{Cake})$  is below the predetermined minimum Support.

It iterates both of the steps until there are no possible further extensions left.

# Program:

```
def draw_network(rules, rules_to_show):
    # Directional Graph from NetworkX
    network = nx.DiGraph()

    # Loop through number of rules to show
    for i in range(rules_to_show):

        # Add a Rule Node
        network.add_nodes_from(["R"+str(i)])
        for antecedents in rules.iloc[i]['antecedents']:
            # Add antecedent node and link to rule
            network.add_nodes_from([antecedents])
            network.add_edge(antecedents, "R"+str(i), weight = 2)

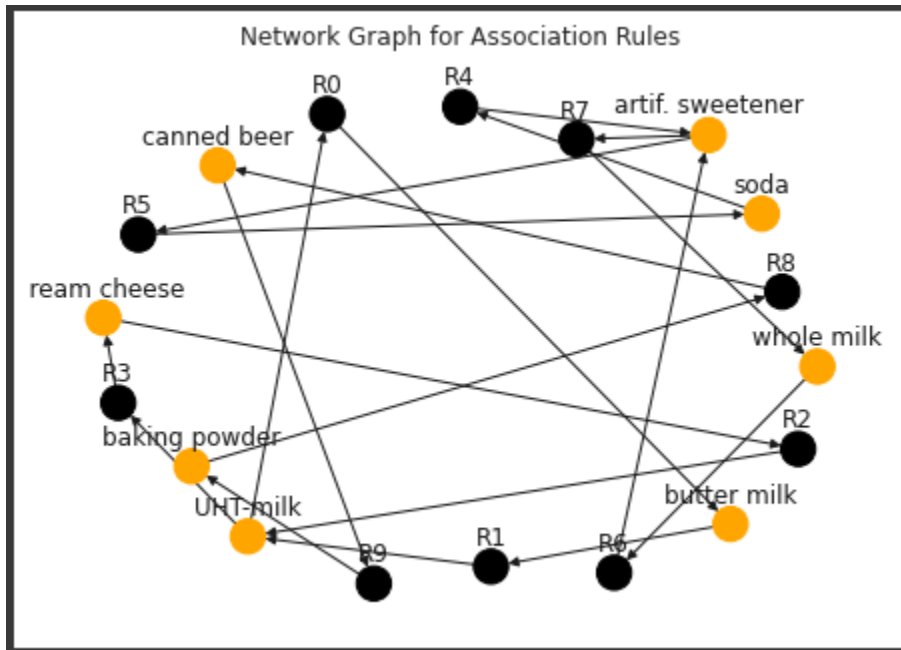
        for consequents in rules.iloc[i]['consequents']:
            # Add consequent node and link to rule
            network.add_nodes_from([consequents])
            network.add_edge("R"+str(i), consequents, weight = 2)
        color_map=[]

    # For every node, if it's a rule, colour as Black, otherwise
    Orange
    for node in network:
        if re.compile("^[R]\d+$").fullmatch(node) != None:
            color_map.append('black')
        else:
            color_map.append('orange')

    # Position nodes using spring layout
    pos = nx.spring_layout(network, k=16, scale=1)
    # Draw the network graph
    nx.draw(network, pos, node_color = color_map, font_size=8)

    # Shift the text position upwards
    for p in pos:
        pos[p][1] += 0.12
    nx.draw_networkx_labels(network, pos)
    plt.title("Network Graph for Association Rules")
    plt.show()
draw_network(rules, 10)
```

## Output:



## Conclusion:

In this, 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.