

Task 1: Exploratory Data Analysis (EDA) Business Report

Objective:

The purpose of this report is to provide insights into the dataset using Exploratory Data Analysis (EDA). EDA helps identify trends, patterns, relationships, and anomalies within the data, which can inform business strategies for optimizing sales, customer engagement, and product offerings.

Overview of Dataset:

The dataset used in this analysis includes attributes related to transactions, products, customers, and dates. Key variables include product price, quantity purchased, total transaction value, product IDs, and the number of days since the customer signed up. By examining these attributes, we aim to uncover patterns in purchasing behavior, pricing dynamics, and customer activity.

Key Findings from EDA:

1. Distribution of Price:

The price distribution of products reveals a large concentration of low-priced items with some high-priced outliers. This suggests that while the majority of sales occur within affordable product categories, high-value items still contribute significantly to total revenue. The business could consider introducing promotional offers for higher-priced items or offering premium versions of popular products to increase high-ticket sales.

2. Quantity Purchased:

The quantity of products purchased in each transaction shows a skewed distribution, with most customers purchasing one to four items per transaction. However, there are occasional instances where customers buy larger quantities. This indicates potential for bulk purchase strategies. The business could implement incentives like bulk discounts or bundle offers to encourage customers to buy more per transaction, which could increase the average revenue per customer.

3. Transaction Value Distribution:

Transaction value analysis indicates that the majority of transactions have a lower total value, suggesting that while customers make frequent purchases, they tend to buy lower-cost items or smaller quantities. However, a few high-value transactions stand out, highlighting the potential for high-revenue sales. The business can explore ways to target customers for higher-value transactions through personalized offers, loyalty programs, or premium product bundles.

4. Days Since Customer Signup:

Most customers in the dataset have signed up within the last 400 days. A smaller number of customers have been signed up for longer periods, indicating that the business has a good influx of new customers. However, this also signals an opportunity to focus on retaining these new customers and converting them into loyal buyers. Strategies such as email marketing, special offers for long-term customers, and loyalty programs could help in improving retention rates.

5. Product Popularity:

The frequency of product purchases varies greatly, with certain products being purchased far more frequently than others. A few products dominate the sales, suggesting strong customer demand. This insight highlights the need for targeted marketing strategies to further promote these popular products. Additionally, less popular products may need to be reevaluated—considering promotional strategies, product bundling, or even discontinuation if they don't meet sales targets.

6. Price vs. Total Transaction Value:

The correlation between price and transaction value is moderately positive, indicating that higher-priced items tend to lead to higher transaction values. This suggests that while most customers purchase low-priced items, the business could further capitalize on this correlation by introducing more high-priced options or increasing the visibility of high-value products in marketing campaigns.

7. Quantity vs. Total Transaction Value:

There is a positive correlation between the quantity of items purchased and total transaction value. This emphasizes the potential to boost sales by encouraging customers to purchase more items per transaction. The business could implement targeted promotions, such as “buy one, get one free” or discounts for larger quantities, to take advantage of this relationship.

8. Outliers:

The dataset contains a few outliers, particularly in the total value and price distributions. While these outliers may represent irregular purchases or exceptional customer behavior, they can provide valuable insights. For instance, high-value outliers may represent premium customer segments or special bulk transactions, which can be studied further to identify strategies for attracting similar customers.

Correlations:

- **Price and Quantity:** A very weak negative correlation exists between price and quantity purchased, suggesting that customers do not necessarily buy more when prices are lower. This highlights the need for effective product differentiation and marketing, as customers are likely to focus on product appeal and need rather than just price.

- **Price and Total Value:** The strong positive correlation between price and total value indicates that higher-priced items are more likely to result in higher total transaction values. This suggests that promoting higher-priced items or introducing tiered pricing strategies could boost overall revenue.
- **Quantity and Total Value:** The positive correlation between quantity and total value suggests that incentivizing customers to buy more products in a single transaction can increase the total transaction value. This can be leveraged through volume discounts, bundle deals, or subscription-based offers.

Recommendations:

1. **Promote Popular Products:**
The business should focus on promoting the most frequently purchased products through targeted advertising, offers, or loyalty programs. This will help increase sales volume and customer engagement.
2. **Encourage Bulk Purchases:**
Since there is a positive correlation between quantity and transaction value, encouraging bulk purchases through discounts or bundling can significantly increase sales. Special offers like "buy more, save more" should be considered to incentivize larger transactions.
3. **Capitalize on High-Value Products:**
High-priced products contribute significantly to total revenue. The business should explore strategies to enhance the visibility and attractiveness of these products. Premium marketing campaigns or loyalty programs targeting high-spending customers could be beneficial.
4. **Customer Retention Strategies:**
The majority of customers are relatively new, so the business must focus on retaining these customers. Implementing customer loyalty programs, special offers for returning customers, or regular follow-up communication can help convert new customers into loyal ones.
5. **Product Diversification:**
Based on the frequency analysis, the business could evaluate whether it is worthwhile to introduce additional variants of underperforming products or discontinue them if they fail to meet the expected sales targets.
6. **Data-Driven Decision Making:**
The insights from this analysis should guide data-driven decision-making for future business strategies. Regularly conducting EDA will ensure that the company stays aligned with customer preferences and adjusts its offerings accordingly.

Conclusion:

The findings from this EDA reveal valuable insights into customer behavior, product preferences, and transaction dynamics. By applying these insights to business strategies such as pricing, marketing, product bundling, and customer retention, the company can enhance

customer satisfaction and maximize revenue growth. The business should continue to use data analytics as a core part of decision-making to stay competitive in the marketplace.