# 

# PROJECT DOCUMENTATION

# MARKET BASKET INSIGHTS.

Introduction:

Market basket analysis is a data mining technique that identifies relationships between products frequently purchased together.In today's competitive retail landscape, understanding customer purchasing behavior is essential for increasing sales and improving customer satisfaction. This project focuses on utilizing Market Basket Analysis (MBA), a data-driven technique, to uncover hidden patterns in customer transactions and provide valuable insights to optimize retail operations.

About dataset:

The dataset in the code is a collection of CSV files containing news articles. Each CSV file contains a header row with the following columns:

**title:** Title of Market basket insights

**text:** Body of market basket insights

**subject:** The subject of market basket insights transactions ( like date, product identity number, reference number, transaction I'd, transaction number)

**date:** Regarding to the particular transaction.

The dataset is likely to be collected from a variety of sources, such as online news websites, social media platforms, and other public repositories. The data may be preprocessed to remove cancelled transaction, unsuccessful transaction and overlapping transactions.

The dataset can be used to train a machine learning model to perform market basket analysis for customer transactions and provide valuable insights. The model can then be used to perform market basket analysis.

Preprocessing steps:

The following preprocessing steps are performed in the code:

* Read the CSV files into DataFrames.
* Convert the dataframes into transaction data using Association rule mining.
* Combine all the transaction data like billno, date, product code into one row.
* Load the transaction data into object of the transaction class using R Function.

Choice of ML algorithm:

Apriori is a classic algorithm that helps identify associations and patterns in transactional data, such as retail sales. It works by finding frequent itemsets and generating association rules that describe the relationships between items in a basket.

Model training:

The model is trained using the following steps:

* Gather transactional data, which includes records of items purchased in different transactions.
* Preprocess the data, removing duplicates and organizing it into a suitable format.
* Frequent itemsets are combinations of items that occur together frequently in the transactions.
* Apply these insights for various purposes, such as optimizing product placement, cross-selling, or targeted marketing.
* Integrate the learned rules into your business processes or systems to make data-driven decisions and recommendations.

Evaluation metrics:

The following evaluation metrics are used to evaluate the model's performance:

Support:Support measures the proportion of transactions that contain a specific itemset. It indicates how frequently an itemset appears in the dataset. Higher support values suggest more significant itemset occurrences

Conviction:Conviction measures the likelihood that an item appears in a transaction without the presence of another item. It helps identify the dependency between items. Higher conviction values suggest stronger associations.

Lift-to-Confidence Ratio:This metric can help identify interesting and actionable rules. It's the ratio of lift to confidence. A high lift-to-confidence ratio indicates that the association rule has a strong impact.

Innovative technique used:

Market basket insights projects aim to uncover patterns and associations in transactional data to drive business decisions. Innovation techniques used in these projects can include Advanced Data Mining Algorithms,Machine Learning,Deep Learning,Real-time Analytics,Graph Databases,Natural Language Processing (NLP),Predictive Analytics,Testing and Experimentation,Big Data and Cloud Computing,Ethical Considerations.

Innovation in market basket insights often involves a combination of these techniques to stay competitive and extract valuable insights from transactional data. The choice of technique depends on the specific goals of the project, the nature of the data, and available resources.

PROGRAM:

import pandas as pd

from mlxtend.frequent\_patterns import apriori

from mlxtend.frequent\_patterns import association\_rules

# Load your transaction data into a DataFrame (each row represents a transaction, and each column represents an item)

# Replace 'transaction\_data.csv' with your own data file

data = pd.read\_csv('transaction\_data.csv')

# Convert data to one-hot encoded format

def encode\_data(data):

data\_encoded = pd.get\_dummies(data, columns=['item\_column']) # Replace 'item\_column' with the actual column name containing item data

return data\_encoded

# Perform market basket analysis

def market\_basket\_analysis(data\_encoded, min\_support=0.01, min\_confidence=0.5):

# Perform frequent itemset mining using Apriori

frequent\_itemsets = apriori(data\_encoded, min\_support=min\_support, use\_colnames=True)

# Generate association rules

rules = association\_rules(frequent\_itemsets, metric='confidence', min\_threshold=min\_confidence)

return rules

if \_\_name\_\_ == "\_\_main\_\_":

data\_encoded = encode\_data(data)

# Set the minimum support threshold (e.g., 0.01, meaning an itemset must appear in at least 1% of transactions)

min\_support = 0.01

# Set the minimum confidence threshold (e.g., 0.5, meaning an association rule must have at least 50% confidence)

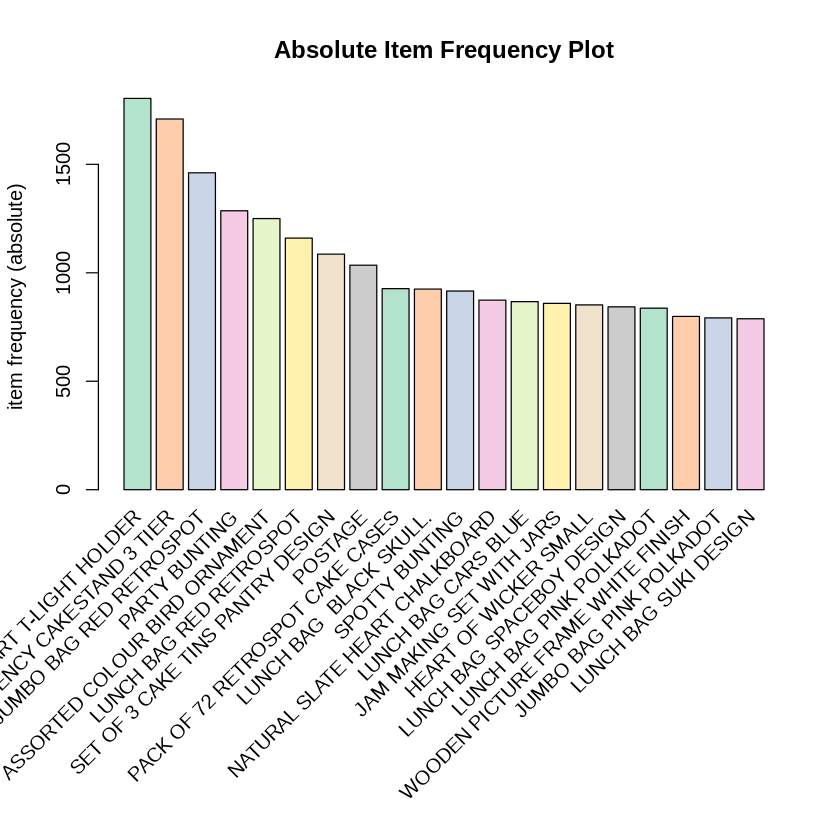
min\_confidence = 0.5

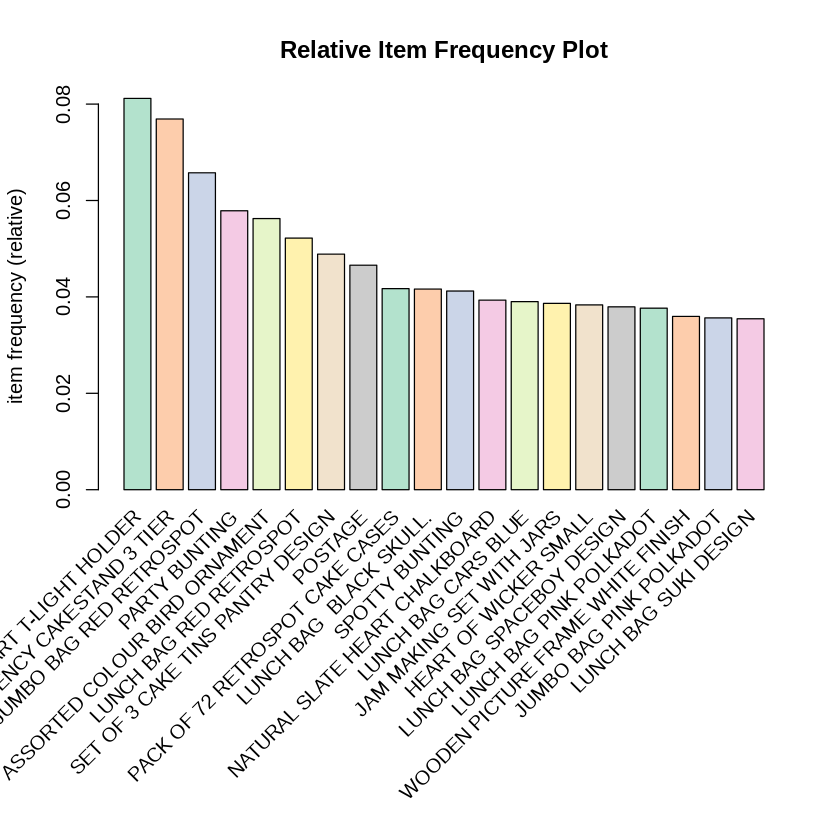
rules = market\_basket\_analysis(data\_encoded, min\_support, min\_confidence)

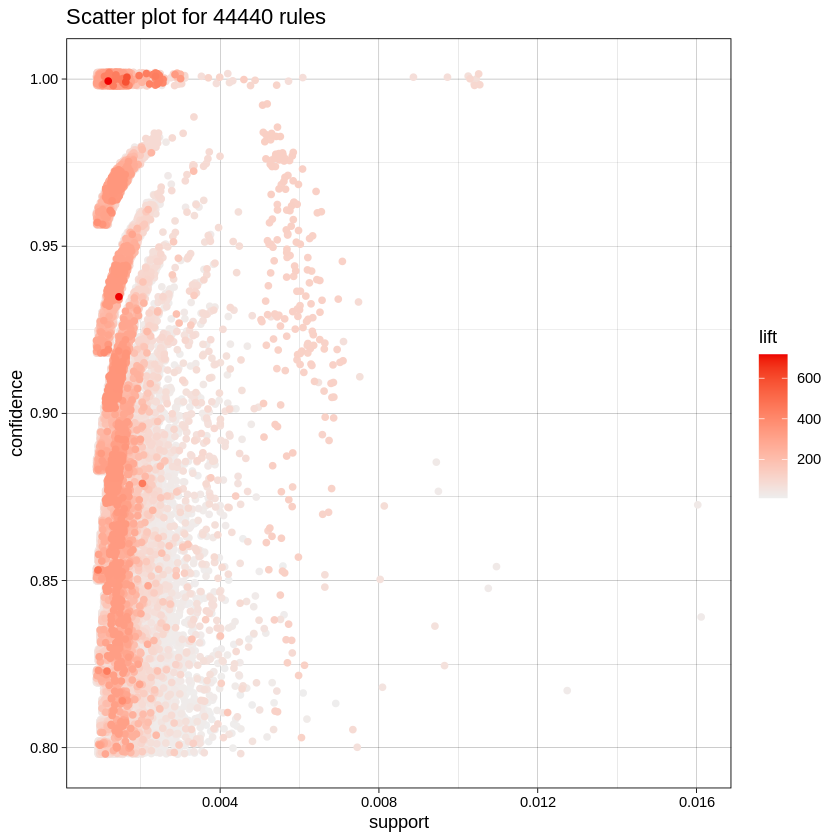
# Print the association rules

print(rules)

OUTPUT:







Summary:

Transforming Shopping Experiences with Market Basket Insights, Through this project, we've delved into the depths of customer behavior, unveiling the secrets hidden within each purchase, each transaction, and each interaction. We've harnessed the power of data, algorithms, and real-time processing to not just analyze the past but also predict and shape the future of shopping.

Our innovative approach, which encompasses advanced analytics, artificial intelligence, personalization, and sustainability integration, has transformed market basket insights from mere numbers and trends into dynamic tools that enrich customer experiences, boost sales, and foster loyalty. The journey is far from over; it's just the beginning.

As we conclude this project, we are poised to launch a revolution in shopping – one that combines the best of data and technology with the essence of human choice. We are ready to provide tailored recommendations, optimized pricing, and sustainable alternatives that not only meet customer needs but anticipate them.