**Market Basket Analysis for Cross-Selling**

**AIM**

To Analyse customer behaviour, identify cross-selling opportunities, and assess marketing campaign effectiveness. Explore various aspects of customer behaviour and marketing effectiveness using the given data set understanding customer demographics, purchase history, and response to marketing campaigns.

**FEATURES:**

* **Customer Age:** Numerical (continuous)
* **Customer Gender:** Categorical (Male, Female)
* **Transaction Amount:** Numerical (continuous)
* **Average Order Value:** Numerical (continuous)
* **Cross-Sell Conversion Rate:** Numerical (percentage)
* **Co-Purchase Frequency:** Numerical (count or frequency)
* **Browsing Behavior:** Categorical (High, Medium, Low)
* **Customer Lifetime Value:** Numerical (continuous)
* **Product Return Rate:** Numerical (percentage)
* **Loyalty Score:** Numerical (continuous)
* **Customer Feedback:** Categorical (Positive, Neutral, Negative)

**METHODOLOGY:**

1. **Data Cleaning and Preparation:**
   * Checked for data consistency using heat map
   * Interpolate missing values in the 'TransactionID' column
   * Filled missing values in numerical columns with the mean
   * Filled missing values in categorical columns with the mode
   * Using statistical mean in Python.
   * Imported the data into Power BI.
   * Transformed the data by using necessary functions like transpose, use first row as header.
   * Created necessary calculated columns or measures.
2. **Data Visualization:**
   * Created visuals to represent the key metrics and trends.
   * Used appropriate chart types to effectively communicate the data.

**OUTPUT:**

A screenshot of a computer

Description automatically generated

**RESULT ANALYSIS**

* Correlation Matrix:

1. **CoPurchaseFreq**: Positively correlated with **ProductAffinity** and **CrossSellConversionRate**, indicating that customers who frequently co-purchase are likely to have a higher affinity for products and a higher cross-sell conversion rate.
2. **ProductAffinity**: Shows positive correlations with **TransactionAmount** and **CrossSellConversionRate**, suggesting that customers with a higher affinity for products tend to spend more and convert more on cross-sells.
3. **TransactionAmount**: Has a positive correlation with **AvgOrderValueIncrease**, meaning higher transaction amounts are associated with an increase in average order value.
4. **PurchaseSequence**: Positively correlated with **CustomerAge** and **LoyaltyScore**, indicating that customers who purchase more frequently tend to be older and have higher loyalty scores.
5. **CustomerAge**: Shows positive correlations with **LoyaltyScore** and **TimeBetweenPurchases**, suggesting that older customers are more loyal and have longer intervals between purchases.
6. **CrossSellConversionRate**: Positively correlated with **AvgOrderValueIncrease** and **Confidence**, indicating that higher cross-sell conversion rates are associated with higher average order values and customer confidence.
7. **AvgOrderValueIncrease**: Positively correlated with **LoyaltyScore** and **Confidence**, suggesting that an increase in average order value is associated with higher loyalty and confidence.
8. **Support**: Shows positive correlations with **Confidence** and **LoyaltyScore**, indicating that customers who receive more support tend to be more confident and loyal.
9. **Confidence**: Positively correlated with **LoyaltyScore** and **TimeBetweenPurchases**, suggesting that more confident customers are more loyal and have longer intervals between purchases.
10. **LoyaltyScore**: Positively correlated with **TimeBetweenPurchases** and **DiscountUsage**, indicating that more loyal customers tend to have longer intervals between purchases and use discounts more frequently.
11. **TimeBetweenPurchases**: Shows positive correlations with **DiscountUsage** and **ProductReturnRate**, suggesting that customers who take longer between purchases tend to use discounts more and have higher return rates.
12. **DiscountUsage**: Positively correlated with **ProductReturnRate** and **CustomerLifetimeValue**, indicating that customers who use discounts more frequently tend to return products more and have higher lifetime values.
13. **ProductReturnRate**: Positively correlated with **CustomerLifetimeValue**, suggesting that customers who return products more frequently tend to have higher lifetime values.
14. **CustomerLifetimeValue**: Positively correlated with **Lift**, indicating that customers with higher lifetime values tend to have higher lift (a measure of the effectiveness of a marketing campaign).
15. **TransactionID**: This variable has low correlations with most other variables, indicating it doesn’t strongly relate to other factors in the dataset.
16. **CustomerID**: Shows a slight positive correlation with **CoPurchaseFreq** and **LoyaltyScore**, suggesting that customers with higher IDs might have higher co-purchase frequencies and loyalty scores.

* Box Plot Analysis

**Median Transaction Amount:** The median transaction amount is approximately around $100. This indicates that half of the transactions fall below this value and the other half above.

**Distribution:** The box plot suggests a right-skewed distribution of transaction amounts. The median is closer to the lower quartile, and the upper whisker is longer than the lower whisker. This implies that there are more smaller transactions and fewer but larger transactions.

**Outliers:** Several data points are plotted as individual points beyond the upper whisker, indicating these are potential outliers or unusually high transaction amounts. These outliers might warrant further investigation to understand their nature and impact on overall analysis.

**Range:** The range of transaction amounts is substantial, extending from around $0 to approximately $5000. This wide range suggests significant variability in purchase behaviour among customers.

* **Average Order Value:**

The average order value is $17.49, indicating room for improvement through cross-selling or upselling strategies.

* **Cross-Sell Conversion Rate:**

At 35.07%, the conversion rate is moderate and could be improved with targeted recommendations.

* **Co-Purchase Frequency:**

The distribution of co-purchase frequency is relatively normal, with a peak around the mid-range.

* **Transaction Amount:**

The transaction amount distribution shows variability, with some high-value transactions. Identifying patterns in these transactions could be beneficial.

* **Customer Lifetime Value:**

The customer lifetime value distribution indicates a range of customer values, with some high-value customers. Focusing on retaining and increasing the value of these customers is crucial.

* **Product Return Rate:**

The product return rate of 5.34% is relatively low, suggesting good product quality or effective return policies.

* **Customer Age:**

The age distribution is relatively even, with a slight skew towards older customers.

* **Customer Gender:**

The distribution of male and female customers appears to be relatively balanced.

* **Customer Feedback:**

A majority of customers have provided positive feedback, indicating overall satisfaction.

* **Browsing Behaviour:**

There's a relatively high percentage of customers with high browsing behaviour, suggesting potential for increased engagement and purchases.

**Summary of the Analysis:**

Improve conversion rate with targeted recommendations.

To Maintain the majority customer’s positive feedback for overall satisfaction.

There are more smaller transactions and fewer but larger transactions.

Increasing Average order value through cross-selling or upselling strategies.

Focusing on retaining and increasing the value of these customers is crucial.