# Market Basket Insights

IBM Naan Mudhalvan Phase 2 Project Submission

# Contents

# Introduction

Market Basket Analysis (MBA) is a data mining technique used by retailers to understand customer purchasing patterns. It helps identify associations between products, allowing businesses to make data-driven decisions, optimize product placement, and design targeted marketing strategies.

## Objective:

To perform Market Basket Analysis on a retail dataset, extract meaningful associations between products, and provide actionable insights for business optimization.

Dataset Link: https://www.kaggle.com/datasets/aslanahmedov/market-basket-analysis

# Project Steps

Data Preprocessing:

- Removed unnecessary columns.
- Handled missing values.
- Cleaned and formatted data, dealing with special characters and unwanted symbols.
- Converted data types for compatibility with analysis techniques.

```
[12]: df1 = df.dropna()

[14]: df1
```

[14]:

| | BillNo | Itemname | Quantity | Date | Price | CustomerID | Country |
|---|---|---|---|---|---|---|---|
| 0 | 536365 | WHITE HANGING HEART T-LIGHT HOLDER | 6 | 01.12.2010 08:26 | 2.55 | 17850.0 | United Kingdom |
| 1 | 536365 | WHITE METAL LANTERN | 6 | 01.12.2010 08:26 | 3.39 | 17850.0 | United Kingdom |
| 2 | 536365 | CREAM CUPID HEARTS COAT HANGER | 8 | 01.12.2010 08:26 | 2.75 | 17850.0 | United Kingdom |
| 3 | 536365 | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | 01.12.2010 08:26 | 3.39 | 17850.0 | United Kingdom |
| 4 | 536365 | RED WOOLLY HOTTIE WHITE HEART. | 6 | 01.12.2010 08:26 | 3.39 | 17850.0 | United Kingdom |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 522058 | 581587 | PACK OF 20 SPACEBOY NAPKINS | 12 | 09.12.2011 12:50 | 0.85 | 12680.0 | France |
| 522059 | 581587 | CHILDREN'S APRON DOLLY GIRL | 6 | 09.12.2011 12:50 | 2.10 | 12680.0 | France |
| 522060 | 581587 | CHILDRENS CUTLERY DOLLY GIRL | 4 | 09.12.2011 12:50 | 4.15 | 12680.0 | France |
| 522061 | 581587 | CHILDRENS CUTLERY CIRCUS PARADE | 4 | 09.12.2011 12:50 | 4.15 | 12680.0 | France |
| 522062 | 581587 | BAKING SET 9 PIECE RETROSPOT | 3 | 09.12.2011 12:50 | 4.95 | 12680.0 | France |

387490 rows × 7 columns

→ Removed the rows where columns contained null values

## Market Basket Analysis (MBA) Preparation:

•Prepared transaction data by grouping items based on common attributes (e.g., same bill number).

•Encoded transaction data into a format suitable for MBA algorithms.

•Used FP-Growth for its more memory efficient approach on large datasets.

```
[17]: transaction_data = transaction_data.drop(['BillNo', 'Date'], axis=1)

[18]: transaction_data

[18]:
```

| | Itemname |
|---|---|
| 0 | WHITE HANGING HEART T-LIGHT HOLDER,WHITE METAL... |
| 1 | HAND WARMER UNION JACK,HAND WARMER RED POLKA DOT |
| 2 | ASSORTED COLOUR BIRD ORNAMENT,POPPY'S PLAYHOUS... |
| 3 | JAM MAKING SET WITH JARS,RED COAT RACK PARIS F... |
| 4 | BATH BUILDING BLOCK WORD |
| ... | ... |
| 18184 | LUNCH BAG RED RETROSPOT,6 CHOCOLATE LOVE HEART... |
| 18185 | RED FLOCK LOVE HEART PHOTO FRAME,6 CHOCOLATE L... |
| 18186 | BLACK TEA TOWEL CLASSIC DESIGN,ASSORTED BOTTLE... |
| 18187 | LARGE CAKE STAND HANGING STRAWBERY,SET OF 3 H... |
| 18188 | CIRCUS PARADE LUNCH BOX,PLASTERS IN TIN CIRCUS... |

18189 rows × 1 columns

Grouped items with similar bill no and date of transaction for applying association rules efficiently.

## Association Rule Mining:

- Applied Apriori or FP-growth algorithms to find frequent itemsets.
- Generated association rules based on support, confidence, and other metrics.
- Explored and analyzed discovered association rules.
- Selected the top 10 most associated itemsets with a confidence level of 1.

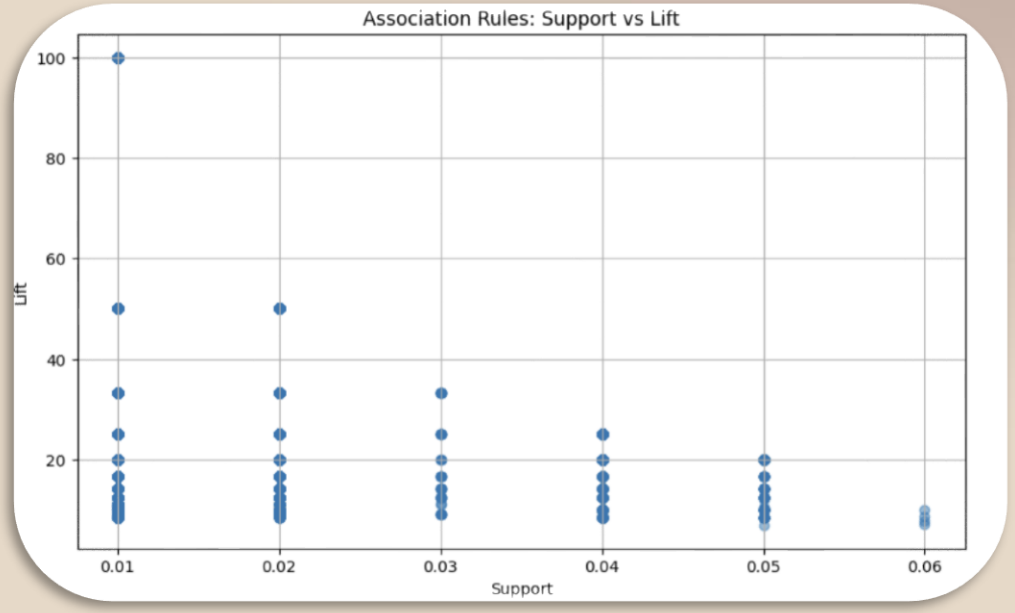
## Sample Output:

```
Top Association Rules (LHS, RHS, Support, Confidence):
{' RETRO SPOT'} => {''} - Support: 0.010, Confidence: 1.000
{'RED RETROSPOT MINI CASES', 'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'RED RETROSPOT TAPE'} - Support: 0.010, Confidence: 1.000
{'ROSE COTTAGE KEEPSAKE BOX', 'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'RED RETROSPOT MINI CASES'} - Support: 0.010, Confidence: 1.000
{'ROSE COTTAGE KEEPSAKE BOX', 'RED RETROSPOT MINI CASES'} => {'FELTCRAFT PRINCESS OLIVIA DOLL'} - Support: 0.010, Confidence: 1.000
{'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'RED RETROSPOT MINI CASES', 'REX CASH+CARRY JUMBO SHOPPER'} - Support: 0.010, Confidence: 1.000
{'RED RETROSPOT MINI CASES', 'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'REX CASH+CARRY JUMBO SHOPPER'} - Support: 0.010, Confidence: 1.000
{'FELTCRAFT PRINCESS OLIVIA DOLL', 'REX CASH+CARRY JUMBO SHOPPER'} => {'RED RETROSPOT MINI CASES'} - Support: 0.010, Confidence: 1.000
{'RED RETROSPOT MINI CASES', 'REX CASH+CARRY JUMBO SHOPPER'} => {'FELTCRAFT PRINCESS OLIVIA DOLL'} - Support: 0.010, Confidence: 1.000
{'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'RED TOADSTOOL LED NIGHT LIGHT', 'RED RETROSPOT MINI CASES'} - Support: 0.010, Confidence: 1.000
{'RED RETROSPOT MINI CASES', 'FELTCRAFT PRINCESS OLIVIA DOLL'} => {'RED TOADSTOOL LED NIGHT LIGHT'} - Support: 0.010, Confidence: 1.000
```

# Visualization

•Visualized the association rules using scatter plots

## Sample Output:



# Innovative Ideas

**Real-time Analysis:**
Implement a system for real-time market basket analysis to provide instant recommendations to customers during online shopping.

**Customer Segmentation:**
Apply clustering techniques to segment customers based on their purchasing behavior. Perform separate market basket analysis for each segment.

**Dynamic Pricing Strategies:**
Utilize MBA insights to implement dynamic pricing strategies, adjusting prices based on product associations and customer preferences.

**Predictive Analytics:**
Use historical transaction data to build predictive models for future purchases, enabling proactive inventory management and product recommendations.

# Conclusion

- **Complementary Purchases:** Customers tend to buy complementary items together, suggesting potential for cross-selling and upselling strategies.
- **Seasonal Influences:** Purchasing behaviors are strongly influenced by seasonal trends and promotions, allowing businesses to tailor marketing efforts accordingly.
- **Targeted Marketing:** Identifying items frequently bought together enables targeted marketing, enhancing the effectiveness of promotional campaigns.
- **Optimized Product Placement:** Understanding customer preferences aids in optimizing product placement, improving visibility and sales of specific items.
- **Customer Satisfaction:** Meeting customer needs by offering complementary items enhances satisfaction, fostering customer loyalty.
- **Business Efficiency:** Utilizing market basket insights streamlines inventory management and sales strategies, leading to efficient business operations.