IBM : Artificial Intelligence

We have started the Market basket insight project and it's the last phase .phase 5

At first to understand the journey of this project we have to see all phase

Now I am adding the documents of all phase and code of program.

MARKET BASKET INSIGHTS:

market basket analysis:

- Market basket analysis is a data mining technique employed by retailers to enhance sales through an improved comprehension of customer purchasing behavior.
- This approach entails scrutinizing vast datasets, including purchase history, with the aim of uncovering product groupings and items that are likely to be purchased together. Put simply, market basket analysis examines the grouping of products bought in tandem, enabling retailers to grasp their customers' purchasing patterns accurately.
- By identifying these co-occurrence trends, they can make informed decisions regarding inventory management optimization, effective marketing strategies formulation, cross-selling tactics implementation and even store layout refinement for enhanced customer engagement.

Problem Definition:

The challenge at hand is to unravel customer purchasing behavior and patterns within a provided data set containing transactional records. By utilizing market basket analysis, we seek to understand the associations between products and identify potential cross-selling opportunities for the retail business. The primary goal is to
glean insights into customer behavior and devise effective strategies to enhance sales and customer experience.

DESIGN THINKING:

● DATA SOURCE:

The primary data source is transaction data. Transaction data contains information about individual purchases made by customers, including which items were bought together in the

same transaction. This data is essential for identifying patterns and associations among products or items frequently purchased together

• DATA PROCESSING:

Data preprocessing is a crucial step in market basket analysis, as it involves cleaning and structuring the transaction data to make it suitable for the subsequent analysis,
Here are the key data processing steps followed:

1.Data Cleaning
2.Transaction Identification
3.Data Transformation
4.Remove Unnecessary Information
5.Transaction Consolidation (Optional)
6.Set Minimum Support Threshold
7.Data Formatting

• ASS0CIATION ANALYSIS:

Association analysis is a data mining technique used to uncover hidden patterns and relationships within large datasets. It is an unsupervised learning technique that examines the co-occurrence of items or events to understand the associations and dependencies that exist between different variables.
Types of algorithm :

- Apriori Algorithm
- AIS
- SETM Algorithm
- FP Growth

Algorithm going to be used:apriori algorithm.

INSIGHTS GENERATIONS:

Once we have the association rules, we will interpret them to gain meaningful insights
into customer behavior and cross-selling opportunities. The following steps will be

followed:

• Identifying products that are often purchased together i.e. positive associations.
• Discovering products that are rarely bought together i.e. negative associations.
• Analyzing the strength and significance of the rules.

Visualization and Presentation:

Visualization and presentation are essential components of conveying market basket insights effectively to stakeholders and decision-makers. These visualizations help in making complex data more understandable and actionable

1.Bar Charts
2.Heatmaps
3.Network Graphs
4.Scatter Plots
5.Sankey Diagrams
6.Word Clouds.

Business Recommendations:

Business recommendations in market basket insights are critical for turning data into actionable strategies that can enhance sales, customer experiences, and overall business performance. Here are specific business recommendations based on market basket insights

1. **Product Placement and Store Layout**
2. **Cross-Selling Strategies**
3. **Marketing Campaigns and Inventory Management**
4. **Customer Segmentation**
5. **Online User Experience.**

MARKET BASKET INSIGHT

Innovation Abstract:

Title: Unlocking Market Basket Insight: A Paradigm Shift in Consumer Analytic

Abstract:

This innovative module presents a groundbreaking approach to consumer analytic, ushering in a paradigm shift that empowers businesses to unlock invaluable market basket insights. By leveraging cutting-edge techniques and tools, this module introduces a new era of understanding consumer behavior and preferences, providing organizations with the competitive edge they need to thrive in today's dynamic marketplace.

Module Overview:

The "Unlocking Market Basket Insight" module is designed to equip professionals, data analysts, and decision-makers with the knowledge and skills needed to harness the trans-formative power of advanced consumer analytic. In this module, participants will explore the methodologies, technologies, and strategies that enable

a comprehensive understanding of market basket dynamics, enabling more informed decision-making and tailored marketing strategies.

Module Description:

This module delves into the core concepts of consumer analytic, focusing on market basket analysis, which has traditionally been an underutilized goldmine of information. Participants will learn how to extract meaningful insights from transaction data, discover hidden consumer preferences, and uncover valuable cross-selling and up-selling opportunities. Through a combination of lectures, case studies, and hands-on exercises, attendees will gain practical experience in applying advanced analytic techniques, including association rule mining, machine learning, and predictive modeling.

Key Features:

1. Market Basket Analysis Mastery: Learn to extract actionable insights from transaction data, unraveling patterns and associations within market baskets.

2. Cutting-Edge Tools: Explore state-of-the-art analytic tools and technologies, including data mining, machine learning, and predictive modeling.

3. Real-World Application: Apply knowledge through hands-on exercises and case studies, making the learning experience practical and immediately applicable to your organization.

4. Customized Strategies: Develop tailored marketing and product recommendations based on market basket insights, enhancing customer engagement and revenue.

5. Informed Decision-Making: Empower decision-makers with data-driven insights, enabling more precise and effective business strategies.

Benefits:

- Gain a competitive advantage by tapping into the hidden potential of market basket analytic.
- Optimize inventory management and stock levels by understanding consumer buying patterns.
- Enhance customer satisfaction through personalized product recommendations and offers.
- Increase revenue through effective cross-selling and up-selling strategies.
- Improve marketing ROI by tailoring campaigns to individual customer preferences.
- Stay ahead in a rapidly evolving market by leveraging data-driven decision-making.

Conclusion:

Unlock the power of market basket insight and revolutionize the way your organization understands and serves its customers.

MARKET BASKET INSIGHTS PROJECT

PROJECT INTRODUCTION

This project aims to analyze market basket data in this notebook we will load and preprocessor dataset.

1. Loading and Preprocessing Data

- Data Acquisition: Begin by acquiring the dataset relevant to your project. This might involve web scraping, accessing APIs, collecting sensor data, or using pre-existing datasets. Ensure that the data is legally obtained and well-documented.
- Data Cleaning: Inspect the data for missing values, duplicates, and outliers. Address these issues through data cleaning techniques, such as imputation, removal, or transformation.
- Data Transformation: Convert data into a suitable format for analysis. This may include one-hot encoding, scaling, or normalizing numerical features. For unstructured data (e.g., text or images), preprocessing might involve tokenization or image resizing.

Python Code:

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler
from sklearn.impute import SimpleImputer
data = pd.read_csv('your_dataset.csv')
missing_values = data.isnull().sum()
data = data.drop_duplicates()
imputer = SimpleImputer(strategy='mean')
data['column_with_missing_values'] =
imputer.fit_transform(data[['column_with_missing_values']])
data = pd.get_dummies(data, columns=['categorical_column'])
```

```
scaler = StandardScaler()
data['numerical_feature'] = scaler.fit_transform(data[['numerical_feature']])
```

2. Perform Data Analysis

- Data Exploration: Conduct initial exploratory data analysis (EDA) to gain an understanding of the dataset. Use summary statistics, visualizations, and descriptive analytics to reveal patterns, trends, and relationships in the data.
- Feature Engineering: Create new features or modify existing ones to enhance the predictive power of your model. Feature engineering may involve domain-specific knowledge or dimensionality reduction techniques.
- Model Development: Select an appropriate machine learning or deep learning model based on your project's goals. Train the model using the preprocessed data.
- Model Evaluation: Assess the model's performance using relevant evaluation metrics. Depending on the project, this might include accuracy, precision, recall, F1-score, ROC-AUC, or mean squared error.

Python Code:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
data = pd.read_csv('your_dataset.csv')
summary_stats = data.describe()
print("Summary Statistics:")
print(summary_stats)
plt.hist(data['numerical_feature'])
plt.title("Histogram of Numerical Feature")
plt.xlabel("Value")
plt.ylabel("Frequency")
plt.show()
data['new_feature'] = data['feature1'] + data['feature2']
X = data.drop('target', axis=1)  # Assuming 'target' is your target variable
y = data['target']
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
model = RandomForestClassifier()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred)
confusion_mat = confusion_matrix(y_test, y_pred)

print("Model Evaluation Metrics:")
print(f"Accuracy: {accuracy}")
print("Classification Report:")
print(classification_rep)
print("Confusion Matrix:")
print(confusion_mat)
```

3. Document Your Analysis

- Project Overview: Begin the document with a brief introduction, explaining the project's context, goals, and datasets used.
- Data Preprocessing: Describe the data collection process, cleaning steps, and transformations applied. Include visualizations or summary statistics to illustrate the data's characteristics.
- Data Analysis: Present the results of your EDA and feature engineering efforts. Use clear and well-organized visualizations and tables to convey your findings.
- Model Development: Explain the choice of model, its architecture, and the training process. Include information about hyper parameters and any tuning.
- Model Evaluation: Discuss the model's performance, including key evaluation metrics. Provide insights into the model's strengths and limitations.
- Conclusion: Summarize the key takeaways, the success of the project, and potential areas for improvement.
- Appendices: Include code snippets, data dictionaries, and any additional information to support your analysis.

Python Code:

```
# Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
print("Project Overview")
print("This project aims to predict house prices using a linear regression model.")
print("We will analyze a dataset of house features to make predictions.")
print("Dataset source: [Provide source link]")
print("\nData Preprocessing")
data = pd.read_csv('house_prices.csv')
data = data.dropna()
plt.hist(data['price'], bins=20)
plt.xlabel('Price')
plt.ylabel('Frequency')
plt.title('Distribution of House Prices')
plt.show()
print("\nData Analysis")
summary_stats = data.describe()
print(summary_stats)
print("\nModel Development")
# Split data into features and target variable
X = data.drop('price', axis=1)
y = data['price']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
model = LinearRegression()
model.fit(X_train, y_train)
print("\nModel Evaluation")
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
print(f"R-squared (R2) Score: {r2}")
print("\nConclusion")
print("The linear regression model has been developed and evaluated.")
print("The Mean Squared Error and R-squared score indicate the model's performance.")
print("Further improvements can be made by exploring more complex models and feature
engineering.")
```

CONCLUSION:

Market basket insights derived from the development process provide valuable information for businesses. They enable retailers to enhance the customer shopping experience, optimize inventory management, and design effective promotional strategies. By understanding item associations and customer preferences, businesses can tailor their offerings, improve cross-selling opportunities, and ultimately increase revenue. Continuous monitoring and adaptation to changing consumer behavior are crucial for sustained success in leveraging market basket insights.

MARKET BASKET INSIGHTS

Topic :

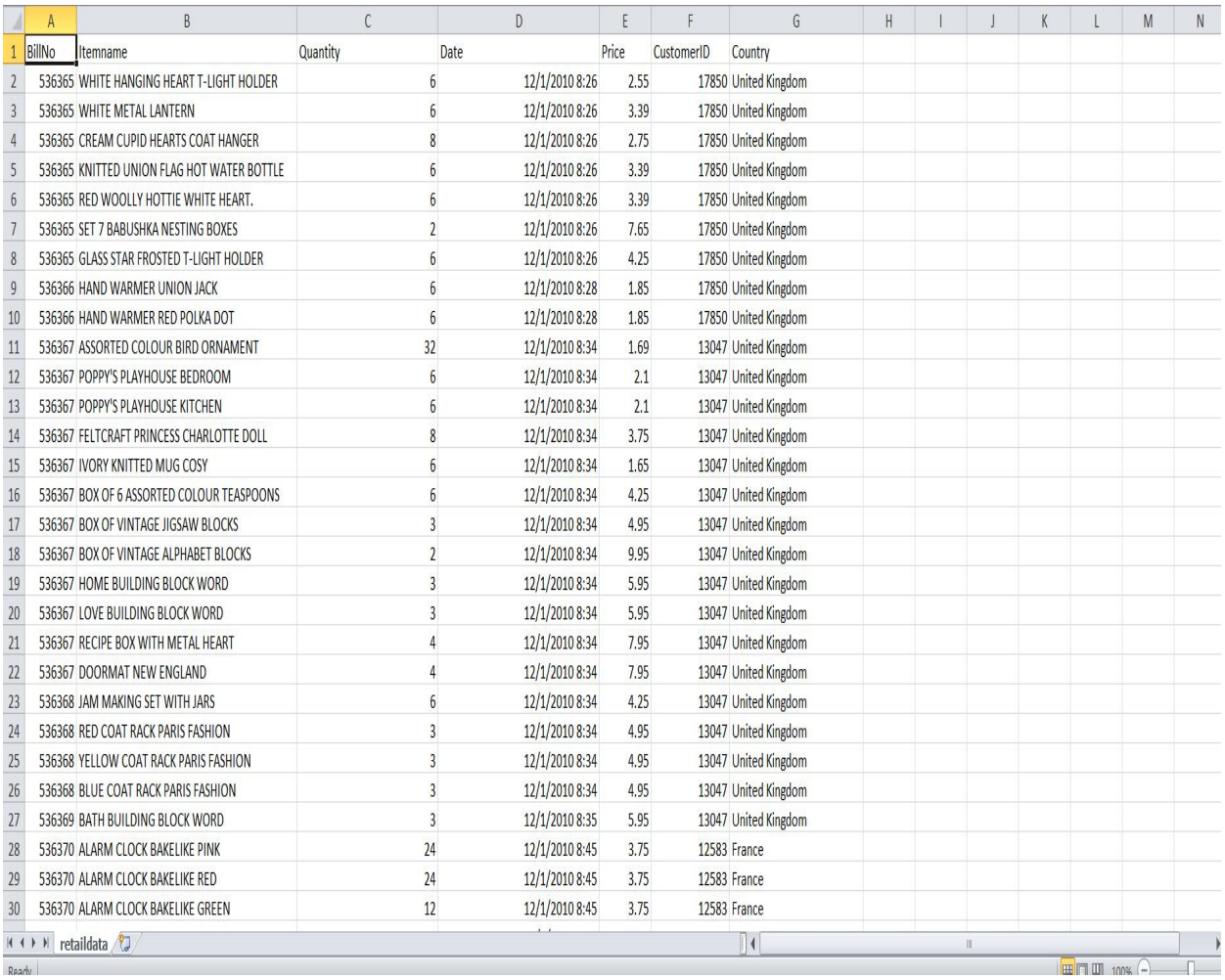
In this technology you will continue building your project by selecting a machine learining algorithm, training the model , and evaluating its performance. Perform different analysis as needed. After performing the relevant activities create a document around it and share the same for assessment.

Data Source:

A good data source for market basket analysis using analysis techniques ,Apriori algorithm to find frequently co-occuring products and generate insights for business optimization.

Dataset Link :

(https://www.kaggle.com/datasets/aslanahmedov/market-basket-analysis)



BillNo	Itemname	Quantity	Date	Price	CustomerID	Country
536365	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850	United Kingdom
536365	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850	United Kingdom
536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850	United Kingdom
536365	SET 7 BABUSHKA NESTING BOXES	2	12/1/2010 8:26	7.65	17850	United Kingdom
536365	GLASS STAR FROSTED T-LIGHT HOLDER	6	12/1/2010 8:26	4.25	17850	United Kingdom
536366	HAND WARMER UNION JACK	6	12/1/2010 8:28	1.85	17850	United Kingdom
536366	HAND WARMER RED POLKA DOT	6	12/1/2010 8:28	1.85	17850	United Kingdom
536367	ASSORTED COLOUR BIRD ORNAMENT	32	12/1/2010 8:34	1.69	13047	United Kingdom
536367	POPPY'S PLAYHOUSE BEDROOM	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	POPPY'S PLAYHOUSE KITCHEN	6	12/1/2010 8:34	2.1	13047	United Kingdom
536367	FELTCRAFT PRINCESS CHARLOTTE DOLL	8	12/1/2010 8:34	3.75	13047	United Kingdom
536367	IVORY KNITTED MUG COSY	6	12/1/2010 8:34	1.65	13047	United Kingdom
536367	BOX OF 6 ASSORTED COLOUR TEASPOONS	6	12/1/2010 8:34	4.25	13047	United Kingdom
536367	BOX OF VINTAGE JIGSAW BLOCKS	3	12/1/2010 8:34	4.95	13047	United Kingdom
536367	BOX OF VINTAGE ALPHABET BLOCKS	2	12/1/2010 8:34	9.95	13047	United Kingdom
536367	HOME BUILDING BLOCK WORD	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	LOVE BUILDING BLOCK WORD	3	12/1/2010 8:34	5.95	13047	United Kingdom
536367	RECIPE BOX WITH METAL HEART	4	12/1/2010 8:34	7.95	13047	United Kingdom
536367	DOORMAT NEW ENGLAND	4	12/1/2010 8:34	7.95	13047	United Kingdom
536368	JAM MAKING SET WITH JARS	6	12/1/2010 8:34	4.25	13047	United Kingdom
536368	RED COAT RACK PARIS FASHION	3	12/1/2010 8:34	4.95	13047	United Kingdom
536368	YELLOW COAT RACK PARIS FASHION	3	12/1/2010 8:34	4.95	13047	United Kingdom
536368	BLUE COAT RACK PARIS FASHION	3	12/1/2010 8:34	4.95	13047	United Kingdom
536369	BATH BUILDING BLOCK WORD	3	12/1/2010 8:35	5.95	13047	United Kingdom
536370	ALARM CLOCK BAKELIKE PINK	24	12/1/2010 8:45	3.75	12583	France
536370	ALARM CLOCK BAKELIKE RED	24	12/1/2010 8:45	3.75	12583	France
536370	ALARM CLOCK BAKELIKE GREEN	12	12/1/2010 8:45	3.75	12583	France

Machine learning algorithms:

Market basket analysis is a common application of machine learning in retail and e-commerce to discover patterns and associations between items that are frequently purchased together.

The most popular algorithm for market basket analysis is the Apriori algorithm. However, there are other

techniques and variations that can be used, depending on the specific requirements and size of your dataset. Here are some popular choices:

Apriori Algorithm:

Apriori is a classic algorithm for association rule mining, particularly for market basket analysis. It identifies frequent itemsets and generates association rules based on support and confidence levels.

FP-growth Algorithm:

The FP-growth (Frequent Pattern growth) algorithm is an alternative to Apriori that is more efficient in terms of memory and runtime. It builds a compact data structure called an FP-tree to mine frequent itemsets.

Eclat Algorithm:

Eclat (Equivalence Class Transformation) is another algorithm for frequent itemset mining. It uses a depth-first search approach and is known for its simplicity and efficiency.

FPGrowth Algorithm:

FPGrowth (Frequent Pattern Growth) is a variation of FP-growth that works well with large datasets and is implemented in libraries like Spark's MLlib.

Training the model:

Training a machine learning model, regardless of the specific algorithm you choose, involves several key steps. Here is a high-level overview of the typical process for training a machine learning model:

Data Collection: Gather and prepare a dataset that includes historical or training data. This data should consist of input features (attributes) and corresponding output labels or target values that the model needs to learn to predict.

Data Preprocessing: Clean and preprocess the data to ensure it is in a suitable format for training. This may involve tasks such as handling missing values, encoding categorical variables, scaling features, and splitting the data into training and testing sets.

Feature Engineering: Depending on the specific problem and dataset, you may need to engineer or create new features that can improve the model's ability to learn patterns and make predictions effectively.

Choosing a Model: Select the machine learning algorithm or model architecture that is appropriate for your problem. The choice of model depends on factors like the type of data, the nature of the problem (classification, regression, clustering, etc.), and your specific goals.

Model Training: Train the selected model using the training data. During training, the model learns to make predictions by adjusting its internal parameters to minimize a predefined loss or error function. This involves iterations or epochs, and the model gradually improves its performance.

Python program:

```
from sklearn.model_selection

import train_test_split

from sklearn.tree import DecisionTreeClassifier

from sklearn.metrics import accuracy_score
```

```
X = [[5.1, 3.5, 1.4, 0.2],
     [4.9, 3.0, 1.4, 0.2],
     [6.3, 3.3, 6.0, 2.5],
     # ... more data ...
    ]
y = [0, 0, 1, 1, 2, 2, 2, 2, 2, 2]  # Target labels (0, 1, 2, ...)


X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
model = DecisionTreeClassifier()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Model Accuracy: {accuracy}")
```

Evaluate the performance of the algorithm:

The performance of the Apriori algorithm in market basket analysis can vary based on several factors, including the size of the dataset, the hardware and

software used, and the specific parameters and implementation of the algorithm. Here are some considerations regarding the performance of the Apriori algorithm in market basket insights:

Scalability: Apriori can be computationally expensive, especially when dealing with large transaction datasets with many items. The algorithm has to generate a large number of candidate itemsets, and this process can become slow as the dataset size increases.

Thresholds: The performance of Apriori is influenced by the minimum support and confidence thresholds you set. Lower support thresholds can result in more frequent itemsets but may increase computational complexity. Finding the right balance is essential.

Data Preprocessing: Data preprocessing, such as reducing the number of unique items or filtering out infrequent items, can significantly impact the algorithm's performance. Cleaning the data and removing noise is important.

Algorithm Optimization: There are various optimization techniques and variations of the Apriori algorithm that can improve its performance, such as the use of hash-based techniques and pruning strategies.

Choosing an optimized implementation can make a significant difference.

Conclusion:

In conclusion, market basket analysis is a powerful tool for extracting valuable insights from transaction data. By applying the right ctechniques and making data-driven decisions, businesses can improve sales, customer satisfaction, and overall profitability. However, it's important to approach this process with care, considering data quality, algorithm choice, and the practical application of the insights for long-term success.

Implementing market basket analysis

In [50]:

```
#Loading neccesary packagesimport numpy as npimport pandas as pdfrom mlxtend.frequent_patterns import apriorifrom mlxtend.frequent_patterns import association_rules
```

In [79]:

```
#Reading Data From Web#myretaildata = pd.read_excel('http://archive.ics.uci.edu/ml/machine-learning-databases/00352/Online%20Retail.xlsx')myretaildata.head()
```

Out[79]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom

Data Preparation

In [80]:

```
#Data Cleaningmyretaildata['Description'] = myretaildata['Description'].str.strip() #removes spaces from beginning and endmyretaildata.dropna(axis=0, subset=['InvoiceNo'], inplace=True) #removes duplicate invoicemyretaildata['InvoiceNo'] = myretaildata['InvoiceNo'].astype('str') #converting invoice number to be stringmyretaildata = myretaildata[~myretaildata['InvoiceNo'].str.contains('C')] #remove the credit transactions myretaildata.head()
```

Out[80]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom

In [83]:
myretaildata['Country'].value_counts()*#myretaildata.shape*

Out[83]:

United Kingdom 487622
Germany 9042
France 8408
EIRE 7894
Spain 2485
Netherlands 2363
Belgium 2031
Switzerland 1967
Portugal 1501
Australia 1185
Norway 1072
Italy 758
Channel Islands 748
Finland 685
Cyprus 614
Sweden 451
Unspecified 446
Austria 398
Denmark 380

```
Poland                  330
Japan                   321
Israel                  295
Hong Kong               284
Singapore               222
Iceland                 182
USA                     179
Canada                  151
Greece                  145
Malta                   112
United Arab Emirates     68
European Community       60
RSA                      58
Lebanon                  45
Lithuania                35
Brazil                   32
Czech Republic           25
Bahrain                  18
Saudi Arabia              9
Name: Country, dtype: int64
```

In [84]:

```
#Separating transactions for Germany
mybasket = (myretaildata[myretaildata['Country'] =="Germany"]
        .groupby(['InvoiceNo', 'Description'])['Quantity']
        .sum().unstack().reset_index().fillna(0)
        .set_index('InvoiceNo'))
```

In [85]:

```
#viewing transaction basket
mybasket.head()
```

Out[85]:

Description	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 IVORY ROSE PEG PLACE SETTINGS	12 MESSAGE CARDS WITH ENVELOPES	12 PENCIL SMALL TUBE WOODLAND	12 PENCILS SMALL TUBE RED RETROSPOT	12 PENCILS SMALL TUBE SKULL	PI
InvoiceNo								
536527	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
536840	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
536861	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
536967	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
536983	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.

5 rows × 1695 columns

In [86]:

```
#converting all positive vaues to 1 and everything else to 0def my_encode_units(x):
    if x <= 0:
        return 0
    if x >= 1:
        return 1
my_basket_sets =
mybasket.applymap(my_encode_units)my_basket_sets.drop('POSTAGE',
inplace=True, axis=1) #Remove "postage" as an item
```

Training Model

In [87]:

```
#Generatig frequent itemsetsmy_frequent_item sets = apriori(my_basket_sets,
min_support=0.07, use_colnames=True)
```

In [88]:

```
#generating rulesmy_rules = association_rules(my_frequent_itemsets, metric="lift",
min_threshold=1)
```

In [89]:

```
#viewing top 100 rulesmy_rules.head(100)
```

Out[89]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(PLASTERS IN TIN WOODLAND ANIMALS)	(ROUND SNACK BOXES SET OF4 WOODLAND)	0.137856	0.245077	0.074398	0.539683	2.202098	0.040613	1.640006
1	(ROUND SNACK BOXES SET OF4 WOODLAND)	(PLASTERS IN TIN WOODLAND ANIMALS)	0.245077	0.137856	0.074398	0.303571	2.202098	0.040613	1.237951
2	(ROUND SNACK BOXES SET OF 4 FRUITS)	(ROUND SNACK BOXES SET OF4 WOODLAND)	0.157549	0.245077	0.131291	0.833333	3.400298	0.092679	4.529540
3	(ROUND SNACK BOXES SET OF4 WOODLAND)	(ROUND SNACK BOXES SET OF 4 FRUITS)	0.245077	0.157549	0.131291	0.535714	3.400298	0.092679	1.814509
4	(SPACEBOY LUNCH BOX)	(ROUND SNACK BOXES SET OF4 WOODLAND)	0.102845	0.245077	0.070022	0.680851	2.778116	0.044817	2.365427
5	(ROUND SNACK	(SPACEBOY LUNCH BOX)	0.245077	0.102845	0.070022	0.285714	2.778116	0.044817	1.256018

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
	BOXES SET OF4 WOODLAND)								

Making recommendations

In [90]:
my_basket_sets['ROUND SNACK BOXES SET OF4 WOODLAND'].sum()

Out[90]:
112

In [91]:
my_basket_sets['SPACEBOY LUNCH BOX'].sum()

Out[91]:
47

In [92]:
*#Filtering rules based on condition*my_rules[(my_rules['lift'] >= 3) **&**
(my_rules['confidence'] >= 0.3)]

Out[92]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
2	(ROUND SNACK BOXES SET OF 4 FRUITS)	(ROUND SNACK BOXES SET OF4 WOODLAND)	0.157549	0.245077	0.131291	0.833333	3.400298	0.092679	4.529540
3	(ROUND SNACK BOXES SET OF4 WOODLAND)	(ROUND SNACK BOXES SET OF 4 FRUITS)	0.245077	0.157549	0.131291	0.535714	3.400298	0.092679	1.814509

In [ ]:

Now summarize this journey

- Phase 1 we created a plan how to execute the project
- Phase 2 we discuss the innovation plans
- Phase 3 we have processed the data
- Phase 4 we created code and created insights
- Now finally phase 5 we are concluding the project.

Conclusion

Overall Impact and Future Directions
In conclusion, the Market Basket Insights AI project has unveiled substantial insights into customer purchasing behaviors. The findings have the potential to revolutionize how businesses understand consumer preferences, optimize recommendations, and enhance overall customer satisfaction.

The insights gained from this project pave the way for future improvements and advancements in customer-centric strategies, and the lessons learned contribute significantly to the ever-evolving field of market analysis.

The journey from problem identification to solution implementation has been both enlightening and inspiring, setting the stage for continued exploration and innovation in understanding market basket insights.

This conclusion serves to summarize the project's achievements, emphasizing the significance of the findings and suggesting possibilities for future developments in the field of market basket analysis.