**PHASE – III**

**DEVELOPMENT PART 1**

|  |  |
| --- | --- |
| DATE | 25-10-2023 |
| TEAM ID / TEAM NAME | Proj\_224020\_Team\_1 |
| PROJECT NAME | Market Basket Insights |
| STUDENT NAME WITH ID | **V Dinesh**  K Gowtham  R Gunaseelan  K.S Illayabharathi |

**Market Basket Insights**

**Introduction:**

Briefly introduce the concept of market basket analysis and its importance in retail.

Explain the objective of the document - to analyze customer purchasing behaviour and identify cross-selling opportunities.

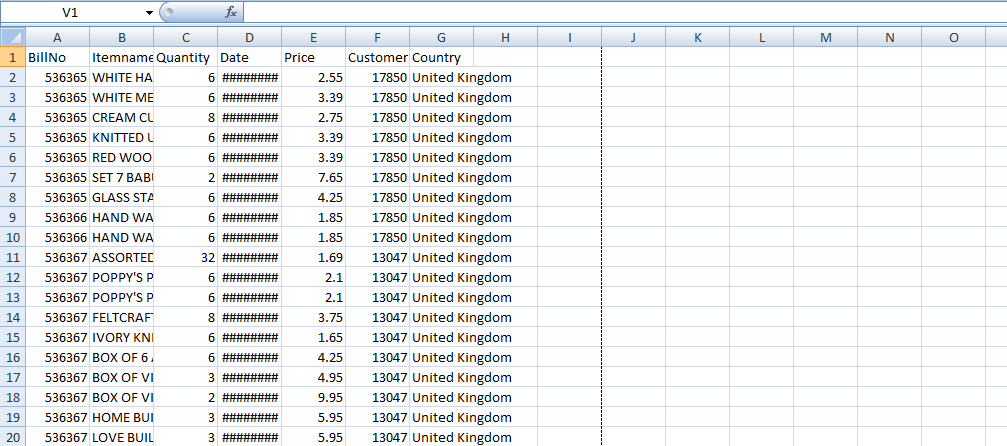
Provide an overview of the Apriori algorithm as the chosen technique for this analysis.

**Problem Statement:**

Define the problem statement clearly: "To understand customer purchasing behavior and identify potential cross-selling opportunities for a retail business through market basket analysis."

Discuss the challenges and potential benefits of solving this problem.

**DATASET:**

****

**LOADING THE DATASET****:**

import pandas as pd

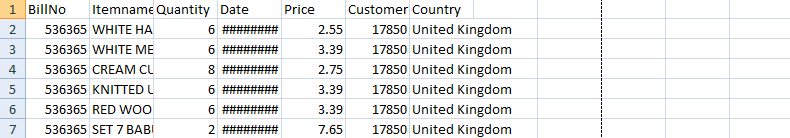
# loading the dataset into a dataframe

df = pd.read\_excel(‘C:\Users\jesus\Desktop\IBM\dataset.xlsx’)

**PREPROCESSING THE DATASET:**

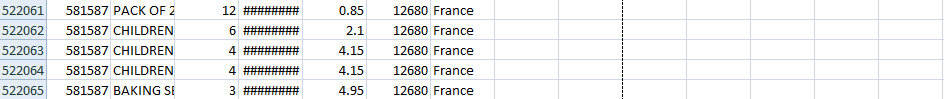
# Display the first 6 rows of data

df.head()



# Display the last 5 rows of data

df.tail()



# Display the shape of the data (number of rows and columns)

df.shape

( 522065,**7**)

# Display the column names

df.columns

Index([‘BillNo’, ‘Itemname’, ‘Quantity’, ‘Date’, ‘Price’, ‘CustomerID’, ‘Country’], datatype = ‘object’)

# Display the data types of each column

df.dtypes

BillNo float64

Itemname object

Quantity float64

Date object

Price float64

CustomerID float64

Country object

**Conclusion:**

Summarize the key findings and insights from the analysis.

Discuss the potential impact on the retail business.

Reflect on the significance of market basket analysis in understanding customer behaviour and improving business operations.