MARKET BASKET INSIGHTS

AI PHASE 5

Problem Statement:

The problem at hand involves performing Market Basket Analysis (MBA) on a retail dataset to discover meaningful patterns in customer purchasing behavior. This analysis aims to identify associations between products purchased by customers, enabling the business to make strategic decisions for cross-selling and upselling opportunities.

Design Thinking Process:

1. Empathize:
   - Understand the dataset and its attributes.
   - Identify the business objective: Enhance sales through cross-selling and upselling strategies.
2. Define:
   - Define the problem statement: Identify product associations for effective cross-selling and upselling.
   - Set goals: Discover high-confidence association rules to guide marketing strategies.
3. Ideate:
   - Plan data preprocessing steps to clean and prepare the dataset.
   - Apply association analysis techniques (Apriori algorithm) to discover patterns.
   - Interpret the discovered association rules for actionable insights.
4. Prototype:
   - Implement data preprocessing techniques.
   - Generate association rules using Apriori algorithm.
5. Test:
   - Evaluate the association rules for confidence, support, and lift metrics.
   - Validate the rules against business knowledge to ensure relevance.
6. Implement:
   - Present the discovered association rules to stakeholders.
   - Develop actionable strategies based on the identified patterns.

Phases of Development:

1. Data Preprocessing:
   - Loaded the dataset and handled data type issues.
   - Cleaned missing and inconsistent data, addressing zero prices and negative quantities.
   - Removed irrelevant columns like 'Country' and 'Adjust bad debt'.
   - Formatted date columns and calculated total prices.
2. Association Analysis Techniques:
   - Utilized the Apriori algorithm to generate frequent itemsets and association rules.
   - Filtered and sorted rules based on confidence, support, and lift metrics.

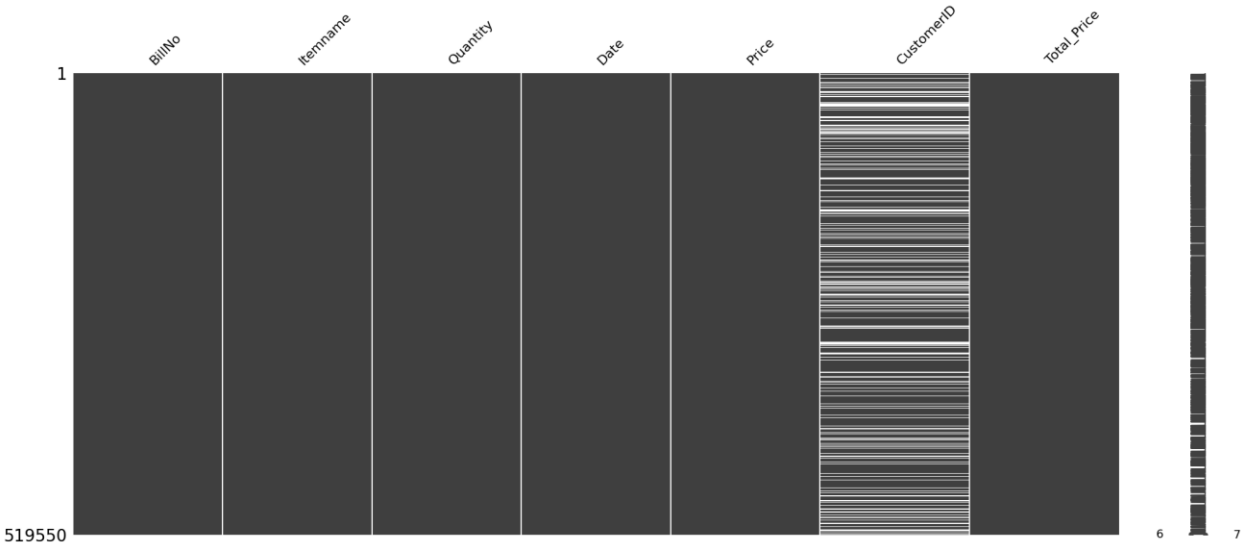
Dataset Description:

The dataset contains information about retail transactions with attributes like 'BillNo', 'Itemname', 'Quantity', 'Price', and 'Date'. It comprises 522,065 rows and 7 attributes.

	Quantity	Date	Price	CustomerID	Total_Price
count	519550.000000	519550	519550.000000	387985.000000	519550.000000
mean	10.398379	2011-07-04 16:03:24.573958656	3.866610	15317.042994	19.807204
min	1.000000	2010-12-01 08:26:00	0.001000	12346.000000	0.001000
25%	1.000000	2011-03-28 10:52:00	1.250000	13950.000000	3.750000
50%	3.000000	2011-07-20 11:55:00	2.080000	15265.000000	9.870000
75%	10.000000	2011-10-19 15:08:00	4.130000	16837.000000	17.400000
max	80995.000000	2011-12-09 12:50:00	13541.330000	18287.000000	168469.600000
std	157.005103	NaN	32.516631	1721.813298	272.430068

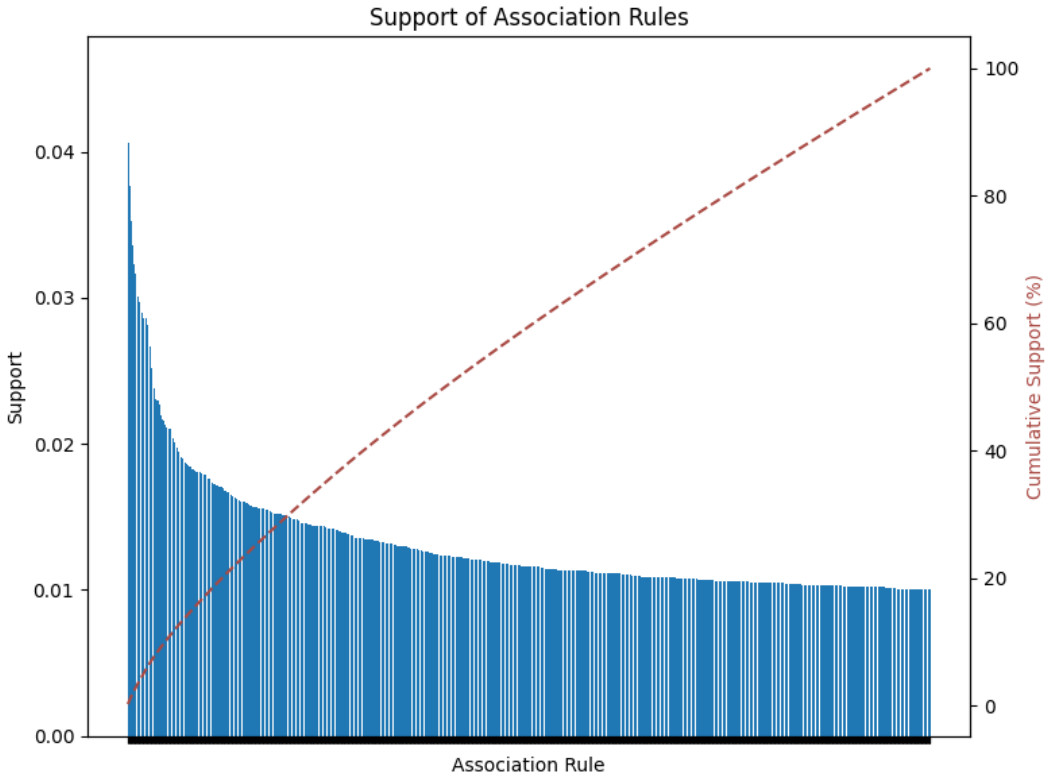
Data Preprocessing Steps:

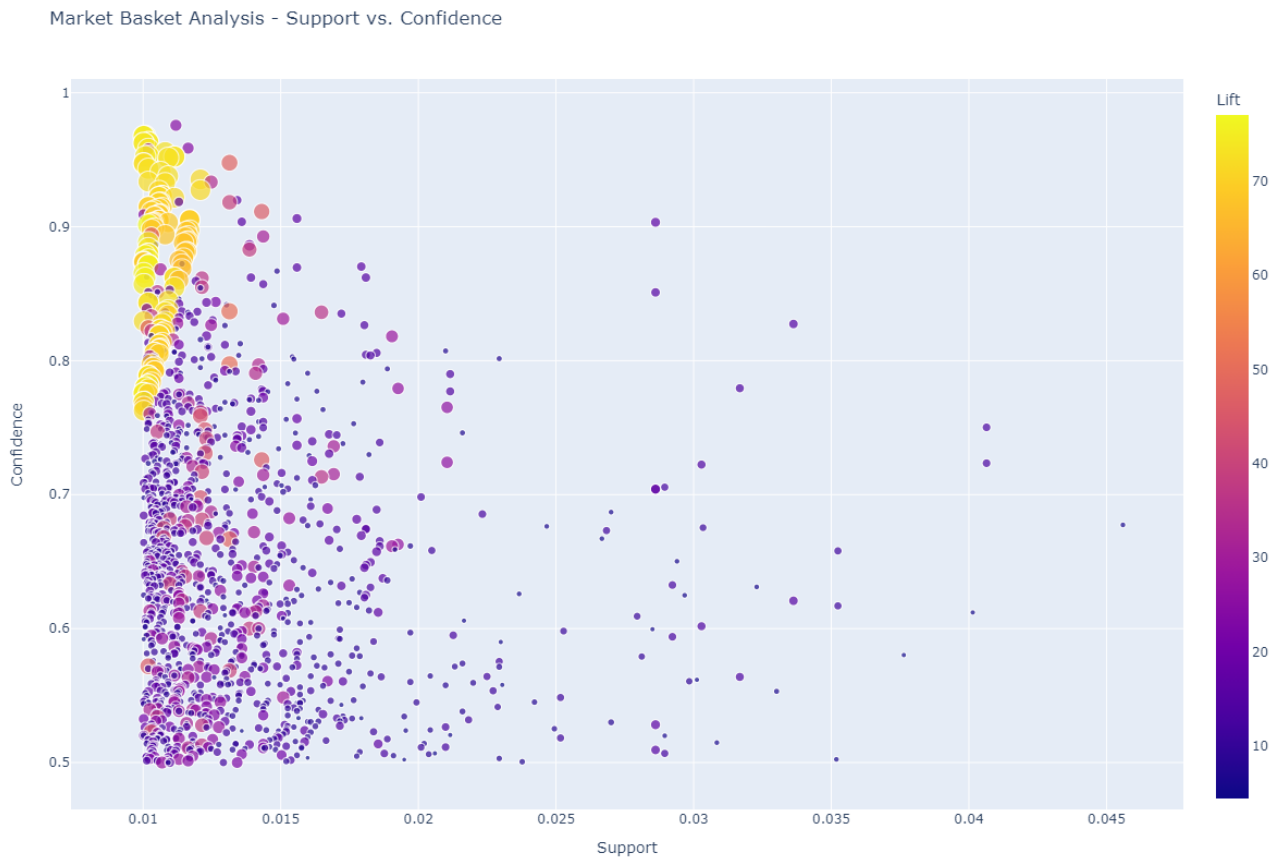
1. Cleaning Data:
   - Handled DtypeWarning and converted 'Price' column to float64.
   - Removed rows with zero prices and negative quantities.
   - Removed irrelevant columns like 'Country' and 'Adjust bad debt'.
2. Handling Missing Data:
   - Checked for missing values using visualization tools like missingno.
   - Removed rows with missing 'Itemname' values.
3. Formatting Data:
   - Formatted date columns for analysis.



Association Analysis Techniques:

1. <u>Generating Association Rules:</u>
   - Filtered items and bills based on occurrences to focus on frequent transactions.
   - Created a pivot table for Apriori algorithm.
   - Generated frequent itemsets and association rules using the Apriori algorithm.





Discovered Association Rules and Business Implications:

Cross-Selling Recommendations:

- Customers who bought ‘Product A’ also bought ‘Product B’, ‘Product C’, etc.
- These associations enable targeted cross-selling promotions, enhancing the likelihood of multiple product purchases.

Cross-Selling Output:

- Customers who bought 'BEADED CRYSTAL HEART PINK ON STICK' also bought 'DOTCOM POSTAGE'.
- Customers who bought 'HERB MARKER THYME' also bought 'HERB MARKER ROSEMARY'.
- Customers who bought 'HERB MARKER ROSEMARY' also bought 'HERB MARKER THYME'.
- Customers who bought 'HERB MARKER CHIVES' also bought 'HERB MARKER PARSLEY'.
- Customers who bought 'REGENCY TEA PLATE PINK' also bought 'REGENCY TEA PLATE GREEN'.

Upselling Recommendations:

- For customers who bought ‘Product X’, recommend upgrades to ‘Product Y’, ‘Product Z’, etc.
- Upselling opportunities help increase revenue by encouraging customers to choose higher-priced or premium products.

Upselling Output:

- For customers who bought 'HERB MARKER CHIVES', recommend the following upgrades: HERB MARKER THYME, HERB MARKER PARSLEY.
- For customers who bought 'HERB MARKER CHIVES', recommend the following upgrades: HERB MARKER PARSLEY, HERB MARKER MINT.
- For customers who bought 'HERB MARKER CHIVES', recommend the following upgrades: HERB MARKER PARSLEY, HERB MARKER ROSEMARY.
- For customers who bought 'HERB MARKER CHIVES', recommend the following upgrades: HERB MARKER THYME, HERB MARKER ROSEMARY.
- For customers who bought 'HERB MARKER THYME', recommend the following upgrades: HERB MARKER PARSLEY, HERB MARKER ROSEMARY.

Insights:

- For customers who bought 'HERB MARKER CHIVES', recommend the following upgrades: HERB MARKER THYME, HERB MARKER PARSLEY.
- For customers who bought 'HERB MARKER THYME', recommend the following upgrades: HERB MARKER PARSLEY, HERB MARKER ROSEMARY.
- For customers who bought 'HERB MARKER PARSLEY', recommend the following upgrades: HERB MARKER THYME, HERB MARKER ROSEMARY.
- For customers who bought 'HERB MARKER ROSEMARY', recommend the following upgrades: HERB MARKER THYME, HERB MARKER PARSLEY.
- For customers who bought 'REGENCY TEA PLATE PINK', recommend the following upgrades: REGENCY TEA PLATE ROSES, REGENCY TEA PLATE GREEN.

Conclusion:

The analysis provides actionable insights for cross-selling and upselling opportunities. By understanding customer purchase patterns, the business can optimize marketing strategies, personalize customer experiences, and ultimately boost sales and customer satisfaction. These insights demonstrate the power of data-driven decision-making in retail operations.