



# Market Basket Analysis Report

COURSE PRESENTER:

(Dr. Omaila Fallatah)

SUBMITTED BY:

Name	Id
امل عوض العتيبي	444001258
ريوف فيصل المحنوني	444003028
رهف ياسين برناوي	444006091

## Contents

<b>Introduction .....</b>	<b>3</b>
<b>Objective.....</b>	<b>3</b>
<b>Data Overview .....</b>	<b>3</b>
<b>Preparing the data.....</b>	<b>4</b>
<b>Missing Values .....</b>	<b>4</b>
<b>Invoice Filtering.....</b>	<b>4</b>
<b>Exploratory Data Analysis .....</b>	<b>4</b>
<b>Top Countries: .....</b>	<b>4</b>
<b>apriori algorithm.....</b>	<b>4</b>
<b>association rules.....</b>	<b>5</b>
<b>Conclusion .....</b>	<b>6</b>

## Introduction

Market Basket Analysis is a method used to discover links between items in retail purchases. This analysis relies on the Apriori algorithm to identify frequently purchased items and create association rules. It helps retailers understand customer buying behavior, improve store layouts, and develop targeted marketing strategies. By relying on data, retailers can make decisions that increase sales and enhance customer experience.

## Objective

- 1.Understanding Customer Behavior
- 2.Improving Store Layout
- 3.Developing Effective Marketing Strategies
- 4.Increasing Sales
- 5.Enhancing Customer Experience

## Data Overview

The dataset consists of retail transactions with the following key columns

Invoice No: Unique identifier for each transaction.

Stock Code: Unique identifier for each product.

Description: Description of the product.

Quantity: Number of items purchased.

Invoice Date: Date and time of the transaction.

Unit Price: Price per item.

Customer: Unique identifier for each customer.

Country: Country of the customer.

## Preparing the data

### Missing Values:

The dataset initially contained missing values. All rows with null values were dropped, resulting in 406,829 entries.

### Invoice Filtering:

Transactions with invoices containing a "C" (indicating cancellations) were removed, resulting in 397,924 entries.

## Exploratory Data Analysis

### Top Countries:

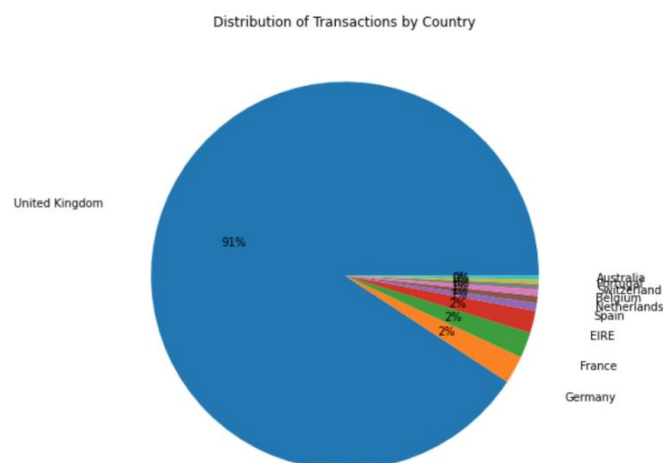
The analysis revealed the distribution of transactions across countries, with the top ten countries being:

United Kingdom: 354,345 transactions

Germany: 9,042 transactions

France: 8,342 transactions

Others include EIRE, Spain, Netherlands, Belgium, Switzerland, Portugal, and Australia



## Apriori algorithm

we applied the apriori algorithm to find the most frequently purchased items in the dataset. From the results, we can see that the "White hanging Heart T-Light Holder" is the most frequently purchased item with a support value of 0.121358, meaning it was purchased in 12% of all transactions.

## association rules

1.Rule: If a customer purchases ROSES REGENCY TEACUP AND SAUCER, they are likely to also purchase GREEN REGENCY TEACUP AND SAUCER.

Support: 3.10%	Confidence: 70.52%	Lift: 17.72
----------------	--------------------	-------------

Application: This indicates a strong association between these two items, suggesting that retailers could offer them as a bundled gift set or promotional deal to increase sales.

2.Rule: If a customer purchases GREEN REGENCY TEACUP AND SAUCER, they are likely to also purchase ROSES REGENCY TEACUP AND SAUCER.

Support: 3.10%	Confidence: 77.78%	Lift: 17.72
----------------	--------------------	-------------

Application: Similar to the first rule, retailers can promote these items together to encourage cross-selling, particularly in tea or home goods sections.

3.Rule: If a customer purchases LUNCH BAG RED RETROSPOT, they are likely to also purchase LUNCH BAG PINK POLKADOT.

Support: 3.06%	Confidence: 42.05%	Lift: 7.63
----------------	--------------------	------------

Application: Retailers can create special offers where purchasing one lunch bag leads to a discount on the other, appealing to parents and schoolchildren.

4.Rule: If a customer purchases LUNCH BAG PINK POLKADOT, they are likely to also purchase LUNCH BAG RED RETROSPOT.

Support: 3.06%	Confidence: 55.61%	Lift: 7.63
----------------	--------------------	------------

Application: Similar to the previous rule, suggesting a promotional strategy focusing on these complementary products for targeted marketing.

5.Rule: If a customer purchases JUMBO BAG RED RETROSPOT, they are likely to also purchase JUMBO BAG PINK POLKADOT.

Support: 3.29%	Confidence: 35.31%	Lift:6.70
----------------	--------------------	-----------

Application: This finding suggests that these bags can be promoted together in fashion or seasonal campaigns, enhancing the shopping experience.

## Conclusion

Based on the results from implementing association rules, we can see that “Roses Regency Teacup and Saucer” and “Green Regency Teacup and Saucer” have the highest “lift” value, and therefore the highest association of any two products. With a combined support of 0.0309, it means both items were purchased together in 3.09% of all transactions.