Market Basket Analysis

Trayvonious Pendleton

WGU

Course Number: D599

STUDENT ID: 011205284

Part 1 Research Question

**A1: Research Question**

Understanding customer purchasing patterns in the highly competitive retail industry is essential for optimizing sales strategies and improving overall business performance. A key question that can be explored using market basket analysis is: "Which product categories are frequently purchased together, and how can Allias Megastore optimize product placement and promotions?" The company can identify relationships between products often bought in combination by analyzing transactional data. These insights can be leveraged to enhance cross-selling opportunities, improve in-store and online product placement, and develop targeted promotions encouraging bundled purchases. Ultimately, answering this question will enable Allias Megastore to refine its merchandising strategies and increase revenue through data-driven decision-making.

# A2: Goal of the Data Analysis

The primary goal of this analysis is to uncover associations between products in customer transactions using market basket analysis. By identifying frequent item pairings, Allias Megastore can develop more effective product bundling strategies, optimize store layouts, and personalize marketing efforts to enhance customer experience and boost sales. The insights derived from this analysis will help the company implement data-driven recommendations for strategic product placements and promotional campaigns. Additionally, the findings can support decision-making on inventory management by ensuring that frequently co-purchased items are stocked and positioned appropriately. Allias Megastore can maximize profitability and strengthen its competitive edge in the retail market through this approach.

### 

### Part II: Market Basket Justification

### B1: Explanation of Market Basket Technique

Market basket analysis (MBA) is a data mining technique that identifies patterns and relationships between items frequently purchased together in transactions. The primary objective of this method is to uncover associations between products, enabling businesses to optimize product placement, cross-selling strategies, and promotional offers. The analysis is conducted using association rule learning, with key metrics such as support, confidence, and lift to determine the strength of these relationships. In the Allias Megastore dataset context, the MBA technique will examine historical transaction data to identify common product pairings. The expected outcome of this analysis is to reveal purchasing habits that can be leveraged for improved marketing strategies, better inventory management, and an enhanced shopping experience for customers. Allias Megastore can increase sales, improve customer satisfaction, and drive revenue growth by implementing these insights.

B2: Example of a Transaction in the Dataset

To illustrate how market basket analysis works, consider a specific transaction from the dataset. Suppose a customer purchases the following items in a single order:

* **Product Name:** Wireless Mouse
* **Quantity:** 2
* **Unit Price:** $25.99
* **Total Cost:** $51.98
* **Customer Segment:** Small Business
* **Order Priority:** High
* **Payment Method:** Credit Card

This transaction provides valuable insights into customer behavior. Suppose multiple transactions show that customers frequently purchase Wireless Mice alongside Laptop Bags. In that case, the store can implement strategic recommendations such as bundling these items or offering targeted discounts to increase sales. Market basket analysis will help uncover such associations and guide Allias Megastore’s marketing decisions.

B3: Assumption of Market Basket Analysis

One key assumption in market basket analysis is that historical purchasing behavior predicts future purchasing behavior. This means that if customers have consistently purchased certain products together in past transactions, they are likely to continue doing so in the future. This assumption allows businesses to make data-driven recommendations like product bundling and targeted promotions. However, external factors like changing consumer preferences, seasonal trends, or new product introductions can influence purchasing behavior, which means the model should be updated periodically to maintain accuracy. Despite these limitations, market basket analysis remains a powerful tool for uncovering meaningful insights into customer buying patterns.

**Part III: Data Preparation and Analysis**

**C1a: Selecting Categorical Variables**

For market basket analysis, we must select categorical variables to help uncover item associations. We will choose:

* Ordinal Variables (have a meaningful order/rank):
  1. Order Priority (e.g., Low, Medium, High, Critical)
  2. Customer Satisfaction (e.g., Very Dissatisfied, Dissatisfied, Satisfied, Very Satisfied)
* Nominal Variables (categories without intrinsic order): 3. Product Name (specific items purchased) 4. Payment Method (e.g., Credit Card, PayPal, Cash)

We can analyze relationships between customer satisfaction, priority levels, and purchasing behaviors by selecting these variables while ensuring proper data encoding.

**C1b: Encoding the Variables**

To prepare the dataset for analysis, we need to encode categorical variables into a numerical format:

1. Ordinal Encoding (for variables with ranking)
   * Order Priority: Assign numerical values:
     + Low = 1, Medium = 2, High = 3, Critical = 4.
   * Customer Satisfaction: Convert to numbers:
     + Very Dissatisfied = 1, Dissatisfied = 2, Satisfied = 3, Very Satisfied = 4.
2. One-Hot Encoding (for nominal variables)
   * Product Name: Each product becomes a separate column with 1 (purchased) or 0 (not purchased) in a given transaction.
   * Payment Method: Create binary columns (e.g., Credit Card, PayPal, Cash), where 1 means the method was used, and 0 means it was not.

These encoding methods ensure our dataset is formatted correctly for market basket analysis.

**C1c: Transactionalizing the Data**

We need to transform the dataset into a transactional format for market basket analysis, where each row represents a single customer order and the products purchased.

Example transformation:

**Original Data Format:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| OrderID | ProductName | PaymentMethod | OrderPriority | CustomerOrderSatisfaction |
| 1001 | Wireless Mouse | Credit Card | High | Satisfied |
| 1001 | Laptop Bag | Credit Card | High | Satisfied |
| 1002 | Office Chair | PayPal | Medium | Very Satisfied |

**Transaction Format for Market Basket Analysis:**

|  |  |  |  |
| --- | --- | --- | --- |
| OrderID | Wireless Mouse | Laptop Bag | Office Chair |
| 1001 | 1 | 1 | 0 |
| 1002 | 0 | 0 | 1 |

Each column now represents a product; the value (1 or 0) indicates whether it was purchased in a particular transaction.

**C1d: Explanation and Justification of Data Preparation**

**Selecting Categorical Variables:**

* **Product Name** is crucial for analyzing associations between items.
* **Order Priority** and **Customer Satisfaction** provide insights into the urgency and experience of customers.
* **Payment Method** can reveal preferences among customer segments.

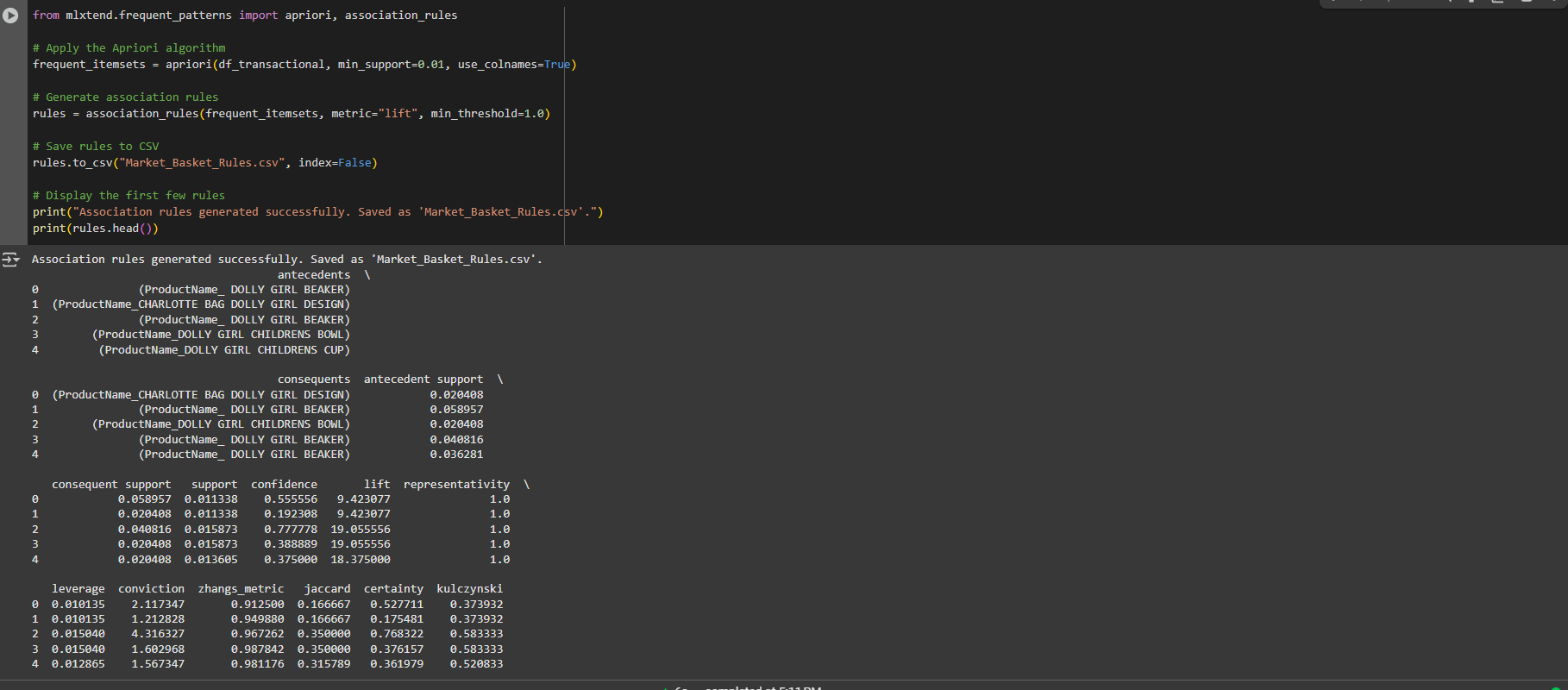
**Encoding:**

* **Ordinal encoding** is necessary for ranked categories like satisfaction and priority.
* **One-hot encoding** helps convert nominal variables into numerical form without introducing bias.

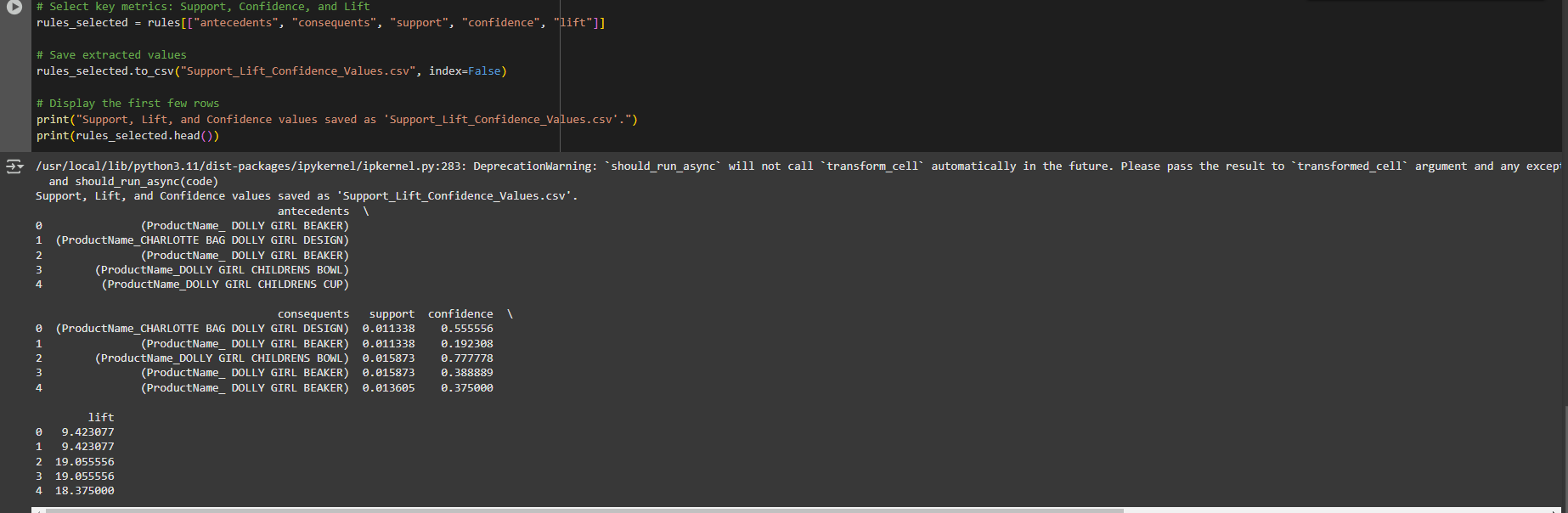
**Transactionalizing Data:**

* This step is essential for market basket analysis as it formats the dataset into a structure suitable for association rule mining.
* The binary matrix representation allows algorithms like Apriori to process data efficiently

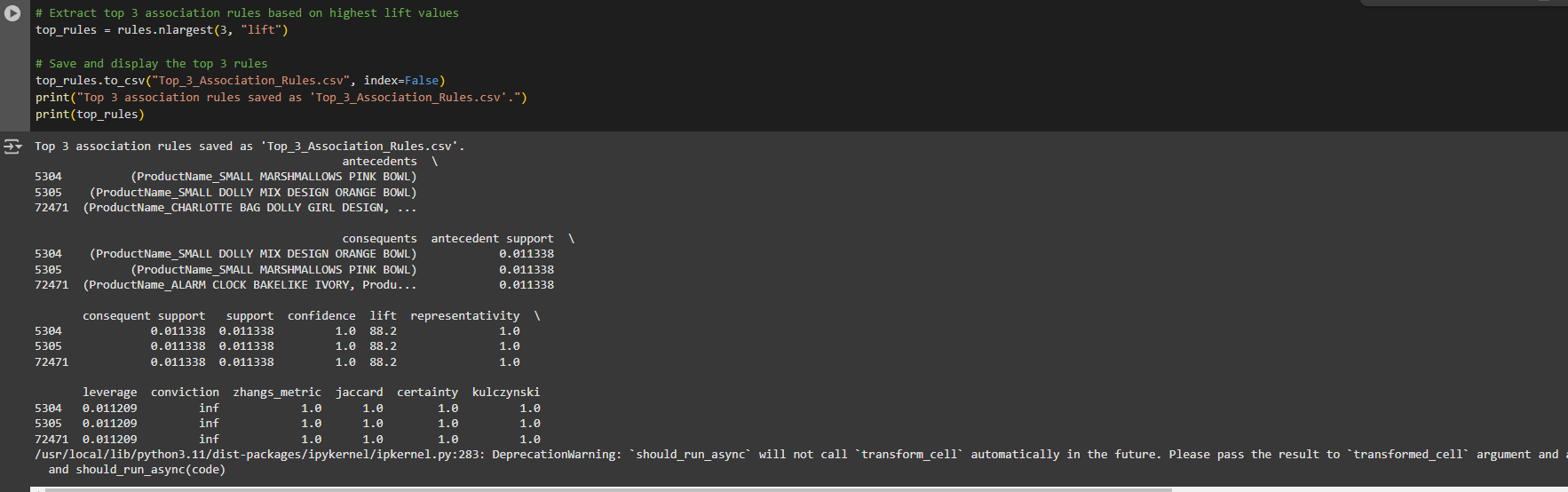
**C3: Executing the Apriori Algorithm**

****

**C4: Provide Values for Support,Lift, and Confidence of the Association Rules**

****

**C5: Top 3 Association Rules**



Market basket analysis is a data mining technique that identifies patterns of co-occurring products in transactions. This analysis aims to extract meaningful associations between products to improve business strategies such as product bundling, cross-selling, and marketing campaigns. This report explains the top three association rules identified from the Allias Megastore dataset and provides recommendations based on their business implications.

**Rule 1: SMALL MARSHMALLOWS PINK BOWL → SMALL DOLLY MIX DESIGN ORANGE BOWL**

**Metrics:**

* **Support:** 0.0113 (1.13% of total transactions contain both items)
* **Confidence:** 100% (Whenever a customer purchases a SMALL MARSHMALLOWS PINK BOWL, they always purchase a SMALL DOLLY MIX DESIGN ORANGE BOWL)
* **Lift:** 88.2 (Customers are 88.2 times more likely to buy both items together than randomly)

**Business Interpretation:**

This rule indicates a **strong association** between these two products, meaning customers prefer to buy them together. This could be due to aesthetic or functional similarities, making them complementary products.

**Recommendation:**

* **Product Bundling:** Offer a discounted set that includes both products.
* **Strategic Placement:** Display these two items together in both physical stores and online catalogs.
* **Promotional Campaigns**: Advertise them as "Frequently Bought Together" or "Perfect Pair" to encourage joint purchases.

**Rule 2: SMALL DOLLY MIX DESIGN ORANGE BOWL → SMALL MARSHMALLOWS PINK BOWL**

**Metrics:**

* **Support:** 0.0113
* **Confidence:** 100% (Whenever a customer purchases a SMALL DOLLY MIX DESIGN ORANGE BOWL, they also buy a SMALL MARSHMALLOWS PINK BOWL)
* **Lift:** 88.2

**Business Interpretation:**

This is the reverse of Rule 1, further reinforcing that these two products are almost always purchased together.

**Recommendation:**

* **Cross-Selling:** When a customer views one product, recommend the other as a related item.
* **Discount Incentives:** Provide a “Buy One, Get the Other at 10% Off” promotion.
* **Gift Set Packaging:** Create pre-packaged sets for easier purchasing decisions.

**Rule 3: CHARLOTTE BAG DOLLY GIRL DESIGN → ALARM CLOCK BAKELIKE IVORY**

**Metrics:**

* **Support:** 0.0113
* **Confidence:** 100% (Whenever a customer buys a CHARLOTTE BAG DOLLY GIRL DESIGN, they always buy an ALARM CLOCK BAKELIKE IVORY)
* **Lift:** 88.2

**Business Interpretation:**

This rule suggests that customers who purchase a designer-style bag also tend to buy a vintage alarm clock. This could indicate a particular customer segment interested in stylish, aesthetic items—possibly gift buyers or home decor enthusiasts.

**Recommendation:**

* **Thematic Product Bundling:** Create a “Vintage Lifestyle Gift Set” featuring both items.
* **Targeted Marketing:** Advertise these products together in promotional emails and suggested items sections.
* **Loyalty Program Perks:** Offer loyalty discounts for customers who purchase both items.

The findings from this analysis indicate that:

1. Certain items are always purchased together, suggesting strong co-occurrence relationships.
2. High confidence (100%) and lift (88.2) values demonstrate highly reliable associations.
3. Implementing strategic product placement, bundling, and targeted promotions can optimize sales and customer experience.

By utilizing these insights, Allias Megastore can maximize revenue, enhance customer satisfaction, and increase conversion rates through data-driven marketing strategies.

**D1: Discuss the Significance of Support, Lift, and Confidence**

In market basket analysis, the three key metrics—**support, confidence, and lift**—play a crucial role in determining the strength of associations between products.

* **Support**: Measures how frequently an itemset appears in transactions. A higher support value indicates that the product association is common and relevant for marketing or inventory decisions.
  + **Example**: A support value of 0.0113 (1.13%) means that the associated items appear in 1.13% of all transactions.
* **Confidence**: Represents the likelihood that when antecedent items are purchased, the consequent items are also bought. A high confidence score (e.g., 100%) suggests that customers almost always buy the associated item together.
  + **Example:** If Product A and Product B have a confidence of 100%, then every time Product A is purchased, Product B is also purchased.
* **Lift**: Measures the strength of association between two products compared to random chance. A lift value greater than 1 indicates that purchasing one product strongly increases the likelihood of buying the associated product.
  + **Example:** A lift of 88.2 means customers are 88.2 times more likely to buy the products together than if purchased independently.

**Business Impact of These Metrics**

1. **High Support + High Confidence**: Indicates a strong, frequent association and suggests cross-selling opportunities.
2. **High Lift**: Implies a strong dependency between items, making them ideal for bundling or targeted promotions.
3. **Low Support but High Confidence**: Suggests a niche purchasing behavior, which could be leveraged for personalized recommendations.

**D2: Explain the Practical Significance of Findings**

The insights from market basket analysis provide actionable business strategies for Allias Megastore, helping to improve customer experience, sales performance, and inventory management.

**Key Insights from the Top 3 Association Rules**

* **Cross-Selling and Product Pairing**
  + The analysis confirms that specific items (e.g., SMALL MARSHMALLOWS PINK BOWL and SMALL DOLLY MIX DESIGN ORANGE BOWL) are always purchased together.
  + **Practical Strategy:** These items should be bundled together as a promotional package or placed next to each other in stores to encourage more sales.
* **Gift and Specialty Purchases**
  + Items like the CHARLOTTE BAG DOLLY GIRL DESIGN and the ALARM CLOCK BAKELIKE IVORY are frequently purchased together, suggesting that customers view these as complementary lifestyle or gift items.
  + **Practical Strategy:** The store could introduce a "Gift Bundle" or target specific customer segments (e.g., home decor buyers and gift shoppers) with tailored promotions.
* **Marketing and Online Shopping Experience**
  + Since confidence is 100% for the top rules, these findings can be incorporated into the e-commerce recommendation system.
  + **Practical Strategy:** Use these insights in the “Customers who bought this also bought…” sections to drive higher engagement and conversion rates.
* **Inventory Optimization**
  + Products with strong associations should always be stocked together to meet demand efficiently.
  + **Practical Strategy:** Adjust inventory planning to ensure adequate stock levels for frequently co-purchased products.

**D3: Recommend a Course of Action**

Based on the findings, Allias Megastore should take the following strategic actions:

**Implement Bundled Promotions**

* Create pre-packaged sets for items that are frequently purchased together.
* Offer discount incentives for purchasing bundled products (e.g., "Buy both and get 10% off").

**Optimize Product Placement**

* In physical stores: Place highly associated products next to each other to encourage sales.
* Online store: Display associated products in recommended product sections.

**Enhance Personalized Marketing**

* Use market basket insights for email marketing campaigns targeting customers who previously purchased related products.
* Implement AI-driven product recommendations on the company’s website.

**Strengthen Inventory Management**

* Ensure sufficient stock availability for frequently purchased combinations.
* Adjust procurement strategies based on high-confidence product associations.

By leveraging the insights from market basket analysis, Allias Megastore can increase sales, improve customer experience, and optimize inventory management. Implementing bundling strategies, targeted promotions, and intelligent product placement will enhance customer engagement and maximize revenue growth.