Amazon Product Bundling and Recommendation
7406 Project Group 115

Ashish Puri
apuri61@gatech.edu

Jagannath Banerjee
jbanerjee7@gatech.edu

Piyush Shivrain
pshivrain3@gatech.edu

1 Project Description

The objective of this project is to conduct a comprehensive analysis of Amazon sales data to gain valuable insights into sales trends, customer purchase behavior, and other factors influencing profitability. Leveraging various data mining techniques such as exploratory data analysis, clustering, association mining, and predictive modeling, we aim to extract meaningful patterns and relationships to build a product bundling strategy and recommend focused products to customers. By delving into multifaceted dimensions such as product categories, geographical distributions, and sales performance metrics, this analysis will furnish actionable insights to refine sales strategies and augment overall profitability.

2 Data

The Amazon sales dataset[1] has been downloaded from Kaggle.It consists of 3204 rows and 9 columns, containing information about order dates, shipping dates, email IDs of users, geographical locations, product categories, product names, sales figures, quantities sold, and profits.

Column Description:

- Order Date - Order_Date.

- Ship Date - Shipping Date.

- Email_ID - Email_ID of Users

- Geography - Location of Orders by Users.

- Category - Product Category

- Product Name - Product Name of Amazon

- Sales - Amazon Product Sales

- Quantity - how many units of a particular product are available.

- Profit - Amazon Sales Profit

3 Research Questions

This project aims to address several key research questions to gain a deeper understanding of Amazon sales dynamics and identify opportunities for up-sell and cross-sell recommendation.

- What are the top-selling product categories on Amazon?

- How do sales & profit vary by geography?

- How do sales & profit trends vary over time?

- Which products are associated and can be bundled together for up-sell and cross-sell?

- Which customer can be recommended for bundled products?

- Can customer segmentation improve targeted marketing efforts?

4 Proposed Methodology

To analyze Amazon sales data, we will employ a combination of descriptive statistics, exploratory data analysis (EDA), and advanced data mining techniques.

EDA:

- **Time Series Analysis:** Visualize sales and profit trends over time to identify seasonal patterns and long-term trends.

- **Profitable Category Analysis:** Identify the most profitable product categories based on sales and profit margins.

- **Popular Profitable Products:** Determine the top-selling and most profitable products within each category.

- **Geography-Based Analysis:** Analyze sales, profit, and popular products based on geographical locations to identify regional preferences and trends.

Modelling:

- **Product Bundling Strategy:** Use association rule mining techniques (e.g., Lift, Support, Confidence) to identify products frequently purchased together, enabling the development of product bundling strategies.

- **Customer Recommendation:** Implement content-based and collaborative filtering algorithms to provide personalized product recommendations to customers based on their purchase history and preferences.

References

[1] anandshaw2001 anandshaw2001. (2024, March 4). *Amazon_Sales_Dataset*. Kaggle. `https://www.kaggle.com/datasets/anandshaw2001/amazon-sales-dataset?resource=download`

[2] Amazon selling stats - sell on Amazon. (n.d.). https://sell.amazon.com/blog/amazon-stats

[3] M. Saranya. (2023, July 25). *Amazon sales analysis*. LinkedIn. `https://www.linkedin.com/pulse/amazon-sales-analysis-saranya-m/`

[4] According to the analysis conducted by J. Miglani (2023), titled *Amazon Sales and Profit Analysis for 2022: Top 10 Insights*, which was posted on Forrester's blog, the analysis can be accessed through the following link: `https://www.forrester.com/blogs/amazon-sales-and-profit-analysis-for-2022-top-10-insights/`.

[5] Vamanan, Dilip. (2024, February 16). *How to grow your profits with Amazon Data Analytics*. SellerApp Blog. `https://www.sellerapp.com/blog/amazon-data-analytics/`