جامعة أم القرى
UMM AL-QURA UNIVERSITY

Collage of computing

Data Analysis 2

COURSE PRESENTER

(DR. Omaima A. Fallatah)

DURATION:

25 Oct 2024

Task 2

“Market Basket”

Student Name	ID
Reuf Bakr Al-sharif	443002428
Jana Hesham Al-Hless	443002347

TABLE OF CONTACT

INTRODUCTION

Market analysis is a vital process that aims to understand consumer behavior and preferences, as well as identify current market trends. It is often used to find growth opportunities and make strategic decisions about product development, marketing, and inventory management. One powerful technique within market analysis is Market Basket Analysis, which identifies patterns in purchase data and reveals relationships between products that customers frequently buy together.

Market Basket Analysis is widely used in the retail sector to recognize the most popular products or understand common purchase behaviors. This technique, also known as Association Rule Mining, helps businesses extract patterns from data to identify items that are commonly purchased together. These insights can support various strategies, such as product placement, cross-selling, and targeted marketing.

DATASET

The data used in this analysis comes from the

[Market Basket Analysis Dataset on Kaggle]

(https://www.kaggle.com/datasets/aslanahmedov/market-basket-analysis) which provides comprehensive transaction records for retail analysis.

Overview of the Dataset:

The dataset contains transaction data from a retail store with the following columns :

BillNo: Unique identifier for each transaction.

Itemname: Description of the item purchased.

Quantity: Number of units purchased per item.

Date: Date of purchase.

Price: Price per item unit.

CustomerID: Unique identifier for each customer.

Country: Country where the transaction occurred.

OBJECTIVES

The primary objective of Market Basket Analysis is to understand patterns in customer purchasing behavior by identifying associations between products commonly bought together. These insights help in making data-driven decisions regarding inventory management, product placement, cross-selling, and personalized marketing strategies. Market Basket Analysis enables businesses to improve customer experience by offering more relevant product recommendations and efficient store layouts, which ultimately supports sales growth.

In this project, we aim to apply Market Basket Analysis techniques to a retail transaction dataset (available here). The dataset contains details of supermarket transactions, with information such as BillNo, Itemname, Quantity, Price, CustomerID, and Country. By mining this dataset, we will uncover frequent itemsets and association rules that can provide actionable insights into customer purchasing patterns within the supermarket context. These insights can be applied to:

- **Optimize Product Placement**:

By placing frequently purchased items closer together, the store layout can be adjusted to encourage sales.

- **Enhance Marketing Strategies**:

Use identified associations to design targeted promotions.

- **Improve Inventory Management**:

Predict which items are more likely to be purchased together to reduce stockouts and manage restocking efficiently.

STEP 1 | DATA CLEANING AND PREPROCESSING

1.1 Handling Missing Values:

Checked for any missing values across all columns to ensure data completeness and accuracy. Identifying missing values allows for informed decisions on data cleaning to avoid inaccurate analyses that may result from incomplete data.

1.2 Removing Extra Spaces in Item Names:

Removed any leading or trailing spaces in item names to ensure consistency. This standardization improves the accuracy of analysis, especially when identifying frequently purchased or top-selling items.

1.3 Converting Bill Number (BillNo) to String:

Converted the BillNo column to a string format to treat it as a unique transaction identifier rather than a numerical value. This prevents any unintended calculations and ensures that each BillNo remains as an identifier for each transaction.

1.4 Converting Date (Date) to Datetime Format:

Transformed the Date column to a datetime format to support time-based analysis. This conversion allows for seasonal or peak purchase pattern analysis, providing insights into the timing of purchases.

1.5 Converting Price (Price) to Numeric Format:

Standardized the Price column to a numeric format, allowing for necessary calculations like total spending per transaction. This conversion enhances the accuracy of analyses that rely on price data.

1.6 Formatting Customer ID (CustomerID):

Converted CustomerID to string format and removed any unnecessary suffixes like .0 to ensure consistency across records. This standardization ensures each ID remains unique without any undesired modifications.

STEP 2 | EXPLORATORY DATA ANALYSIS

2.1 Best Selling Items in Each Country:

During the exploratory analysis, the best-selling items were identified for each country. Below is a table and visualization that summarizes the top-selling products by quantity across different countries.

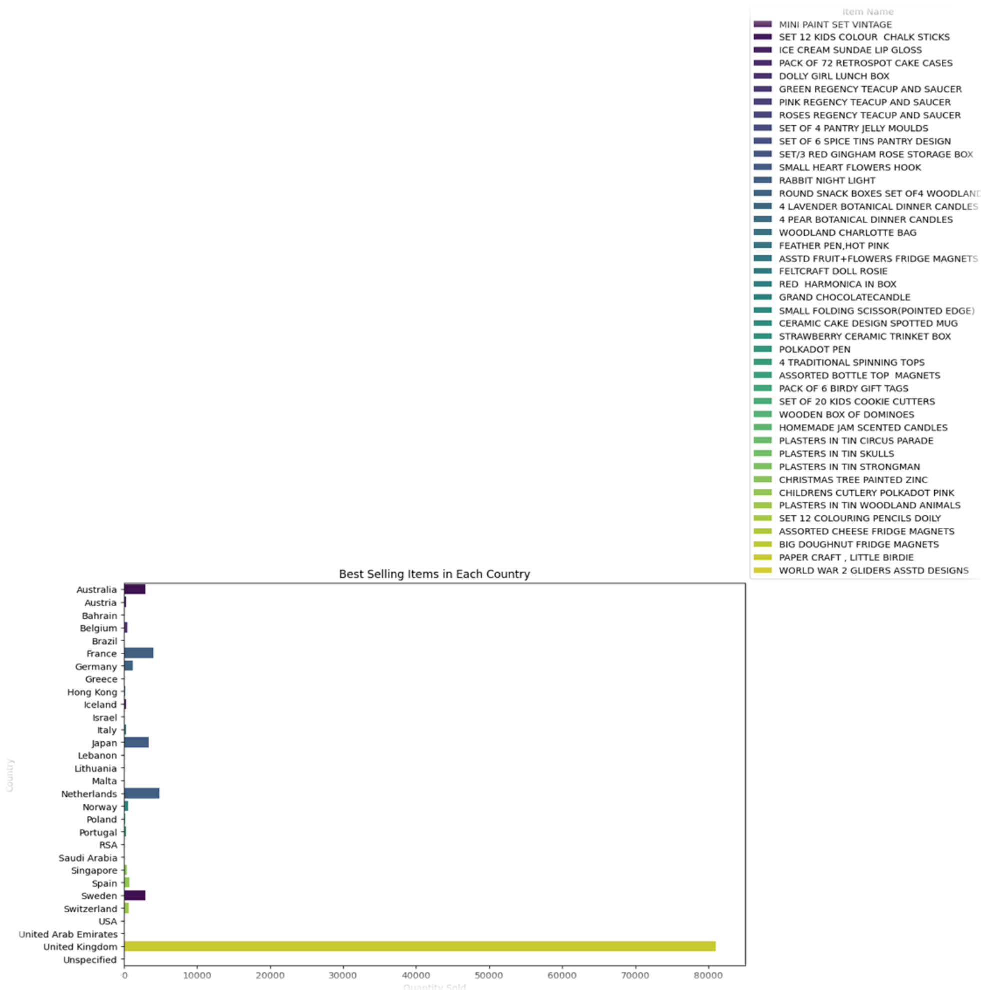
Key Findings:

The United Kingdom had the highest sales of "PAPER CRAFT, LITTLE BIRDIE" with a quantity of 80,995 units sold.

The Netherlands and France shared a common best-seller,

"RABBIT NIGHT LIGHT", with significant sales figures.

Countries like Brazil, Italy, and Saudi Arabia had different best-selling items with smaller quantities.



STEP 2 | (cont)..

2.2 Most Popular Items in the United Kingdom:
As part of our exploratory analysis, we focused on identifying the best-selling items in the United Kingdom. The analysis revealed that certain products performed exceptionally well in this market, with a clear preference for specific items.

PAPER CRAFT, LITTLE BIRDIE emerged as the highest-selling product, with an impressive total of 80,995 units sold. This item significantly outperformed others, indicating strong consumer demand for craft-related products in the UK.

MEDIUM CERAMIC TOP STORAGE JAR followed closely, with 77,036 units sold. This demonstrates a considerable interest in storage solutions, particularly ceramic storage products.

Other popular products include ***WORLD WAR 2 GLIDERS ASSTD DESIGNS***, which sold 49,430 units, ***and JUMBO BAG RED RETROSPOT***, with 44,165 units sold.

Additionally, products such ***as WHITE HANGING HEART T-LIGHT HOLDER*** also showed substantial sales, with 35,726 units sold, indicating consumer preferences toward home décor items.

On the other hand, some items, while part of the inventory, had minimal sales. For instance, ***SET/4 2 TONE EGG SHAPE MIXING BOWLS, FILIGREE DIAMANTE CHAIN***, and Boombox Ipod Classic were among the least sold items, each selling only 1 unit. These results suggest a lower demand for these particular items in the UK.

STEP 3 | Transaction for Association Rule Mining

The original dataset was transformed into a transactional format to facilitate the extraction and analysis of association rules. This transformation was done by grouping together items that were purchased within each invoice, so that each transaction represents a list of items bought together.

3.1 Steps Taken:

1 - Converting the Data into Transactional Format:

The data was grouped by invoice numbers (BillNo), collecting all purchased items together per invoice.

Each transaction now represents a list of items bought within the same invoice, making it possible to analyze which items are frequently purchased together.

2 - Creating the One-Hot Matrix:

After converting the data into transactional format, a One-Hot matrix was created. In this matrix, each product is represented as a column, with values of 1 (True) if the item was purchased in the transaction, and 0 (False) if it was not.

The BillNo column was then re-added to the One-Hot matrix, allowing us to retain the association between each transaction and its corresponding invoice.

3 - Organizing the One-Hot Matrix by Invoice:

The final One-Hot matrix was grouped by invoice numbers, creating a structured dataset that records which items were bought together within each transaction. This arrangement makes the data ready for further analysis.

STEP 3 | (cont)..

Results:

The final matrix includes 19,735 transactions (representing the number of (invoices) and 4,057 columns (representing each unique product.

Each invoice now has a list of items purchased, enabling us to perform analysis to identify purchasing patterns and extract association rules between items.

BillNo	Itemname	*Boombox Ipod Classic	*USB Office Mirror Ball	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 DAISY PEGS IN WOOD BOX	12 EGG HOUSE PAINTED WOOD	12 HANGING EGGS HAND PAINTED	12 IVORY ROSE PEG PLACE SETTINGS	12 MESSAGE CARDS WITH ENVELOPES	...	returned	taig adjust	test	to push order througha s stock was	website fixed	wrongly coded 20713	wrongly coded 23343	wrongly marked	wrongly marked 23343	wrongly sold (22719) barcode
536365	[WHITE HANGING HEART T-LIGHT HOLDER, WHITE MET...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
536366	[HAND WARMER UNION JACK, HAND WARMER RED POLKA...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
536367	[ASSORTED COLOUR BIRD ORNAMENT, POPPY'S PLAYHO...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
536368	[JAM MAKING SET WITH JARS, RED COAT RACK PARIS...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
536369	[BATH BUILDING BLOCK WORD]	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
581586	[LARGE CAKE STAND HANGING STRAWBERY, SET OF 3...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
581587	[CIRCUS PARADE LUNCH BOX, PLASTERS IN TIN CIRC...	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
A563185	[Adjust bad debt]	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
A563186	[Adjust bad debt]	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
A563187	[Adjust bad debt]	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

19735 rows × 4057 columns

3.2 Converting Product Columns to Binary Values:

As part of the data preparation for Market Basket Analysis, the product columns were converted into binary values (0 and 1). This step involved transforming any positive quantity values in the dataset to a 1 to indicate that the item was purchased, and 0 if it was not. This binary transformation allows us to focus on the presence or absence of each product in transactions rather than the quantity, which simplifies the data for analysis.

Results:

This binary format offers a streamlined representation of the data, where each product is indicated by its presence (1) or absence (0) in each transaction.

The resulting binary matrix facilitates the identification of purchasing patterns and relationships between products, as this simplified format is optimal for association rule mining and other pattern detection methods.

STEP 3 | (cont)..

3.3 Frequent Itemset Mining:
To uncover commonly purchased items and frequent combinations, we applied frequent itemset mining on the transaction data. Using the Apriori algorithm, we analyzed which items and item pairs are frequently bought together, identifying patterns in customer purchases. ***Steps:***

1 - Binary Conversion: Each product column was converted into a Boolean format (True/False) to indicate whether the item was purchased in each transaction.

2 - Applying the Apriori Algorithm: We then used the Apriori algorithm to find itemsets that meet a minimum support threshold of 1%. This threshold ensures that only itemsets appearing in at least 1% of transactions are included, filtering out infrequent items and highlighting popular choices.

Results :

The analysis produced a list of frequent itemsets that are commonly purchased together, helping identify products that frequently co-occur in transactions.

~~***Some of the top itemsets include :***~~

Colour Spaceboy Pen" with a support of 1.58% 10".

Message Cards with Envelopes" with a support of 1.26% 12".

Pairs of items like "Jumbo Bag Red Retrospot" and

"Jumbo Storage Bag Su", which appear together with a support of 1.12%.

STEP 4 | Association Rule Generation

To further analyze customer purchasing behavior, we applied Association Rule Mining on the frequent itemsets identified in the previous steps. The goal was to discover relationships between items frequently bought together, helping us identify which items or item pairs drive combined purchases. ***Steps:***

4.1 Setting Parameters: We generated association rules using the metric of lift with a minimum threshold of 1, ensuring that only significant and impactful associations were captured.

Results: The resulting association rules include information on

~~Antecedents:~~ Items or itemsets that, when bought, often lead to the purchase of another item (the consequent).

~~Consequents:~~ The items that are likely to be bought given the purchase of the antecedents.

~~Support:~~ The frequency with which the entire rule occurs in transactions.

~~Confidence:~~ The likelihood that a consequent is bought given the antecedent.

~~Lift:~~ A measure of how much the likelihood of purchasing the consequent increases when the antecedent is purchased, indicating the strength of the association.

Sample Results: One of the notable rules is that customers who purchased "DOTCOM POSTAGE" are also likely to buy "6 RIBBONS RUSTIC CHARM," with a high lift value of 5.98. This suggests a strong relationship between these items.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(DOTCOM POSTAGE)	(6 RIBBONS RUSTIC CHARM)	0.035875	0.047732	0.010236	0.285311	5.977290	0.008523	1.332422	0.863685
1	(6 RIBBONS RUSTIC CHARM)	(DOTCOM POSTAGE)	0.047732	0.035875	0.010236	0.214437	5.977290	0.008523	1.227305	0.874439
2	(JAM MAKING SET PRINTED)	(6 RIBBONS RUSTIC CHARM)	0.056549	0.047732	0.011806	0.208781	4.373992	0.009107	1.203545	0.817611
3	(6 RIBBONS RUSTIC CHARM)	(JAM MAKING SET PRINTED)	0.047732	0.056549	0.011806	0.247346	4.373992	0.009107	1.253499	0.810041
4	(JAM MAKING SET WITH JARS)	(6 RIBBONS RUSTIC CHARM)	0.055181	0.047732	0.010337	0.187328	3.924538	0.007703	1.171773	0.788715
...	...	...	...	...	...	...	...	...	...	...
3337	(CHARLOTTE BAG PINK POLKADOT)	(WOODLAND CHARLOTTE BAG, RED RETROSPOT CHARLOT...	0.037395	0.013073	0.010286	0.275068	21.040551	0.009797	1.361406	0.989475
3338	(RED RETROSPOT CHARLOTTE BAG)	(WOODLAND CHARLOTTE BAG, STRAWBERRY CHARLOTTE ...	0.052090	0.011198	0.010286	0.197471	17.633876	0.009703	1.232107	0.995127
3339	(WOODLAND CHARLOTTE BAG)	(STRAWBERRY CHARLOTTE BAG, RED RETROSPOT CHARL...	0.041905	0.012364	0.010286	0.245466	19.853534	0.009768	1.308934	0.991166
3340	(STRAWBERRY CHARLOTTE BAG)	(WOODLAND CHARLOTTE BAG, RED RETROSPOT CHARLOT...	0.036281	0.012668	0.010286	0.283520	22.381034	0.009827	1.378031	0.991284
3341	(CHARLOTTE BAG SUKI DESIGN)	(WOODLAND CHARLOTTE BAG, RED RETROSPOT CHARLOT...	0.044337	0.012212	0.010286	0.232000	18.998008	0.009745	1.286183	0.991315

3342 rows × 10 columns

STEP 4 | (cont)..

4.2 Top 10 Association Rules by Lift:

To identify the strongest associations between products, we examined the top 10 association rules sorted by lift. Lift indicates the strength of the association, where a higher lift value suggests a stronger correlation between items.

<u>Results:</u>

The top rules showed exceptionally high lift values, particularly among combinations of HERB MARKER items, indicating that these items are frequently purchased together.

Key rules include:

~~Rule 1:~~ Customers who purchased "HERB MARKER PARSLEY" and "HERB MARKER ROSEMARY" were also likely to buy "HERB MARKER THYME". This rule has a lift of 80.67 and a confidence of 95.2%, making it a strong association.

~~Rule 2:~~ Purchases of "HERB MARKER THYME" are associated with "HERB MARKER PARSLEY" and "HERB MARKER ROSEMARY" with a lift of 80.67 and a confidence of 85.8%.

Other Notable Rules: Many other herb markers are commonly bought together, with similar lift values above 76. These high values suggest that customers are highly likely to purchase multiple herb markers in one transaction.

Top 10 Rules by Lift:

	antecedents	consequents	support	confidence	lift
409	(HERB MARKER PARSLEY, HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.010134	0.952381	80.666258
414	(HERB MARKER THYME)	(HERB MARKER PARSLEY, HERB MARKER ROSEMARY)	0.010134	0.858369	80.666258
410	(HERB MARKER PARSLEY, HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.010134	0.952381	80.321530
413	(HERB MARKER ROSEMARY)	(HERB MARKER PARSLEY, HERB MARKER THYME)	0.010134	0.854701	80.321530
121	(HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.010996	0.927350	78.546183
122	(HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.010996	0.931330	78.546183
411	(HERB MARKER ROSEMARY, HERB MARKER THYME)	(HERB MARKER PARSLEY)	0.010134	0.921659	78.400604
412	(HERB MARKER PARSLEY)	(HERB MARKER ROSEMARY, HERB MARKER THYME)	0.010134	0.862069	78.400604
120	(HERB MARKER THYME)	(HERB MARKER PARSLEY)	0.010641	0.901288	76.667715

STEP 4 | (cont)..

4.3 Top 10 Association Rules by Confidence:
The top 10 rules with the highest confidence levels demonstrated significant patterns where certain products were highly predictive of subsequent purchases.

<u>Key Findings:</u>

- The highest confidence was observed with the purchase of ***"BEADED CRYSTAL HEART PINK ON STICK"***, which strongly predicted the purchase of "DOTCOM POSTAGE" with a 97.57% likelihood. This suggests that customers buying decorative items are likely to add postage-related products to their carts.

- ***"JAM MAKING SET PRINTED" and "SUKI SHOULDER BAG"*** were also highly associated with "DOTCOM POSTAGE", showing a 95.87% confidence level. This pattern indicates the popularity of specific combinations where gift items and bags often require additional packaging services.

- The ***"HERB MARKER"*** items (such as ***"HERB MARKER PARSLEY"***, "HERB MARKER ROSEMARY", and "HERB MARKER THYME") showed a strong associative relationship, with confidence levels exceeding 95%. This association reveals that customers are likely to buy various herb markers together.

- Another high-confidence association was seen among holiday-themed products such as ***"WOODEN TREE CHRISTMAS SCANDINAVIAN"*** and "WOODEN HEART CHRISTMAS SCANDINAVIAN", which were highly predictive of the purchase of ***"WOODEN STAR CHRISTMAS SCANDINAVIAN"*** with a 93.33% confidence level.

Top 10 Rules by Confidence:

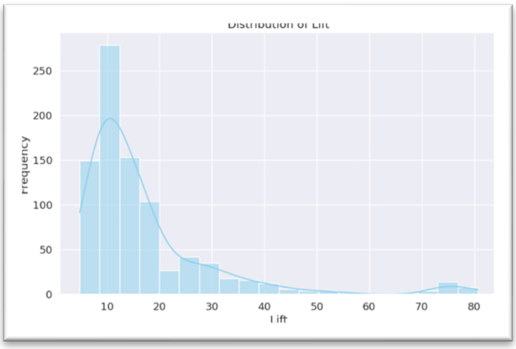
	antecedents	consequents	support	confidence	lift
17	(BEADED CRYSTAL HEART PINK ON STICK)	(DOTCOM POSTAGE)	0.010185	0.975728	27.197733
339	(JAM MAKING SET PRINTED, SUKI SHOULDER BAG)	(DOTCOM POSTAGE)	0.010590	0.958716	26.723520
409	(HERB MARKER PARSLEY, HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.010134	0.952381	80.666258
410	(HERB MARKER PARSLEY, HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.010134	0.952381	80.321530
735	(REGENCY TEA PLATE ROSES, REGENCY TEA PLATE PINK)	(REGENCY TEA PLATE GREEN)	0.011958	0.947791	52.689179
747	(WOODEN TREE CHRISTMAS SCANDINAVIAN, WOODEN HE...	(WOODEN STAR CHRISTMAS SCANDINAVIAN)	0.011350	0.933333	37.210774
122	(HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.010996	0.931330	78.546183
121	(HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.010996	0.927350	78.546183
411	(HERB MARKER ROSEMARY, HERB MARKER THYME)	(HERB MARKER PARSLEY)	0.010134	0.921659	78.400604
775	(WOODLAND CHARLOTTE BAG, STRAWBERRY CHARLOTTE ...	(RED RETROSPOT CHARLOTTE BAG)	0.012212	0.919847	17.658742

STEP 5 | Distribution of Metric in Association Rules

5.1 graph shows the distribution of Lift:
X-axis (Lift): Represents the lift values. Lift measures the strength of the association between items in a rule. A higher lift value indicates a stronger association between the items.

Y-axis (Frequency): Represents the number of rules that have certain lift values. This shows how frequently different lift values appear in the dataset.

The majority of the association rules have lift values between 8 and 20, indicating that most of the rules show moderately strong associations between items.

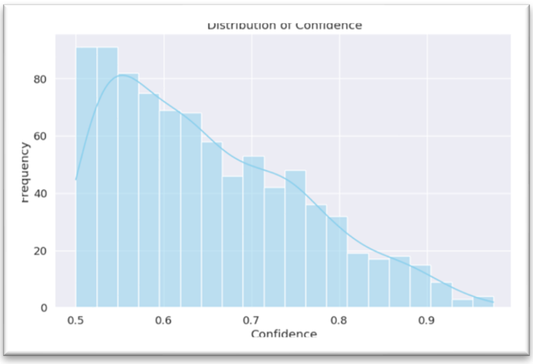


A small number of rules have very high lift values exceeding 60, although these are rare.

5.2 graph shows the distribution Confidence:
x-axis represents the confidence values, ranging from 0.5 to 1.0 (50% to 100%). Confidence closer to 1 indicates stronger relationships between the antecedent and consequent.

Y-axis represents the frequency of rules within each confidence range, showing how many rules fall into specific confidence intervals.

Most rules have confidence values between 0.5 and 0.7, indicating that the majority of the relationships are moderately strong.



Fewer rules have very high confidence (closer to 1), indicating that highly confident rules are less common in this dataset.

STEP 5 | (cont)..

5.3 Support vs Confidence with Lift as Size:
This scatter plot visualizes the relationship between support, confidence, and lift for the generated association rules.

~~**Support (x-axis):**~~ The proportion of transactions in the dataset that contain the itemset. Higher support indicates more frequent itemsets.

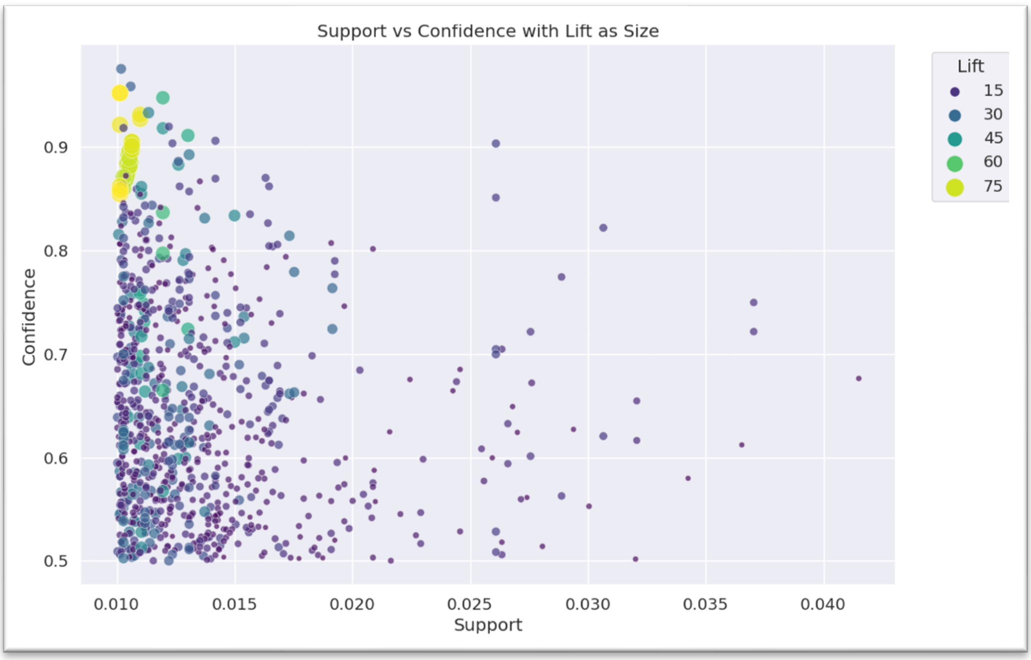
~~**Confidence (y-axis):**~~ The likelihood that the consequent occurs given the antecedent. Higher confidence suggests stronger relationships.

~~**Lift (circle size and color):**~~ Indicates how much more likely the consequent is to occur compared to its baseline frequency. Higher lift suggests a stronger association between items.

Most rules have lower support (below 0.015) but vary widely in confidence and lift.

High lift values (in yellow) are often concentrated at lower support levels, but these rules tend to have high confidence.

The size and color variation in the chart help highlight which rules are more influential (high lift) compared to others.



STEP 5 | (cont)..

5.4 visualizations:

- Support vs Confidence Scatter Plot:

Support is plotted on the x-axis, and Confidence is on the y-axis.

The plot shows how frequent itemsets (those with higher support) tend to have varying levels of confidence.

Higher confidence values indicate that given the antecedent, the consequent is more likely to occur.

There is a high concentration of rules with support between 0.01 and 0.02 and confidence values ranging from 0.6 to 0.9.

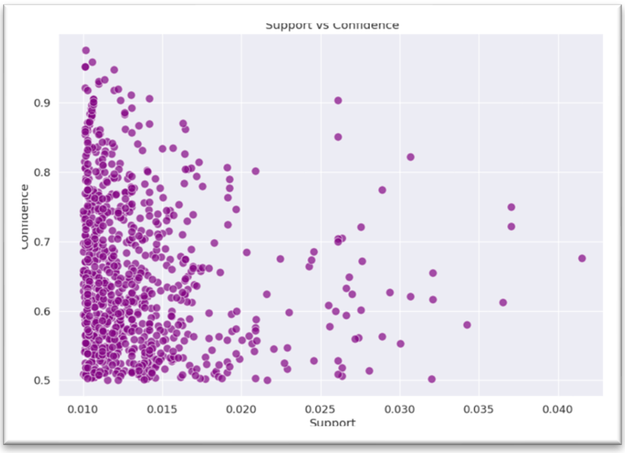
- Lift vs Confidence Scatter Plot:

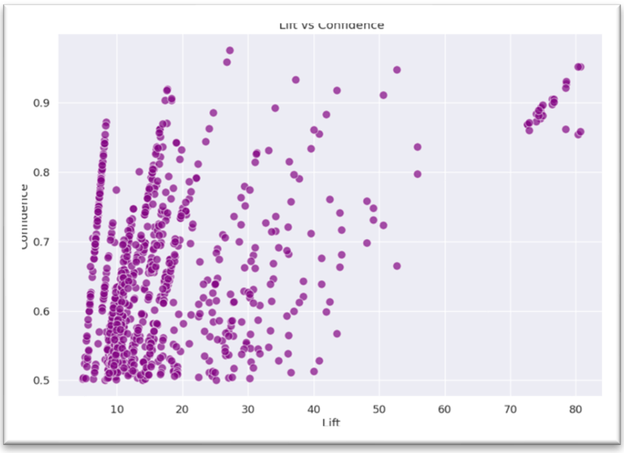
Lift is on the x-axis, and Confidence is on the y-axis

The plot highlights that rules with higher Lift values tend to indicate stronger associations between itemsets.

Rules with higher Lift (values greater than 60) have confidence values over 0.9, suggesting strong and reliable associations between items.

There is also a notable spread of rules with lift values between 10 to 40, indicating a range of potential strengths in association.





CONCLUSION

This project involved performing a market basket analysis using retail data. Starting with data cleaning, several important steps were taken to ensure the dataset was ready for analysis. We addressed missing values, removed duplicates, and ensured the integrity of key columns such as BillNo, Itemname, Quantity, and Date. Additionally, we converted numerical values like Price into a standardized format and created a SumPrice column to compute total prices per transaction.

After cleaning the data, we converted the dataset into a transactional format, preparing it for frequent itemset mining. Using a one-hot encoding technique, each product was transformed into binary features, enabling us to apply the Apriori algorithm for mining frequent itemsets. Following that, we generated association rules to discover meaningful relationships between items frequently purchased together.

The visualizations generated from this analysis, including histograms for lift and confidence, and scatter plots comparing support, confidence, and lift, provided clear insights into the strength of the associations found. We identified the most frequently purchased items in each country and extracted valuable patterns, such as popular product pairs. These results can help businesses optimize product placements and create targeted marketing strategies based on customer behavior.

The results revealed strong associations between certain products based on support, confidence, and lift. These associations provide valuable insights into purchasing behaviors that businesses can leverage for competitive advantage. By optimizing product display strategies, tailoring campaigns based on geographic regions and popular items, and capitalizing on highly demanded products, companies can significantly increase revenue and enhance customer satisfaction.

RECOMMENDATIONS

Based on the analysis results, several suggestions can be made to enhance performance and derive more business value:

~~*1 - Improve Display and Promotion Strategies:*~~ Companies can use the discovered association rules to display products that are frequently purchased together in nearby locations. For instance, products with strong associations in the data can be positioned together in stores to attract more customers.

~~*2 - Geographically Tailored Marketing:*~~ By identifying the best-selling products in each country, marketing strategies can be tailored according to local customer behavior. This allows companies to increase sales by promoting the most demanded items in each region.

~~*3- Leverage High-Support Products:*~~ Products with high support represent items that are frequently purchased. This information can be used to expand the stock of highly demanded products, ensuring availability, while also offering promotions or discounts to encourage more purchases.

~~*4 - Targeted Promotional Campaigns*~~*:* The relationships discovered between products can be utilized to develop targeted marketing campaigns. For example, if customers who purchase one product frequently buy another product along with it, "bundled products" can be promoted as part of marketing efforts to increase cross-sales.