MARKET BASKET INSIGHTS

TEAM MEMBERS:
1.EMMANUAL JUDE ANTONY.J
2.ARUN.A
3.MURUGAVELL.S.T
4.CHANDRU.T
5.DHANUSH.BP

Problem Statement:

"Analyze customer purchase data from a retail store to uncover valuable insights into market basket behavior. Develop a data-driven strategy to understand which products are frequently purchased together, identify cross-selling opportunities, and optimize product placement within the store. The goal is to increase sales, enhance customer satisfaction, and improve overall business profitability through a deeper understanding of customer shopping patterns."

This problem statement sets the stage for a project that involves analyzing transaction data to gain insights into customer behavior and make data-driven decisions to improve the retail store's operations and sales strategies.

1. DATA SOURCE:

1. Empathize: Understand the Stakeholders

Identify and empathize with the stakeholders involved, such as business owners, store managers, and data analysts.
Conduct interviews and surveys to gather their perspectives, pain points, and objectives regarding market basket insights.
Explore the existing data sources and understand the limitations and opportunities they present.

2. Define: Problem Definition and Objectives

Clearly define the problem statement and project objectives, aligning them with the stakeholders' needs and goals.
Specify the key metrics and KPIs that will be used to measure the success of the project (e.g., increase in sales, better customer satisfaction, improved inventory management).
Set a timeline and budget for the project.

3. Ideate: Data Source Selection and Collection Strategy

Brainstorm potential data sources that can provide insights into market basket behavior. This may include transaction data, customer profiles, inventory data, and even external data sources like weather or events data.
Evaluate the feasibility, quality, and accessibility of each data source.

Choose the most relevant and valuable data sources that align with the project objectives.
Develop a data collection and integration strategy, considering the frequency and volume of data, data storage, and data quality assurance.

4. Prototype: Data Collection and Preprocessing

Create a prototype data collection process. This may involve setting up automated data feeds from point-of-sale systems, e-commerce platforms, or loyalty programs.
Implement data preprocessing steps, such as data cleaning, transformation, and enrichment, to ensure that the data is ready for analysis.
Establish data governance practices to maintain data quality and security.

5. Test: Exploratory Data Analysis (EDA) and Modeling

Conduct exploratory data analysis to gain a deeper understanding of the data and identify initial patterns and trends in market basket behavior.
Apply data mining and machine learning techniques to build predictive models for market basket insights. Common approaches include association rule mining, collaborative filtering, and recommendation algorithms.
Continuously iterate on the modeling process, refining models based on feedback and results.

6. Implement: Insights and Recommendations

Translate the model results into actionable insights for stakeholders.
Identify which products are frequently purchased together, recommend product placements, and uncover cross-selling opportunities.
Develop data visualizations and dashboards to present the insights in a clear and understandable manner.
Collaborate with stakeholders to implement recommendations in the retail store, whether through in-store promotions, layout changes, or targeted marketing campaigns.
Continuously monitor the impact of the implemented recommendations on key performance indicators (KPIs).
Gather feedback from store staff and customers to assess the effectiveness of the changes.
Use the feedback and data-driven insights to iterate on the strategies, refining the data sources, models, and recommendations as needed.8.
Scale: Extend Insights and Optimize

If successful, consider scaling the insights and strategies to other stores or locations.
Explore opportunities to use the same data-driven approach for other aspects of the business, such as inventory management, supply chain optimization, and customer segmentation.
Throughout this design thinking process, collaboration and communication with stakeholders and data experts are essential for creating a successful market basket insights solution. Additionally, flexibility and adaptability in response to evolving customer behavior and market dynamics are key to long-term success.

2. DATA PREPROCESSING:

1. Empathize: Understand Data Sources and Stakeholder Needs

Identify the various data sources available, such as transaction records, product databases, customer profiles, and any external data that might be relevant.
Conduct interviews or meetings with data analysts, business owners, and domain experts to understand their specific requirements and expectations regarding data preprocessing.

2. Define: Define Data Preprocessing Objectives

Clearly define the goals and objectives of data preprocessing in the context of market basket insights. These objectives should align with the overall project goals.
Determine what specific issues need to be addressed during preprocessing, such as missing data, data inconsistencies, or outliers.

3. Ideate: Data Preprocessing Strategies

Brainstorm data preprocessing techniques and strategies that can address the identified issues. Common preprocessing steps for market basket insights may include:
Data cleaning: Removing duplicates, handling missing values, and correcting errors.
Data transformation: Encoding categorical variables, scaling numeric features, and creating derived variables.
Outlier detection: Identifying and handling outliers that may affect the analysis.

Data integration: Combining data from multiple sources into a unified dataset.
Feature engineering: Creating new features that may enhance market basket analysis, such as transaction frequency or customer segmentation.
Consider the sequence of preprocessing steps and their dependencies.

4. Prototype: Data Preprocessing Workflow

Create a prototype data preprocessing workflow that outlines the sequence of preprocessing steps, tools, and technologies to be used.
Develop scripts or code snippets to perform each preprocessing task.
Define data quality metrics to assess the effectiveness of data preprocessing.

5. Test: Data Preprocessing Validation and Quality Assurance

Test the preprocessing workflow on a subset of the data to ensure it functions correctly and achieves the defined objectives.
Establish data quality validation criteria and conduct data quality checks to measure the impact of preprocessing on data quality.
Use visualization tools to explore the data before and after preprocessing, highlighting improvements and changes.

6. Implement: Automate Data Preprocessing

Automate the data preprocessing workflow to ensure it can be applied consistently to new data as it becomes available.
Set up data pipelines or workflows that trigger preprocessing tasks as new data arrives.
Implement error handling and logging mechanisms to monitor the preprocessing process for issues or failures.

7. Evaluate: Measure Data Preprocessing Impact

Measure the impact of data preprocessing on downstream analysis.
Assess whether the preprocessing steps have improved the quality and suitability of the data for market basket analysis.
Gather feedback from data analysts and domain experts to identify any potential improvements or adjustments needed in the preprocessing workflow.

8. Scale: Scalability and Ongoing Maintenance

Consider the scalability of the data preprocessing workflow as the volume of data increases.
Establish a schedule for regular data preprocessing maintenance to account for changes in data sources, schema updates, or evolving business requirements.
Continuously monitor the performance and effectiveness of the preprocessing process, making adjustments as needed.

9. Feedback Loop: Continuous Improvement

Establish a feedback loop with stakeholders to gather ongoing feedback and adapt the data preprocessing workflow to evolving needs.
Explore opportunities to incorporate advanced data preprocessing techniques, such as machine learning-based anomaly detection or natural language processing for unstructured data.
By following this design thinking framework, you can ensure that the data preprocessing phase is systematic, effective, and adaptable to the evolving needs of market basket insights. The goal is to provide clean and well-structured data that facilitates accurate and meaningful analysis.

3. ASSOCIATION ANALYSIS:

1. Empathize: Understand the Business Context

Begin by understanding the business objectives and challenges. Talk to key stakeholders, such as store managers, marketing teams, and data analysts, to gain insights into their specific needs and goals related to market basket insights.

2. Define: Define the Scope and Objectives

Clearly define the scope of the association analysis project. Determine the specific questions you aim to answer, such as identifying frequently co-purchased items, understanding customer preferences, or optimizing product placement.
Set measurable objectives, such as improving cross-selling rates or increasing average transaction values through association analysis.

3. Ideate: Data Collection and Preprocessing

Identify the data sources required for association analysis, such as transaction data, product databases, and customer profiles.

Brainstorm strategies for collecting and preprocessing the data to make it suitable for association analysis. This may include data cleaning, encoding categorical variables, and handling missing values.
Consider how you will handle transactional data, including item sets, timestamps, and customer identifiers.

4. Prototype: Association Analysis Techniques

Explore various association analysis techniques and algorithms, such as Apriori, FP-Growth, or Eclat. Evaluate which one(s) are most suitable for your specific objectives and dataset.
Create a prototype workflow that outlines the steps involved in association analysis, from generating itemsets to discovering association rules.
Develop code or scripts for implementing association analysis on sample data.

5. Test: Validate and Tune

Test the association analysis prototype on a subset of your data to validate its effectiveness and efficiency.
Fine-tune the parameters of the chosen association analysis algorithm(s) to optimize rule generation and discover meaningful associations.

6. Implement: Full-scale Association Analysis

Implement the association analysis process on the entire dataset. This may involve setting up data pipelines or workflows for automated analysis.
Generate and evaluate association rules based on the objectives defined earlier.
Visualize and present the discovered associations using charts, graphs, or reports for better understanding.

7. Evaluate: Interpretation and Insights

Interpret the results of association analysis. Identify key insights, such as frequently co-purchased items, lift values, and confidence levels.
Collaborate with domain experts and stakeholders to understand the business implications of the discovered associations.
Determine how the insights can be translated into actionable strategies, such as recommending related products, adjusting pricing, or optimizing store layouts.

8. Scale: Integration and Ongoing Analysis

Explore opportunities to integrate association analysis insights into the daily operations of the business, such as e-commerce recommendation systems or in-store promotions.
Consider the scalability of the association analysis process as data volumes grow and adapt the infrastructure accordingly.
Establish a schedule for regular updates and re-analysis to capture changing customer preferences and market trends.

9. Feedback Loop: Continuous Improvement

Maintain an open feedback loop with stakeholders to gather feedback on the effectiveness of association analysis insights.
Continuously improve the association analysis process by refining algorithms, adding new data sources, or enhancing the interpretability of results.
By following this design thinking framework, you can systematically design and execute an association analysis project that uncovers valuable insights into market basket behavior, helping to improve sales, customer satisfaction, and overall business profitability.

4. INSIGHTS GENERATION:

1. Empathize: Understand Stakeholder Needs

Begin by understanding the needs and goals of the stakeholders, including business owners, store managers, and marketing teams.
Conduct interviews and gather insights into their specific challenges and objectives related to market basket insights.

2. Define: Define Objectives and Key Metrics

Clearly define the objectives of the insights generation process. What specific insights are you aiming to extract from the market basket data?
Identify key metrics and KPIs that will be used to measure the success of the insights generation process, such as increased revenue, improved customer retention, or enhanced product recommendations.

3. Ideate: Data Collection and Preprocessing

Determine the data sources required for generating insights, which typically include transaction data, customer profiles, and product information.
Brainstorm strategies for collecting, preprocessing, and integrating these data sources to create a unified dataset that's ready for analysis. Consider data cleaning, feature engineering, and data enrichment.

4. Prototype: Exploratory Data Analysis (EDA) and Modeling

Conduct exploratory data analysis (EDA) to gain a deeper understanding of the data. Visualize the data to identify trends, patterns, and anomalies.
Develop prototype models or analytical techniques to extract insights from the data. This may involve techniques such as clustering, classification, association analysis, or predictive modeling.

5. Test: Validation and Model Tuning

Test the prototype insights generation models on a subset of the data to validate their effectiveness. Ensure that the models are generating meaningful and actionable insights.
Fine-tune the models and analytical techniques to optimize their performance and accuracy.

6. Implement: Full-scale Insights Generation

Implement the insights generation process on the entire dataset. This may involve setting up automated workflows or pipelines to ensure regular insights generation.
Generate insights on a regular basis (e.g., daily, weekly, or monthly) to keep the information fresh and relevant.

7. Evaluate: Interpretation and Actionable Insights

Interpret the results generated by the models and analytical techniques. Identify key insights, trends, and correlations in the market basket data.
Collaborate with stakeholders to understand the business implications of the insights and determine how they can be translated into actionable strategies.

8. Scale: Integration and Deployment

Explore opportunities to integrate the generated insights into various aspects of the business, such as marketing campaigns, product recommendations, or store operations.
Consider the scalability of the insights generation process as the volume of data and complexity of analyses increase.

9. Feedback Loop: Continuous Improvement

Maintain an ongoing feedback loop with stakeholders to gather feedback on the usefulness and impact of the generated insights.
Continuously improve the insights generation process by updating models, incorporating new data sources, and adapting to changing business needs.

10. Ethical Considerations: Data Privacy and Fairness

Ensure that the insights generation process adheres to data privacy regulations and ethical standards. Protect customer information and handle sensitive data appropriately.
Monitor for biases in the data and models to ensure fair and equitable insights.
By following this design thinking framework, you can systematically design and execute an insights generation process that leverages market basket data to drive informed business decisions and improve overall business performance.

5. VISUALIZATION:

1. Empathize: Understand the Audience and Objectives

Begin by understanding the audience for your visualizations, such as business owners, store managers, or data analysts. Consider their needs, preferences, and levels of data literacy.
Clarify the specific objectives of the visualizations. What insights or messages are you trying to convey through the visualizations?

2. Define: Define Key Metrics and Data Sources

Identify the key metrics and KPIs that you want to visualize, such as sales trends, product associations, or customer segmentation.
Determine the data sources and datasets that will be used for creating the visualizations, ensuring they are clean and well-preprocessed.

3. Ideate: Choose Visualization Types

Brainstorm various types of visualizations that are suitable for conveying the chosen metrics and insights. Common visualization types for market basket analysis include:
Bar charts and histograms to show product frequency.
Scatter plots or bubble charts to visualize associations between products.
Heatmaps to display item co-occurrence patterns.
Line charts to illustrate sales trends over time.
Sankey diagrams to depict customer journeys.
Consider interactive elements like filters, drill-down options, or tooltips to enhance user engagement.

4. Prototype: Create Initial Visualizations

Develop prototypes of the selected visualizations using sample data. You can use data visualization tools or libraries like Tableau, Power BI, Matplotlib, or D3.js.
Design the visualizations with a focus on clarity, simplicity, and aesthetics. Ensure that they effectively communicate the intended insights.
Test the prototypes with a small group of users or stakeholders to gather initial feedback.

5. Test: Gather Feedback and Refine

Conduct usability testing to gather feedback on the prototypes. Ask users to provide insights into the effectiveness and usability of the visualizations.
Use feedback to refine the visualizations, making adjustments to layout, colors, labels, and interactive features as needed.
Ensure that the visualizations are accessible to all users, including those with disabilities.

6. Implement: Develop Final Visualizations

Based on the feedback and refinements, create the final versions of the visualizations using real data.
Implement any necessary data connectors or APIs to ensure that the visualizations can be updated with fresh data regularly.

7. Evaluate: Assess Effectiveness

Assess the effectiveness of the visualizations in conveying insights. Analyze whether they meet the defined objectives and provide actionable information.
Collect feedback from users and stakeholders after the final visualizations are deployed to ensure they are valuable and useful.

8. Scale: Deployment and Accessibility

Deploy the visualizations to the appropriate platforms or channels where they will be accessed by the target audience. This could be within a business intelligence dashboard, a web application, or reports.
Ensure that the visualizations are accessible on various devices and screen sizes.

9. Feedback Loop: Continuous Improvement

Establish a feedback loop with users and stakeholders to gather ongoing feedback about the visualizations.
Continuously improve the visualizations based on user suggestions, changing business needs, or evolving data patterns.

10. Ethical Considerations: Data Privacy and Fairness

Ensure that the visualizations respect data privacy regulations and ethical standards. Avoid displaying sensitive customer information.
Monitor for biases in the visualizations and data representations to ensure fairness and accuracy.
By following this design thinking framework, you can create visualizations that effectively communicate market basket insights, enabling better decision-making and business optimization.

6. BUSINESS RECOMMENDATIONS:

1. Empathize: Understand Business Goals and Challenges

Begin by gaining a deep understanding of the business's overarching goals and the specific challenges it faces. Engage with key stakeholders, including business owners, marketing teams, and operational staff.
Listen to their concerns, objectives, and aspirations related to market basket insights and how they impact the business.

2. Define: Define Business Objectives and Metrics

Clearly define the business objectives that market basket insights can address. These objectives should align with the broader business goals. Identify the key performance metrics and KPIs that will be used to measure the success of the business recommendations. For example, increased revenue, improved customer satisfaction, or enhanced cross-selling rates.

3. Ideate: Generate Data-Driven Recommendations

Brainstorm and ideate data-driven recommendations that can address the defined business objectives. These recommendations may include strategies related to product placement, pricing, promotions, and marketing campaigns.
Consider the insights gained from market basket analysis, such as frequently co-purchased items, customer preferences, and seasonality trends, to inform your recommendations.

4. Prototype: Develop Recommendation Strategies

Develop prototypes of the recommendation strategies, outlining the steps and processes required to implement them. Consider the following aspects:
Product recommendations: Determine how personalized product recommendations can be integrated into e-commerce websites or in-store displays.
Pricing strategies: Explore dynamic pricing approaches based on market basket insights.
Promotions: Design targeted promotions and marketing campaigns based on customer purchase patterns.
Inventory management: Create strategies for optimizing inventory levels and product assortment based on market basket data.

5. Test: Validation and Refinement

Test the prototype recommendation strategies on a smaller scale or in a controlled environment to validate their effectiveness and feasibility.
Gather feedback from users or stakeholders involved in the testing phase to identify any issues or areas for improvement.
Fine-tune the recommendation strategies based on the feedback and results.

6. Implement: Full-Scale Recommendation Implementation

Implement the refined recommendation strategies across the business operations. This may involve collaborating with various departments, including marketing, sales, and operations.
Ensure that the necessary technology infrastructure, such as recommendation engines or pricing tools, is in place to support the implementation.

7. Evaluate: Measure Impact and Effectiveness

Continuously monitor and measure the impact of the implemented recommendations on the defined business metrics and KPIs.
Assess the effectiveness of the strategies in achieving the business objectives. Use A/B testing and before-after comparisons to evaluate performance.

8. Scale: Extend and Optimize

If the implemented recommendations prove successful, consider scaling them to other business locations, products, or customer segments.
Optimize the recommendations further based on ongoing analysis and feedback, adapting them to changing market conditions and customer preferences.

9. Feedback Loop: Continuous Improvement

Establish a feedback loop with stakeholders and users to gather ongoing feedback about the recommendations' effectiveness and impact.
Use the feedback to refine and improve the recommendations continually. This iterative process ensures that the business recommendations remain relevant and effective over time.

10. Ethical Considerations: Fairness and Transparency

Ensure that the recommendations are fair and transparent, avoiding biases and discriminatory practices.
Communicate the basis of recommendations clearly to customers to build trust and transparency.
By following this design thinking framework, businesses can leverage market basket insights to generate data-driven recommendations that enhance operations, drive revenue growth, and improve customer satisfaction.

Ensemble techniques:
Ensemble techniques can be valuable in market basket analysis to improve the accuracy of association rule mining and gain deeper insights into customer behavior. Here are some details on how to use ensemble techniques in this context:

1. Association Rule Mining: Market basket analysis typically relies on association rule mining algorithms like Apriori or FP-growth to discover relationships between items purchased together.

2. Single Algorithms: Start by running association rule mining using a single algorithm on your transaction data. This will provide you with basic insights into item associations.

3. Ensemble Techniques: To enhance the accuracy and robustness of your analysis, consider using ensemble techniques such as:

a. Bootstrap Aggregating (Bagging): You can apply bagging to create multiple subsets of your transaction data and mine association rules on each subset. Then, combine the results to obtain more robust rules.

b. Boosting: Boosting methods like AdaBoost can assign weights to transactions or items that are more relevant, thus giving more importance to certain associations.

c. Stacking: Stacking involves training multiple association rule models and then

training another model to combine their predictions. This can be useful for handling different aspects of your market basket analysis.

4. Feature Engineering: You can also use ensemble techniques to create new features or item sets that are combinations of existing items, which might reveal hidden patterns.

5. Pruning and Filtering: Ensemble techniques can help in filtering out less significant associations by considering only those that are consistent across multiple models, reducing the risk of overfitting.

6. Cross-Validation: Use cross-validation to ensure that your ensemble approach doesn't overfit your data. This involves

splitting your data into training and testing sets multiple times to validate the robustness of your models.

7. Evaluation Metrics: Evaluate the performance of your ensemble models using metrics like support, confidence, lift, and conviction to assess the significance and quality of discovered rules.

8. Interpretability: Keep in mind that as you combine multiple models, the interpretability of your insights may become more challenging. Consider using visualization techniques to present the results in a more understandable way.

9. Continuous Monitoring: Market basket insights can change over time, so consider setting up a process for continuous monitoring and updating your ensemble

models as new transaction data becomes available.

Remember that the choice of ensemble technique and parameters should be based on the specific characteristics of your dataset and the insights you aim to extract. It's important to strike a balance between complexity and interpretability in order to derive meaningful and actionable market basket insights.

Deep Learning Techniques:
Deep learning techniques can be employed in market basket analysis to uncover complex patterns and relationships within transaction data. Here are some details on how deep learning can be used for market basket insights:

1. Neural Networks: Deep learning typically

involves neural networks. In the context of market basket analysis, you can use feedforward neural networks, recurrent neural networks (RNNs), or even more advanced architectures like convolutional neural networks (CNNs) for specific tasks.

2. Data Preparation: Convert your transaction data into a format suitable for deep learning. This may involve one-hot encoding for item representation or embedding layers to learn item representations directly from the data.

3. Sequential Data: In market basket analysis, the sequence of items in a transaction is important. RNNs and LSTM (Long Short-Term Memory) networks are well-suited for modeling sequences of items and capturing temporal dependencies.

4. Prediction Tasks: You can use deep learning for different tasks in market basket analysis, including:

a. Item Recommendations: Use neural networks to recommend additional items to customers based on their current selection, which can increase cross-selling opportunities.

b. Basket Completion: Predict which items are likely to be added to a customer's basket to help with inventory management and personalization.

c. Market Basket Segmentation: Employ deep learning to cluster customers based on their purchase history, allowing for more targeted marketing strategies.

5. Autoencoders: Deep learning autoencoders can be used for dimensionality reduction and anomaly detection. They can help identify rare but interesting patterns in market baskets.

6. Collaborative Filtering: Deep learning-based collaborative filtering techniques, such as matrix factorization and deep matrix factorization models, can be applied to identify item associations based on user-item interactions.

7. Hybrid Models: Combine traditional market basket analysis techniques like association rules with deep learning models for improved insights. For example, you can use association rules to identify frequent itemsets and then build a neural network to recommend items based on those frequent associations.

8. Evaluation Metrics: When using deep learning models, consider relevant evaluation metrics such as precision, recall, F1-score, and Mean Average Precision (MAP) to assess the quality of your recommendations or predictions.

9. Regularization and Hyperparameter Tuning: Deep learning models can be prone to overfitting, so apply regularization techniques and perform hyperparameter tuning to optimize model performance.

10. Real-time and Batch Processing: Depending on your specific needs, deploy your deep learning models for real-time recommendation systems or batch processing to generate periodic market basket insights reports.

11. Data Privacy and Security: Keep in mind that deep learning models may require sensitive customer transaction data. Ensure you have appropriate measures in place to protect data privacy and comply with regulations like GDPR.

Deep learning can be a powerful tool for market basket analysis, especially when dealing with large and complex datasets. However, it also comes with computational costs and data requirements, so consider the scalability and feasibility of implementing these techniques in your specific business context.

Advanced Association Analysis: Advanced association analysis in market basket insights involves using more sophisticated techniques and approaches

to uncover deeper and more valuable patterns within transaction data. Here are details on some advanced methods for association analysis in this context:

1. Sequential Pattern Mining:
   -Traditional association rules (like Apriori) focus on item co-occurrence, but sequential pattern mining considers the order of item purchases within transactions.
   - It's useful for understanding the sequence of events or item additions in a customer's shopping journey.

2. Temporal Association Analysis:
   - Incorporates time-related information into association rules, allowing you to identify trends and seasonality in customer purchases.
   - Useful for optimizing inventory

management and marketing strategies based on time-sensitive patterns.

3. Multi-Level Association Analysis:
   - Considers hierarchical or multi-level item relationships, where items are organized into categories or subcategories.
   - Helps to identify associations between items at different levels of granularity.

4. Lift and Leverage Analysis:
   - Instead of relying solely on support and confidence, use metrics like lift and leverage to evaluate the strength and uniqueness of associations.
   - Lift measures how much more often items are bought together compared to what would be expected by chance, while leverage quantifies the increase in the number of occurrences due to the association.

5. Contextual Association Analysis:
   - Consider contextual information, such as customer demographics or location, to discover associations that are specific to certain customer segments or store locations.
   - This can lead to more targeted marketing campaigns and product placement strategies.

6. Market Basket Segmentation:
   - Cluster customers into segments based on their purchase history and analyze associations within each segment.
   - This allows for personalized marketing and product recommendations tailored to different customer groups.

7. Advanced Data Mining Algorithms:
   - Utilize advanced data mining

techniques like FP-growth or Eclat, which are more efficient than the Apriori algorithm for large datasets.
   - These algorithms can help in finding frequent itemsets more quickly.

8. Graph-Based Association Analysis:
   - Represent transaction data as a graph where items are nodes and edges represent associations.
   - Graph-based approaches can reveal more complex relationships between items and are particularly useful for visualizing and exploring associations.

9. Hierarchical Clustering:
   - Apply hierarchical clustering techniques to group items or products that are frequently purchased together, creating a hierarchy of associations.

10. Cross-Domain Association Analysis:
    - Combine transaction data from multiple domains (e.g., online and offline sales) to discover cross-domain associations and gain a holistic view of customer behavior.

11. Anomaly Detection:
    - Use advanced anomaly detection techniques to identify rare or unexpected item associations that may be indicative of fraudulent activities or other unusual patterns.

12. A/B Testing and Experimentation:
    - Conduct controlled experiments to test the impact of specific changes in product placement, pricing, or marketing strategies on association patterns.

Advanced association analysis goes

beyond the basics and provides more valuable insights for businesses. It requires a deep understanding of the data, domain expertise, and the use of advanced analytics tools and algorithms to extract meaningful and actionable information from transaction data.

Visualization tools:
Visualization is a powerful tool in market basket insights as it helps to present complex patterns and associations in a more understandable and actionable way. Here are some details on visualization tools and techniques used in market basket analysis:

1. Sankey Diagrams:
   - Sankey diagrams visualize the flow of items from one category to another, making it easy to understand item

transitions in a customer's shopping journey.

2. Heatmaps:
   - Heatmaps represent item associations through color-coding, with darker colors indicating stronger associations.
   - They provide an at-a-glance view of frequently co-purchased items.

3. Item Co-Occurrence Matrix:
   - Visualize a matrix where rows and columns represent items, and cells display the frequency or strength of associations between item pairs.
   - It's an efficient way to identify patterns and relationships.

4. Network Graphs:
   - Use network graphs to show items as nodes and associations as edges.

- Graphs are great for displaying complex relationships between items in a visually appealing way.

5. Tree Maps:
- Tree maps can illustrate hierarchical relationships between items and item categories.
- They are useful for displaying associations within subcategories.

6. Bar Charts and Pie Charts:
- Simple bar and pie charts can be used to display item support, confidence, and lift values.
- They make it easy to compare and understand the significance of associations.

7. Market Basket Segmentation Visualizations:

- Use visualizations like scatter plots or parallel coordinate plots to showcase how customer segments (derived from clustering) interact with different items.
- It helps in targeted marketing strategies.

8. Time Series Plots:
- Plot item purchases over time to identify temporal patterns and seasonality.
- Time series visualizations are vital for understanding how item associations change with time.

9. Geospatial Maps:
- If your data includes location information, visualize item associations on geographical maps to identify location-specific patterns.
- This is particularly useful for businesses with multiple physical stores.

10. Word Clouds:
    - Create word clouds to display frequently co-purchased items in a visually appealing manner.
    - The size of each item's name corresponds to its frequency in associations.

11. 3D Visualization:
    - For more complex data, consider 3D visualization tools that allow you to explore associations in three dimensions.
    - This can provide a deeper perspective on item relationships.

12. Dashboard Tools:
    - Utilize dashboarding tools like Tableau, Power BI, or custom-built dashboards to combine various visualizations and create interactive, dynamic market basket insights dashboards.

13. Association Rule Visualization:
    - Some tools are specifically designed to visualize association rules generated from algorithms like Apriori or FP-growth.
    - They display rules in an easy-to-understand format, often with interactive filtering options.

14. Custom Data Visualization:
    - Depending on the unique needs of your business, you may need to create custom data visualizations to represent specific market basket insights.

Select the visualization tools and techniques that best suit your data and the insights you aim to convey. The goal is to make the patterns and associations in your market basket analysis results easily digestible and actionable for your team or

stakeholders.

```
# This Python 3 environment comes with many helpful analytics libraries ins
# It is defined by the kaggle/python Docker image: https://github.com/kaggl
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that
# You can also write temporary files to /kaggle/temp/, but they won't be sa
```

```
/kaggle/input/market-basket-analysis/Assignment-1_Data.xlsx
/kaggle/input/market-basket-analysis/Assignment-1_Data.csv
```

```
from matplotlib import pyplot as plt
df=pd.read_excel("/kaggle/input/market-basket-analysis/Assignment-1_Data.xl


df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 522064 entries, 0 to 522063
Data columns (total 7 columns):
 #   Column      Non-Null Count   Dtype         
---  ------      --------------   -----         
 0   BillNo      522064 non-null  object        
 1   Itemname    520609 non-null  object        
 2   Quantity    522064 non-null  int64         
 3   Date        522064 non-null  datetime64[ns]
 4   Price       522064 non-null  float64       
 5   CustomerID  388023 non-null  float64       
 6   Country     522064 non-null  object        
dtypes: datetime64[ns](1), float64(2), int64(1), object(3)
memory usage: 27.9+ MB
```

Step 1: Data Hygiene We're going to do the following steps:

1. Drop any rows where item name column is null.
2. Drop any rows where item quantity sold is 0 or less.
3. Fill missing customer IDs with a placeholder ID (99999)
4. Create a new column, Sumprice, that tells us total sales revenue (Quantity * Price) of the item

```
df.isnull().sum()
```

```
BillNo             0
Itemname        1455
Quantity           0
Date               0
Price              0
CustomerID    134041
Country            0
dtype: int64
```

```
#Dropping rows where ItemName isn't available
df.dropna(subset=["Itemname"],inplace=True)
#Dropping rows where Quantity <=0
df = df[df["Quantity"]>0]
df.isnull().sum()
```

```
BillNo             0
Itemname           0
Quantity           0
Date               0
Price              0
CustomerID    132113
Country            0
dtype: int64
```

```
#Fill missing customer IDs
df['CustomerID'].fillna(99999, inplace=True)
#Create SumPrice column
df["SumPrice"]=df["Quantity"]*df["Price"]
```

Step 2: EDA Let's explore the data for any insights. Let's find which countries sell the most items, and what items are the most popular in each country.

```
#Find the best selling items in each country
best_selling_items = df.groupby(['Country', 'Itemname']).agg({'Quantity': '
best_selling_items = best_selling_items.groupby('Country').apply(lambda x:
best_selling_items.sort_values("Quantity",ascending=False)
```

	Country	Itemname	Quantity
47	United Kingdom	PAPER CRAFT , LITTLE BIRDIE	80995
25	Netherlands	RABBIT NIGHT LIGHT	4801
12	France	RABBIT NIGHT LIGHT	4024
20	Japan	RABBIT NIGHT LIGHT	3408
0	Australia	MINI PAINT SET VINTAGE	2952
42	Sweden	MINI PAINT SET VINTAGE	2916
13	Germany	ROUND SNACK BOXES SET OF4 WOODLAND	1233
41	Spain	CHILDRENS CUTLERY POLKADOT PINK	729
43	Switzerland	PLASTERS IN TIN WOODLAND ANIMALS	639
26	Norway	SMALL FOLDING SCISSOR(POINTED EDGE)	576
3	Belgium	PACK OF 72 RETROSPOT CAKE CASES	480
40	Singapore	CHRISTMAS TREE PAINTED ZINC	384
1	Austria	SET 12 KIDS COLOUR CHALK STICKS	288
17	Iceland	ICE CREAM SUNDAE LIP GLOSS	240
19	Italy	FEATHER PEN,HOT PINK	240
29	Portugal	POLKADOT PEN	240
16	Hong Kong	ROUND SNACK BOXES SET OF4 WOODLAND	150
28	Poland	STRAWBERRY CERAMIC TRINKET BOX	144
27	Poland	CERAMIC CAKE DESIGN SPOTTED MUG	144
18	Israel	WOODLAND CHARLOTTE BAG	130
48	Unspecified	WORLD WAR 2 GLIDERS ASSTD DESIGNS	96
2	Bahrain	ICE CREAM SUNDAE LIP GLOSS	96
44	USA	SET 12 COLOURING PENCILS DOILY	88
24	Malta	GRAND CHOCOLATECANDLE	81
46	United Arab Emirates	BIG DOUGHNUT FRIDGE MAGNETS	72
45	United Arab Emirates	ASSORTED CHEESE FRIDGE MAGNETS	72

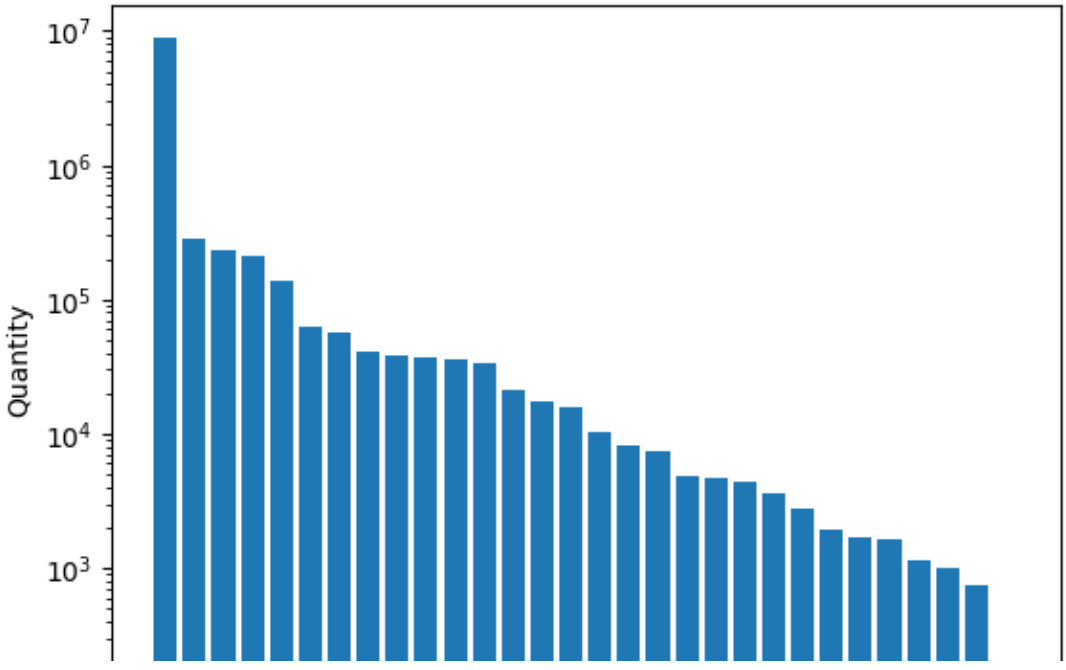
22	Lithuania	FELTCRAFT DOLL ROSIE	48
23	Lithuania	RED HARMONICA IN BOX	48
15	Greece	4 PEAR BOTANICAL DINNER CANDLES	48
14	Greece	4 LAVENDER BOTANICAL DINNER CANDLES	48
4	Brazil	DOLLY GIRL LUNCH BOX	24
5	Brazil	GREEN REGENCY TEACUP AND SAUCER	24
6	Brazil	PINK REGENCY TEACUP AND SAUCER	24
7	Brazil	ROSES REGENCY TEACUP AND SAUCER	24
21	Lebanon	ASSTD FRUIT+FLOWERS FRIDGE MAGNETS	24
8	Brazil	SET OF 4 PANTRY JELLY MOULDS	24
9	Brazil	SET OF 6 SPICE TINS PANTRY DESIGN	24
10	Brazil	SET/3 RED GINGHAM ROSE STORAGE BOX	24
11	Brazil	SMALL HEART FLOWERS HOOK	24
34	RSA	WOODEN BOX OF DOMINOES	12
36	Saudi Arabia	HOMEMADE JAM SCENTED CANDLES	12
37	Saudi Arabia	PLASTERS IN TIN CIRCUS PARADE	12
38	Saudi Arabia	PLASTERS IN TIN SKULLS	12
39	Saudi Arabia	PLASTERS IN TIN STRONGMAN	12
33	RSA	SET OF 20 KIDS COOKIE CUTTERS	12
32	RSA	PACK OF 6 BIRDY GIFT TAGS	12
31	RSA	ASSORTED BOTTLE TOP MAGNETS	12
30	RSA	4 TRADITIONAL SPINNING TOPS	12
35	Saudi Arabia	ASSORTED BOTTLE TOP MAGNETS	12

```
#Find the total sales by country.
total_sales_country = df.groupby(['Country']).agg({'SumPrice': 'sum'}).rese
total_sales_country = total_sales_country.sort_values('SumPrice', ascending
total_sales_country
```

	Country	SumPrice
0	United Kingdom	9003097.964
1	Netherlands	285446.340
2	Germany	228867.140
3	France	209715.110
4	Australia	138521.310
5	Spain	61577.110
6	Switzerland	57089.900
7	Belgium	41196.340
8	Sweden	38378.330
9	Japan	37416.370
10	Norway	36165.440
11	Portugal	33747.100
12	Singapore	21279.290
13	Italy	17483.240
14	Hong Kong	15691.800
15	Austria	10198.680
16	Israel	8135.260
17	Poland	7334.650
18	Greece	4760.520
19	Unspecified	4749.790
20	Iceland	4310.000
21	USA	3580.390
22	Malta	2725.590
23	United Arab Emirates	1902.280
24	Lebanon	1693.880

```
#Visualizing Total sales by country.
plt.bar(total_sales_country["Country"],total_sales_country["SumPrice"])
plt.yscale('log')
plt.ylabel('Quantity')
```

```
plt.xticks(rotation=90)
plt.show()
```



So far we've noticed that the UK has the most amount of sales and the most popular item sold in UK is 'PAPER CRAFT, LITTLE BIRDIE'. However, this outsells the most popular items in other countries by a large magnitude. Let's dig in by only looking at UK's grocery store data.

```
#Isolate the UK data and let's sort the most popular items in UK by quantit
only_uk = df[df["Country"]=="United Kingdom"]
only_uk.groupby("Itemname")["Quantity"].sum().sort_values(ascending=False)
```

```
Itemname
PAPER CRAFT , LITTLE BIRDIE           80995
MEDIUM CERAMIC TOP STORAGE JAR        77036
WORLD WAR 2 GLIDERS ASSTD DESIGNS     49526
JUMBO BAG RED RETROSPOT               44268
WHITE HANGING HEART T-LIGHT HOLDER    35744
                                      ...
HEN HOUSE W CHICK IN NEST                 1
BLACKCHRISTMAS TREE 30CM                  1
GOLD COSMETICS BAG WITH BUTTERFLY         1
WATERING CAN SINGLE HOOK PISTACHIO        1
*Boombox Ipod Classic                     1
Name: Quantity, Length: 4046, dtype: int64
```

```
#Let's find out what items, across the globe, bring in the most revenue.
total_sales_item = df.groupby(['Itemname']).agg({'Price': 'mean', 'Quantity

# Create a new column with the count of rows for each group
```

```
total_sales_item['Count'] = df.groupby(['Itemname']).size().values

# Sort the dataframe by 'SumPrice' column in descending order
total_sales_item = total_sales_item.sort_values("SumPrice", ascending=False

total_sales_item
```

	Itemname	Price	Quantity	SumPrice	Count
1060	DOTCOM POSTAGE	291.311822	708	206248.77	708
2386	PAPER CRAFT , LITTLE BIRDIE	2.080000	80995	168469.60	1
2848	REGENCY CAKESTAND 3 TIER	14.043347	13119	165689.19	1930
3840	WHITE HANGING HEART T-LIGHT HOLDER	3.220569	36527	102588.37	2269
2411	PARTY BUNTING	5.808664	17812	97367.48	1677
...	...	...	...	...	...
4025	allocate stock for dotcom orders ta	0.000000	4	0.00	1
4026	amazon	0.000000	161	0.00	8
4027	amazon adjust	0.000000	10	0.00	1
4028	amazon sales	0.000000	20	0.00	1
255	Adjust bad debt	-3687.353333	3	-11062.06	3

4056 rows × 5 columns

Interesting. We find out that the most sold item globally, 'PAPER CRAFT, LITTLE BIRDIE' was sold in just one transaction. Perhaps this was a large corporate

order. If we were to ever do a marketing or promotional push in the future, that required us to analyse our most popular products, this would be an anomaly that we would need to adjust for.

Step 3: EDA Market Basket Analysis using Apriori Algorithm and Association Rule Mining

1. Convert the Dataset into transactional format (Each row is one bill number with every item sold in that bill in a list)
2. Create a one-hot matrix of the products (Product sold = 1, Not sold = 0)
3. Merge the transactional matrix and the one hot matrix
4. Import the mlxtend library and perform association mining and generate association rules

```
#Convert the dataset into transactional format
transactions = df.groupby(['BillNo'])['Itemname'].apply(list)
transactions
```

```
BillNo
536365     [WHITE HANGING HEART T-LIGHT HOLDER, WHITE MET...
536366     [HAND WARMER UNION JACK, HAND WARMER RED POLKA...
536367     [ASSORTED COLOUR BIRD ORNAMENT, POPPY'S PLAYHO...
536368     [JAM MAKING SET WITH JARS, RED COAT RACK PARIS...
536369                              [BATH BUILDING BLOCK WORD]
                                  ...
581586     [LARGE CAKE STAND  HANGING STRAWBERY, SET OF 3...
581587     [CIRCUS PARADE LUNCH BOX, PLASTERS IN TIN CIRC...
A563185                                      [Adjust bad debt]
A563186                                      [Adjust bad debt]
A563187                                      [Adjust bad debt]
Name: Itemname, Length: 19735, dtype: object
```

```
#Create a one-hot matrix of the products
one_hot = pd.get_dummies(df['Itemname'])
one_hot
```

	*Boombox Ipod Classic	*USB Office Mirror Ball	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 DAISY PEGS IN WOOD BOX	12 EGG HOUSE PAINTED WOOD	HANG
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	
...	...	...	...	...	...	...	
522059	0	0	0	0	0	0	

```
#Add the BillNo column back to the one-hot encoded matrix
one_hot['BillNo']=df['BillNo']
one_hot
```

```
#Now, we group the One-Hot Matrix by BillNo and sum the values
one_hot = one_hot.groupby('BillNo').sum()
one_hot
```

BillNo	*Boombox Ipod Classic	*USB Office Mirror Ball	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 DAISY PEGS IN WOOD BOX	12 EGG HOUSE PAINTED WOOD	HANG H PAIN
536365	0	0	0	0	0	0	
536366	0	0	0	0	0	0	
536367	0	0	0	0	0	0	
536368	0	0	0	0	0	0	
536369	0	0	0	0	0	0	
...	...	...	...	...	...	...	
581586	0	0	0	0	0	0	
581587	0	0	0	0	0	0	
A563185	0	0	0	0	0	0	
A563186	0	0	0	0	0	0	
A563187	0	0	0	0	0	0	

19735 rows × 4056 columns

```
#Now, we merge the one-hot encoded matrix, with the transactional data
transaction_matrix = pd.merge(transactions, one_hot, on='BillNo')
transaction_matrix
```

BillNo	Itemname	*Boombox Ipod Classic	*USB Office Mirror Ball	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 DAISY PEGS IN WOOD BOX	P
536365	[WHITE HANGING HEART T-LIGHT HOLDER, WHITE MET...	0	0	0	0	0	
536366	[HAND WARMER UNION JACK, HAND WARMER RED POLKA...	0	0	0	0	0	
536367	[ASSORTED COLOUR BIRD ORNAMENT, POPPY'S PLAYHO...	0	0	0	0	0	
536368	[JAM MAKING SET WITH JARS, RED COAT RACK PARIS...	0	0	0	0	0	
536369	[BATH BUILDING BLOCK WORD]	0	0	0	0	0	
...	...	...	...	...	...	...	
581586	[LARGE CAKE STAND HANGING STRAWBERY, SET OF 3...	0	0	0	0	0	
	[CIRCUS PARADE						

581587	PARADE LUNCH BOX, PLASTERS IN TIN CIRC...	0	0	0	0	0

```
#Now we have to convert the product columns to 0s and 1s. We are converting
transaction_matrix[one_hot.columns[:-1]] = (transaction_matrix[one_hot.colu
transaction_matrix
```

Itemname	*Boombox Ipod Classic	*USB Office Mirror Ball	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	12 DAISY PEGS IN WOOD BOX	P
BillNo						
536365	[WHITE HANGING HEART T-LIGHT HOLDER, WHITE MET...	0	0	0	0	0
536366	[HAND WARMER UNION JACK, HAND WARMER RED POLKA...	0	0	0	0	0
	[ASSORTED					

```
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules
print(transaction_matrix.dtypes)
```

```
Itemname                         object
*Boombox Ipod Classic             int64
*USB Office Mirror Ball           int64
10 COLOUR SPACEBOY PEN            int64
12 COLOURED PARTY BALLOONS        int64
                                  ...
wrongly coded 20713               int64
wrongly coded 23343               int64
wrongly marked                    int64
wrongly marked 23343              int64
wrongly sold (22719) barcode      uint8
Length: 4057, dtype: object
```

```
transaction_matrix.iloc[:, 1:] = transaction_matrix.iloc[:, 1:].astype(bool
#Perform frequent itemset mining
frequent_itemsets = apriori(transaction_matrix.iloc[:, 1:], min_support=0.0
frequent_itemsets
```

	support	itemsets
0	0.015809	(10 COLOUR SPACEBOY PEN)
1	0.012567	(12 MESSAGE CARDS WITH ENVELOPES)
2	0.017887	(12 PENCIL SMALL TUBE WOODLAND)
3	0.018242	(12 PENCILS SMALL TUBE RED RETROSPOT)
4	0.017887	(12 PENCILS SMALL TUBE SKULL)
...	...	...
1891	0.011249	(JUMBO BAG RED RETROSPOT, JUMBO SHOPPER VINTAG...
1892	0.011249	(LUNCH BAG CARS BLUE, LUNCH BAG BLACK SKULL.,...
1893	0.010388	(LUNCH BAG CARS BLUE, LUNCH BAG BLACK SKULL.,...
1894	0.010286	(LUNCH BAG SUKI DESIGN, LUNCH BAG BLACK SKULL...
1895	0.010286	(CHARLOTTE BAG PINK POLKADOT, CHARLOTTE BAG SU...

1896 rows × 2 columns

```
# generate association rules
rules = association_rules(frequent_itemsets, metric="lift", min_threshold=1
rules
```

	antecedents	consequents	antecedent support	consequent support	support	c
0	(DOTCOM POSTAGE)	(6 RIBBONS RUSTIC CHARM)	0.035875	0.047732	0.010236	
1	(6 RIBBONS RUSTIC CHARM)	(DOTCOM POSTAGE)	0.047732	0.035875	0.010236	
2	(JAM MAKING SET PRINTED)	(6 RIBBONS RUSTIC CHARM)	0.056549	0.047732	0.011806	
3	(6 RIBBONS RUSTIC CHARM)	(JAM MAKING SET PRINTED)	0.047732	0.056549	0.011806	
4	(6 RIBBONS RUSTIC CHARM)	(JAM MAKING SET WITH JARS)	0.047732	0.055181	0.010337	
...	...	...	...	...	...	
3337	(CHARLOTTE BAG PINK POLKADOT)	(RED RETROSPOT CHARLOTTE BAG, STRAWBERRY CHARL...	0.037395	0.013073	0.010286	
3338	(CHARLOTTE BAG SUKI DESIGN)	(STRAWBERRY CHARLOTTE BAG, CHARLOTTE BAG PINK ...	0.044337	0.012212	0.010286	
	(STRAWBERRY	(RED RETROSPOT				

```
#Let's see the top 10 rules by lift
rules.sort_values('lift', ascending=False).head(10)
```

	antecedents	consequents	antecedent support	consequent support	support	c
2080	(HERB MARKER THYME)	(HERB MARKER PARSLEY, HERB MARKER ROSEMARY)	0.011806	0.010641	0.010134	
2077	(HERB MARKER PARSLEY, HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.010641	0.011806	0.010134	
2081	(HERB MARKER ROSEMARY)	(HERB MARKER PARSLEY, HERB MARKER THYME)	0.011857	0.010641	0.010134	
2076	(HERB MARKER PARSLEY, HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.010641	0.011857	0.010134	
534	(HERB MARKER THYME)	(HERB MARKER ROSEMARY)	0.011806	0.011857	0.010996	
535	(HERB MARKER ROSEMARY)	(HERB MARKER THYME)	0.011857	0.011806	0.010996	
2079	(HERB MARKER PARSLEY)	(HERB MARKER THYME, HERB MARKER ROSEMARY)	0.011756	0.010996	0.010134	
2078	(HERB MARKER THYME, HERB MARKER	(HERB MARKER PARSLEY)	0.010996	0.011756	0.010134	

```
import mpld3

# create scatter plot with x and y as lift and confidence values
fig, ax = plt.subplots()
```

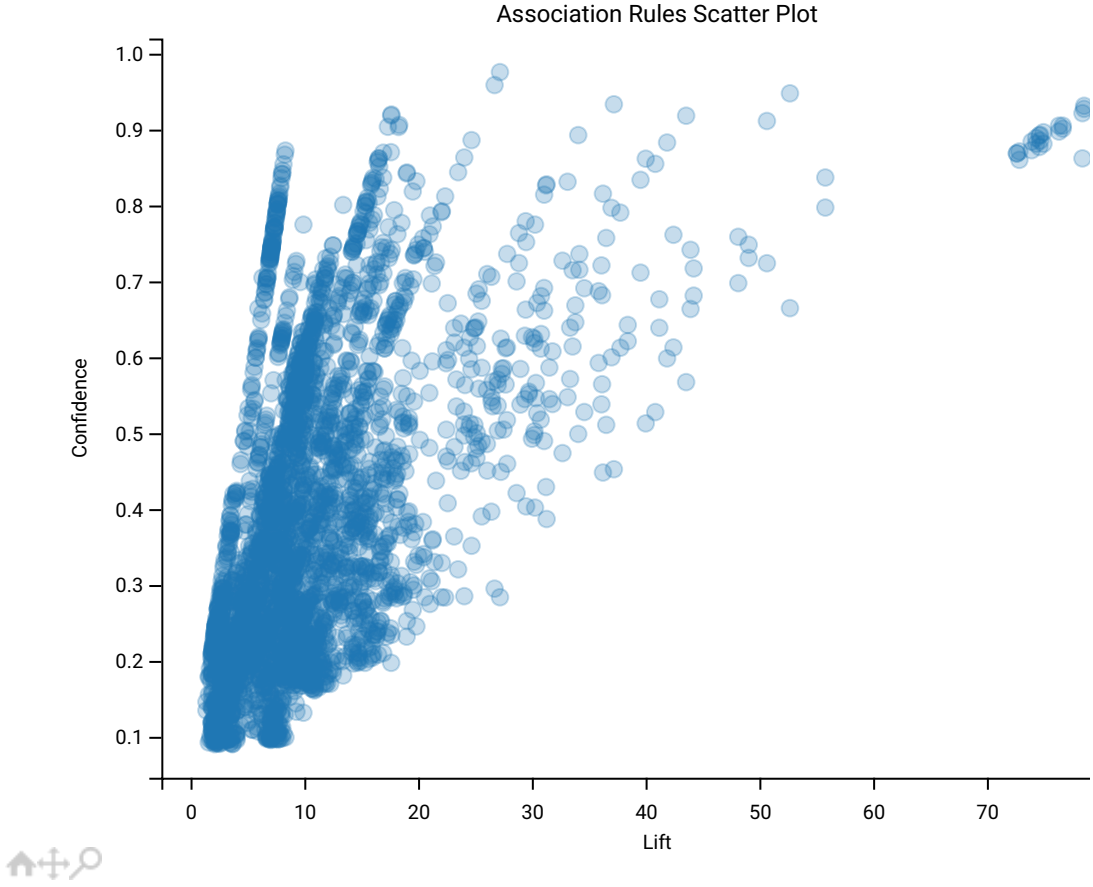
```
scatter = ax.scatter(rules['lift'], rules['confidence'], alpha=0.5)

# Define tooltips
tooltips = []
for i in range(len(rules)):
    rule = rules.iloc[i]
    tooltip = f"Rule: {rule['antecedents']} -> {rule['consequents']}\nSuppc
    tooltips.append(tooltip)

# Add tooltips to scatter plot using mpld3
mpld3.plugins.connect(fig, mpld3.plugins.PointHTMLTooltip(scatter, tooltips

# Set axis labels and title
ax.set_xlabel("Lift")
ax.set_ylabel("Confidence")
ax.set_title("Association Rules Scatter Plot")

# Show the plot
mpld3.display()
```




```
rules[(rules['lift'] > 40) & (rules['lift'] < 50)]
```

	antecedents	consequents	antecedent support	consequent support	support
124	(BLUE POLKADOT CUP)	(PINK POLKADOT CUP)	0.016418	0.015505	0.010489
125	(PINK POLKADOT CUP)	(BLUE POLKADOT CUP)	0.015505	0.016418	0.010489
264	(CHILDRENS CUTLERY SPACEBOY)	(CHILDRENS CUTLERY DOLLY GIRL)	0.017938	0.014441	0.010996
265	(CHILDRENS CUTLERY DOLLY GIRL)	(CHILDRENS CUTLERY SPACEBOY)	0.014441	0.017938	0.010996
552	(JAM JAR WITH PINK LID)	(JAM JAR WITH GREEN LID)	0.016874	0.015100	0.011198
553	(JAM JAR WITH GREEN LID)	(JAM JAR WITH PINK LID)	0.015100	0.016874	0.011198
1556	(REGENCY SUGAR BOWL GREEN)	(REGENCY MILK JUG PINK)	0.014897	0.015252	0.011148
1557	(REGENCY MILK JUG PINK)	(REGENCY SUGAR BOWL GREEN)	0.015252	0.014897	0.011148
1568	(REGENCY TEA PLATE ROSES)	(REGENCY TEA PLATE PINK)	0.021079	0.014289	0.012617
1569	(REGENCY TEA PLATE PINK)	(REGENCY TEA PLATE ROSES)	0.014289	0.021079	0.012617
1616	(SET OF 3 WOODEN TREE DECORATIONS)	(SET OF 3 WOODEN STOCKING DECORATION)	0.014492	0.015759	0.010996
1617	(SET OF 3 WOODEN STOCKING DECORATION)	(SET OF 3 WOODEN TREE DECORATIONS)	0.015759	0.014492	0.010996
3024	(POPPY'S PLAYHOUSE LIVINGROOM, POPPY'S	(POPPY'S PLAYHOUSE BEDROOM)	0.012921	0.020927	0.011046

	PLAYHOU...				
3025	(POPPY'S PLAYHOUSE LIVINGROOM, POPPY'S PLAYHOU...	(POPPY'S PLAYHOUSE KITCHEN)	0.012820	0.021535	0.011046
3026	(POPPY'S PLAYHOUSE KITCHEN, POPPY'S PLAYHOUSE ...	(POPPY'S PLAYHOUSE LIVINGROOM)	0.015404	0.016215	0.011046
3027	(POPPY'S PLAYHOUSE LIVINGROOM)	(POPPY'S PLAYHOUSE KITCHEN, POPPY'S PLAYHOUSE ...	0.016215	0.015404	0.011046
3028	(POPPY'S PLAYHOUSE KITCHEN)	(POPPY'S PLAYHOUSE LIVINGROOM, POPPY'S PLAYHOU...	0.021535	0.012820	0.011046
3029	(POPPY'S PLAYHOUSE BEDROOM)	(POPPY'S PLAYHOUSE LIVINGROOM, POPPY'S PLAYHOU...	0.020927	0.012921	0.011046
3038	(REGENCY TEA PLATE GREEN, REGENCY TEA PLATE PINK)	(REGENCY TEA PLATE ROSES)	0.013023	0.021079	0.011958
3039	(REGENCY TEA PLATE ROSES)	(REGENCY TEA PLATE GREEN, REGENCY TEA PLATE PINK)	0.021079	0.013023	0.011958