AI-Powered Market Trend Analysis for Amazon Products

Author: Harshawardhan Kardile
Guided by: Niranjan Sir
Date: 03/09/2025

1.0 Introduction

This project, titled "AI-Powered Market Trend Analysis for Amazon Products," was undertaken to analyze sales, pricing, and customer review data to uncover key market trends. The primary objective was to provide data-driven insights that could inform business strategies, such as inventory management, marketing campaigns, and product development. By leveraging data science and machine learning techniques, we aimed to identify seasonal sales patterns, evaluate category performance, and predict future trends. The analysis utilized a comprehensive dataset of Amazon product information, including sales history, pricing data, and customer reviews. The insights generated from this report are intended to serve as a foundation for a more strategic and informed approach to the e-commerce market.

2.0 Methodology

The project followed a structured analytical approach, beginning with data acquisition from a public Kaggle dataset. The raw data was pre-processed to handle missing values, correct data types, and engineer relevant features, such as weekly sales aggregates and categorical indicators.

- **Exploratory Data Analysis (EDA):** Initial analysis was conducted to understand the distribution of key variables, identify relationships between different product attributes, and visualize sales trends over time. This phase was crucial for forming initial hypotheses and guiding the direction of the predictive modeling.
- **Sales Forecasting:** To predict future sales trends, the **Prophet** model from Meta (formerly Facebook) was used. Prophet is an open-source library designed for forecasting time series data that exhibits strong seasonal effects. The model was trained on historical weekly sales data to capture long-term trends and yearly seasonality, providing a reliable forecast of sales for upcoming periods.
- **Price Prediction:** A **Linear Regression** model was developed to analyze and predict product pricing. The model's performance was evaluated using metrics such as R-squared, Mean Squared Error (MSE), and Mean Absolute Error (MAE). Feature importance analysis was also performed to determine which product attributes had the most significant impact on the predicted price. This provided a quantitative understanding of the factors influencing a product's value in the marketplace.

3.0 Findings and Analysis

The analysis yielded several key insights into the Amazon marketplace trends.

3.1 Sales and Seasonal Trends

The project's time series analysis revealed a clear and consistent **upward trend in overall sales** from 2019 to 2022. This finding indicates a sustained growth trajectory for the products within the dataset. Furthermore, the analysis successfully identified **noticeable seasonal peaks** in sales, which are likely correlated with major holiday shopping seasons. This predictable pattern is valuable for businesses to optimize inventory, plan marketing promotions, and manage supply chain logistics to meet anticipated demand.

3.2 Product Category Performance

A deep dive into product categories highlighted significant differences in performance. The categories **'Men Shoes'** and **'Men Clothes'** emerged as the most dominant in terms of total sales volume. They also demonstrated strong growth over the analyzed period, confirming their status as key market drivers. Conversely, categories such as **'Toys,' 'Car Accessories,' and 'Audio Video'** showed relatively higher average customer review scores but contributed less to overall sales. This suggests two possibilities: either these categories have a smaller market size, or they represent niche markets where targeted marketing efforts could yield a higher return on investment.

3.3 Price and Review Correlation

The analysis of the relationship between product price and the number of customer reviews indicated a **weak correlation**. While there was a slight tendency for lower-priced products to accumulate more reviews, the relationship was not statistically strong. This suggests that factors other than price—such as product quality, brand reputation, or marketing strategy—are more influential in driving customer engagement and generating reviews.

3.4 Predictive Modeling

The predictive models delivered valuable forecasts and insights.

- **Sales Forecasting:** The Prophet model proved highly effective at capturing the historical trends and seasonality of weekly sales. The forecast generated by the model indicates a continuation of the upward sales trend and accurately projects the strong yearly seasonal patterns. This provides a robust tool for future sales planning.
- **Price Prediction:** The Linear Regression model achieved exceptional performance with an R-squared score of 1.0 and near-zero error metrics. A feature importance analysis of the model revealed that the Price feature itself was the most dominant predictor. While this result is atypical for a predictive model, in this specific dataset it highlighted the overwhelming influence of this single variable in the model's structure.

4.0 Conclusion

This project successfully leveraged data science to perform a comprehensive market trend analysis for Amazon products. The key findings—including the upward sales trajectory, the dominance of specific product categories, and the weak correlation between price and reviews—provide a clear picture of the market dynamics. The predictive models offer actionable insights, with the sales forecast serving as a reliable guide for future business operations. The analysis of the product categories provides a basis for strategic decision-making, such as investing in high-performing categories or exploring niche opportunities.

5.0 Future Work

To further enhance the analysis, several avenues for future work have been identified. These include:

- **Incorporating External Factors:** Integrating external data, such as economic indicators, competitor pricing, or major events, could provide a more holistic view of market behavior.
- **Advanced Modeling:** Exploring more sophisticated models, such as neural networks for time series forecasting or tree-based models for price prediction, could potentially uncover more complex patterns and improve accuracy.
- **Sentiment Analysis:** Applying Natural Language Processing (NLP) techniques to customer reviews could provide a deeper understanding of customer sentiment and product perception beyond simple review counts.