***Development part-2***

Market Basket Insights

Introduction:

* **Market basket analysis is a**[**data mining**](https://searchsqlserver.techtarget.com/definition/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.**
* **The adoption of market basket analysis was aided by the advent of electronic point-of-sale (POS) systems. Compared to handwritten records kept by store owners, the digital records generated by POS systems made it easier for applications to process and**[**analyze large volumes of purchase data**](https://www.techtarget.com/searchcustomerexperience/tip/6-ways-to-use-analytics-to-improve-customer-engagement)**.**
* **Implementation of market basket analysis requires a background in statistics and**[**data science**](https://www.techtarget.com/searchenterpriseai/definition/data-science)**, as well as some algorithmic computer programming skills. For those without the needed technical skills, commercial, off-the-shelf tools exist.**

**Given dataset:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| S.No | Bill No | Itemname | Quantity | Date | Price | Customer ID | Country |
| 1 | 536365 | WHITE HANGING HEART T-LIGHT HOLDER | 6 | ####### | 2.55 | 17850 | United Kingdom |
| 2 | 536365 | WHITE HANGING HEART | 6 | ####### | 3.39 | 17850 | United Kingdom |
| 3 | 536365 | CREAM CUPID HEARTS COAT HANGER | 8 | ####### | 2.75 | 17850 | United Kingdom |
| 4 | 536365 | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | ####### | 3.39 | 17850 | United Kingdom |
| … | … | … | … | … | … | … | … |
| 522061 | 581587 | PACK OF 20 SPACEBOY NAPKINS | 12 | ####### | 0.85 | 12680 | France |
| 522062 | 581587 | CHILDREN'S APRON DOLLY GIRL | 6 | ####### | 2.1 | 12680 | France |
| 522063 | 581587 | CHILDRENS CUTLERY DOLLY GIRL | 4 | ####### | 4.15 | 12680 | France |
| 522064 | 581587 | CHILDRENS CUTLERY CIRCUS PARADE | 4 | ####### | 4.15 | 12680 | France |
| 522065 | 581587 | BAKING SET 9 PIECE RETROSPOT | 3 | ####### | 4.95 | 12680 | France |

500000 Rows × 8 Columns

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

**Overview of the process:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

To continue building the Market Basket Insights project, we'll perform association analysis and generate insights using the Apriori algorithm and some post-analysis techniques. Association analysis helps identify patterns and relationships between items in a transaction dataset, allowing you to gain insights into customer behavior and make data-driven decisions. Here's a step-by-step guide:

**1.Data Preprocessing:**

Before performing association analysis, ensure your data is clean and formatted correctly. Your dataset should have a transaction ID and a list of items purchased in each transaction.

**2. Library Setup:**

You'll need to use Python and some libraries for this analysis. Ensure you have the necessary libraries installed, such as Pandas, NumPy, and MLxtend.

Program:

import pandas as pd

from mlxtend.frequent\_patterns import apriori

from mlxtend.frequent\_patterns import association\_rules

3. Load and Prepare Data:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |

Load your transaction data into a Pandas DataFrame and encode the items as binary variables (1 if the

item is present in the transaction, 0 otherwise).

Program:

# Load your data

data = pd.read\_csv('your\_transaction\_data.csv')

# Convert data to a binary format

basket\_sets = pd.get\_dummies(data, columns=['item\_column\_name']).groupby('transaction\_id').sum()

**4. Run Apriori Algorithm:**

Apply the Apriori algorithm to identify frequent item sets with a minimum support threshold. This step helps you find items that are often purchased together.

Program:

frequent\_itemsets = apriori(basket\_sets, min\_support=0.01, use\_colnames=True)

**5. Generate Association Rules:**

Discover association rules from the frequent item sets. Association rules consist of antecedents (items in the "if" part) and consequents (items in the "then" part).

Program:

association\_rules = association\_rules(frequent\_itemsets, metric="lift", min\_threshold=1.0)

6. Interpretation and Insights:

Now that you have the association rules, you can generate insights:

a. High Confidence and Lift:

Focus on rules with high confidence and lift values. High confidence indicates that the consequent item is often purchased when the antecedent item(s) are present. High lift suggests a stronger association compared to random chance.

b. Support:

Analyze the support of rules to identify how often they occur. Higher support values indicate that the rule applies to a significant portion of your dataset.

c. Visualization:

Create visualizations like scatter plots, bar charts, or network graphs to present the most interesting and actionable association rules. Tools like Matplotlib or network analysis libraries can be helpful.

d. Business Insights:

Based on the association rules and your domain knowledge, draw actionable insights. For example, you can recommend item placements, bundling strategies, or marketing campaigns based on what items tend to be bought together.

e. Continuous Monitoring:

Set up a process to regularly update your association analysis to adapt to changing customer preferences and market dynamics.

7. Implementation:

Use the insights gained from the association analysis to make data-driven decisions. Implement strategies to maximize sales, improve customer satisfaction, or optimize inventory based on the discovered patterns.

Procedure: **Feature selection:**

* Identifying product associations

Market basket analysis helps businesses uncover hidden relationships and associations between products. By understanding which items are frequently purchased together, companies can optimize their inventory management, identify product bundling opportunities, and improve overall sales.

* Cross-selling and upselling opportunities

MBA enables businesses to identify cross-selling and upselling opportunities. By analyzing customer purchase patterns, companies can recommend complementary or higher-value products to customers, thereby increasing average order value and driving additional revenue.

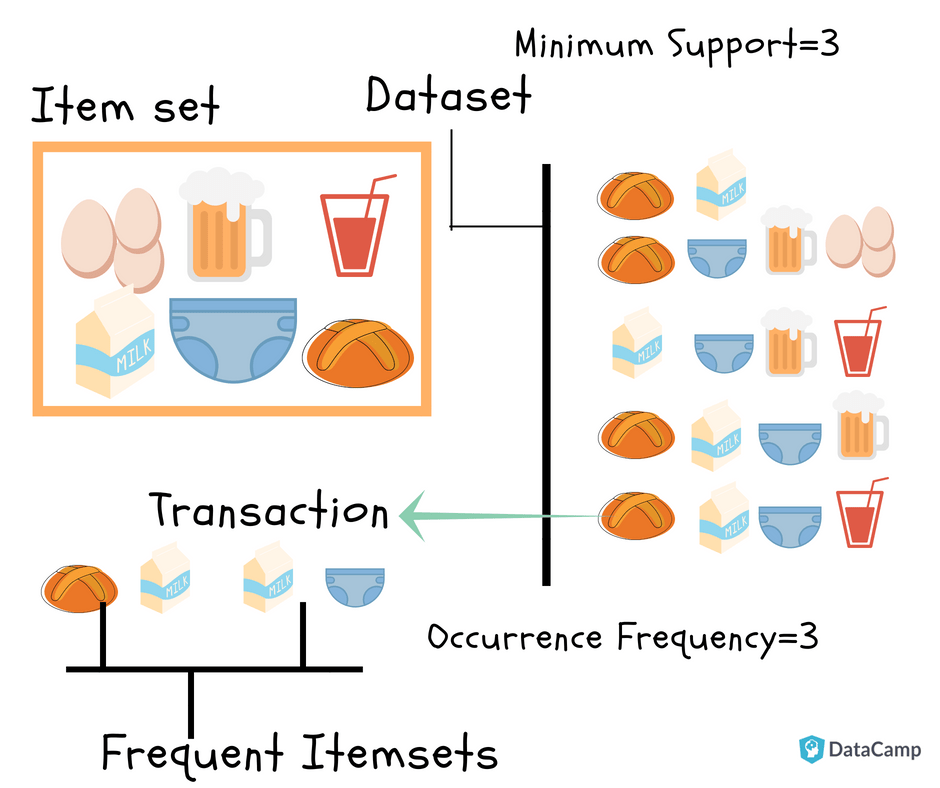
* Enhancing customer understanding

By gaining insights into customer behavior and preferences, market basket analysis allows businesses to better understand their customers. This understanding can help tailor marketing campaigns, personalize communications, and create targeted offers, leading to improved customer satisfaction and loyalty.

* Optimizing product placementMarket basket analysis can assist in optimizing product placement within physical stores or on e-commerce platforms. By identifying frequently associated products, businesses can strategically position items to enhance visibility, encourage cross-category purchases, and increase overall sales.
* Personalizing recommendations

Leveraging market basket analysis, businesses can provide personalized recommendations to customers. By recommending products based on a customer's past purchase history or items frequently purchased together, companies can enhance the customer experience, drive repeat purchases, and foster long-term customer relationships.

**Market Basket Analysis Using Apriori**



Remember that association analysis is an ongoing process. Regularly update your data and re-run the analysis to adapt to changing market conditions and customer behaviors.

By following these steps, you can perform association analysis and generate actionable insights for your Market Basket Insights project.

Program:

import pandas as pd

from mlxtend.frequent\_patterns import apriori

from mlxtend.frequent\_patterns import association\_rules

data = pd.read\_csv('your\_transaction\_data.csv')

basket\_sets = pd.get\_dummies(data, columns=['item\_column\_name']).groupby('transaction\_id').sum()

frequent\_itemsets = apriori(basket\_sets, min\_support=0.01, use\_colnames=True)

association\_rules = association\_rules(frequent\_itemsets, metric="lift", min\_threshold=1.0)

Output:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| S.No | Bill No | Itemname | Quantity | Date | Price | Customer ID | Country |
| 1 | 536365 | WHITE HANGING HEART T-LIGHT HOLDER | 6 | ####### | 2.55 | 17850 | United Kingdom |
| 2 | 536365 | WHITE HANGING HEART | 6 | ####### | 3.39 | 17850 | United Kingdom |
| 3 | 536365 | CREAM CUPID HEARTS COAT HANGER | 8 | ####### | 2.75 | 17850 | United Kingdom |
| 4 | 536365 | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | ####### | 3.39 | 17850 | United Kingdom |
| … | … | … | … | … | … | … | … |
| 522061 | 581587 | PACK OF 20 SPACEBOY NAPKINS | 12 | ####### | 0.85 | 12680 | France |
| 522062 | 581587 | CHILDREN'S APRON DOLLY GIRL | 6 | ####### | 2.1 | 12680 | France |
| 522063 | 581587 | CHILDRENS CUTLERY DOLLY GIRL | 4 | ####### | 4.15 | 12680 | France |
| 522064 | 581587 | CHILDRENS CUTLERY CIRCUS PARADE | 4 | ####### | 4.15 | 12680 | France |
| 522065 | 581587 | BAKING SET 9 PIECE RETROSPOT | 3 | ####### | 4.95 | 12680 | France |

500000 Rows × 8 Columns

Conclusion:

In addition its popularity as a retailer’s technique, Market Basket Analysis is applicable in many other areas:

* In the manufacturing industry for predictive analysis of equipment failure.
* In Pharmaceutical/Bioinformatics for the discovery of co-occurrence relationships among diagnosis and pharmaceutical active ingredients prescribed to different patient groups.
* In Banking/Criminology for fraud detection based on credit card usage data.
* For analyzing customer behavior by associating purchases with demographic and socio-economic data.

More and more organizations are discovering ways of using market basket analysis to gain useful insights into associations and hidden relationships. As industry leaders continue to explore the technique’s value, a predictive version of market basket analysis is making in-roads across many sectors in an effort to identify sequential purchases.