Project Synopsis: BigBasket Analysis

Submitted by- Kunal Yogesh Dhamal

Title:

BigBasket Analysis Using Python

Introduction:

This project analyzes Bigbasket, a leading online grocery platform in India, leveraging transaction data to uncover insights that can optimize its offerings and enhance customer satisfaction. As the e-commerce sector grows, Bigbasket plays a crucial role in delivering groceries and essentials to a diverse customer base. With the rise of online shopping driven by convenience, especially during the pandemic, the online grocery industry has seen rapid growth. This project explores consumer preferences, product trends, and payment methods, providing valuable insights that can help Bigbasket refine its strategies and stay competitive in this dynamic industry.

Objectives :

1. To analyze sales data to identify trends and patterns in customer purchases and product performance.
2. To assess the effectiveness of discounts and their impact on total revenue.
3. To evaluate product ratings and payment preferences to enhance customer satisfaction and experience.
4. To provide actionable insights through visualizations for informed decision-making in inventory, pricing, and marketing strategies.

Scope of Work:

The project will involve the following tasks:

- **Data Exploration:** Analyze the dataset, including product categories, customer demographics, order details, payment modes, and ratings. Understand the distribution and characteristics of sales and customer behavior.

- **Data Preprocessing:** Clean the dataset by handling missing values, removing outliers, and ensuring consistency in data formatting. Prepare the data for in-depth analysis and visualization.
- **Feature Selection:** Identify the key features influencing customer choices, including product type, quantity, price, payment methods, and ratings, to understand purchasing patterns.
- **Data Visualization:** Create visualizations (e.g., bar charts, histograms, scatter plots, pie charts, heatmaps) to explore relationships between product sales, customer preferences, and transaction patterns, offering insights into high-performing categories and popular payment methods.
- **Interpretation of Results:** Analyze visual outputs to determine customer behavior, preferred product categories, and spending trends, providing insights into purchasing decisions and sales drivers.
- **Reporting:** Document findings and prepare a final report with visualizations, conclusions, and actionable recommendations to help Bigbasket improve its offerings, customer experience, and marketing strategies.

Methodology:

The project will follow a structured approach:

1. **Data Collection**: The dataset will be sourced from a Kaggle Website.

2. **Data Preprocessing:** Handle missing data using imputation techniques, detect and remove outliers, and normalize or standardize data where necessary to ensure consistency and accuracy.
3. **Exploratory Data Analysis (EDA):** Utilize descriptive statistics to summarize the dataset and create visualizations (e.g., bar charts, pie charts, line graphs) to explore sales trends, customer preferences, and relationships between product categories, ratings, and payment modes.
4. **Feature Selection:** Conduct correlation analysis to identify significant features such as product price, discount, stock availability, and order quantity, and how they impact sales performance.
5. **Evaluation and Interpretation:** Analyze the findings to interpret customer purchasing patterns, factors influencing the most popular products, and payment preferences. Identify insights to optimize inventory management, pricing, and marketing strategies.
6. **Visualization:** Generate various visualizations (treemaps, bar charts, histograms, scatter plots) to present key insights, highlighting trends, product performance, and correlations.
7. **Reporting:** Compile the analysis, results, and actionable insights into a detailed report, including visualizations and recommendations for improving the platform's offerings and customer experience.

Tools and Technologies:

The project will utilize the following tools and technologies:

- **Programming Language**: Python
- **Libraries**:
    - **Pandas**: For data manipulation and analysis.
    - **NumPy**: For numerical operations and handling arrays.
    - **Matplotlib**: For creating static and interactive visualizations.
    - **Squarify:** visualizing hierarchical data with rectangles proportional to their values.
    - **Mysql.connector:** Establishing a connection to the MySQL database using the MySQL Connector library.
- **IDE**: Jupyter Notebook
- **Data Source**: dataset from Kaggle website (BigBasket Entire Product List).

Expected Outcomes :

1. A detailed report highlighting sales trends across different product categories and subcategories on Bigbasket.
2. Identification of the most popular products, as well as those with the highest and lowest sales, to inform inventory management.
3. Insights into customer preferences based on payment methods, ratings, and product categories, helping to enhance customer experience and satisfaction.
4. Visualizations and data-driven recommendations for targeted marketing strategies, based on product discounts, price ranges, and customer behavior.
5. Understanding the relationship between product pricing, ratings, and sales, offering actionable insights for pricing strategies.
6. Key recommendations for optimizing the discount strategy and improving sales performance across different product categories.

Timeline:

Timeline The project is expected to be completed within a [specific timeframe, e.g., 4 weeks], with the following milestones:

- **Week 1**: Data Collection and Preprocessing

- **Week 2**: Exploratory Data Analysis and Feature Selection
- **Week 3**: Visualization
- **Week 4**: Reporting, and Final Submission

Conclusion:

This project provides a comprehensive analysis of Bigbasket's sales data, highlighting key trends, customer preferences, and product performance across various categories and subcategories. By utilizing different data visualization techniques such as bar charts, treemaps, scatter plots, and heatmaps, we have identified the top-selling products, the distribution of discounts, and the preferred payment methods. Furthermore, the analysis of product ratings, sales amounts, and quantity reveals significant insights into customer behavior and purchasing patterns.

Based on these findings, strategic recommendations have been provided to enhance Bigbasket's inventory management, optimize pricing and discount strategies, and target marketing efforts more effectively. By focusing on the most popular product categories and leveraging customer preferences, Bigbasket can drive better sales performance, improve customer satisfaction, and maintain competitive advantage in the grocery delivery market.