Amazon Product Sales Analysis – Portfolio Project

1.Overview

This project explores a dataset of Amazon products to analyze **sales trends, product ratings, discounts**, and **category-wise performance**. The objective is to extract business insights that can help e-commerce sellers optimize product listings, pricing, and customer satisfaction.

2.Business Objectives

- Identify top-performing and underperforming product categories
- Analyze the impact of **discounts** on **sales** and **ratings**
- Discover **low-rated** and **low-engagement** products
- Offer actionable insights for inventory, marketing, and quality control

3.Tools & Libraries

- **Python**: Pandas, NumPy, Seaborn, Matplotlib
- **Jupyter Notebook**
- **Dataset**: Amazon product sales (source: kaggle)

4.Dataset Overview

The dataset includes information about product pricing, discount percentage, ratings, rating counts, categories, and customer reviews. Data cleaning was performed to convert formats, remove nulls, and engineer simplified category features for analysis.

5.Data Cleaning

- Dropped missing values from key columns (e.g., rating_count)
- Converted price and rating columns to numeric
- Cleaned discount percentages
- Created a 'Reduced_Category' field to simplify category paths

6. Exploratory Data Analysis (EDA)

The EDA was performed to extract trends and insights from the product-level data. It covered discount patterns, sales volume, customer ratings, and product category performance.

6.1 Discount & Pricing Analysis

- Distribution of discount percentages
- Average discount by product category

6.2 Ratings Analysis

- Overall product rating distribution
- Average ratings per category
- Lowest-rated products

6.3 Sales Volume Analysis

- Total sales volume per category (using rating count as a proxy)
- Least sold products
- Relationship between discount and sales

6.4 Combined Insights

- Heatmap showing average discount across rating ranges and categories
- Analysis of categories performing well despite low discounts

7. Key Visualizations

Discount distribution histogram
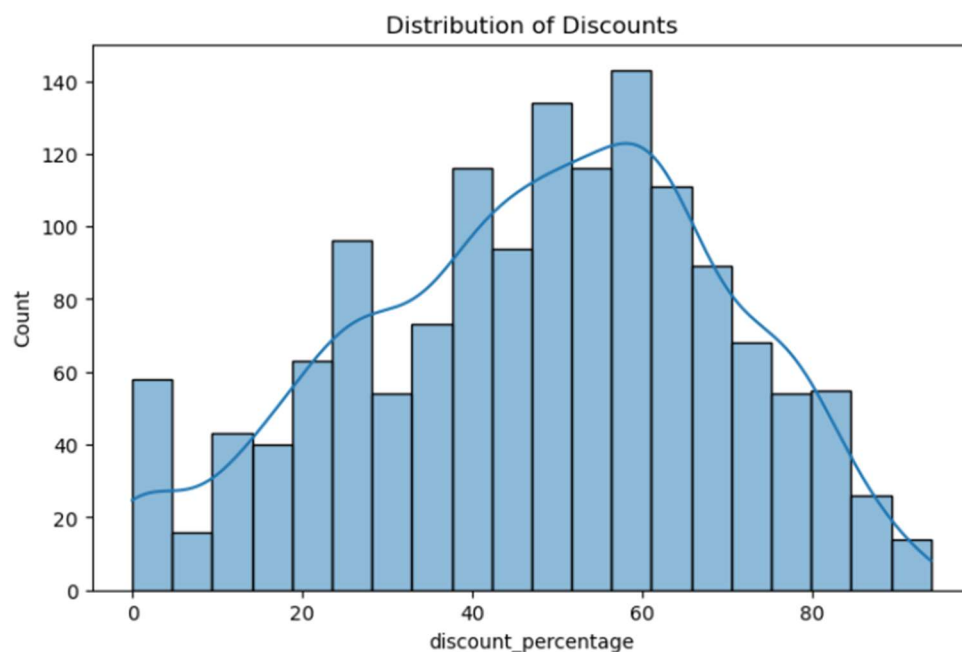
- The majority of discounts on Amazon fall between 50-60%, indicating that sellers prefer moderate discounting over extreme price cuts. Products with very high (80%+) or very low (10%-) discounts are less frequent, suggesting that moderate discounts might be the most effective in attracting buyers.

Average discount by category (bar chart)

Average Discount Percentage by top 10 category
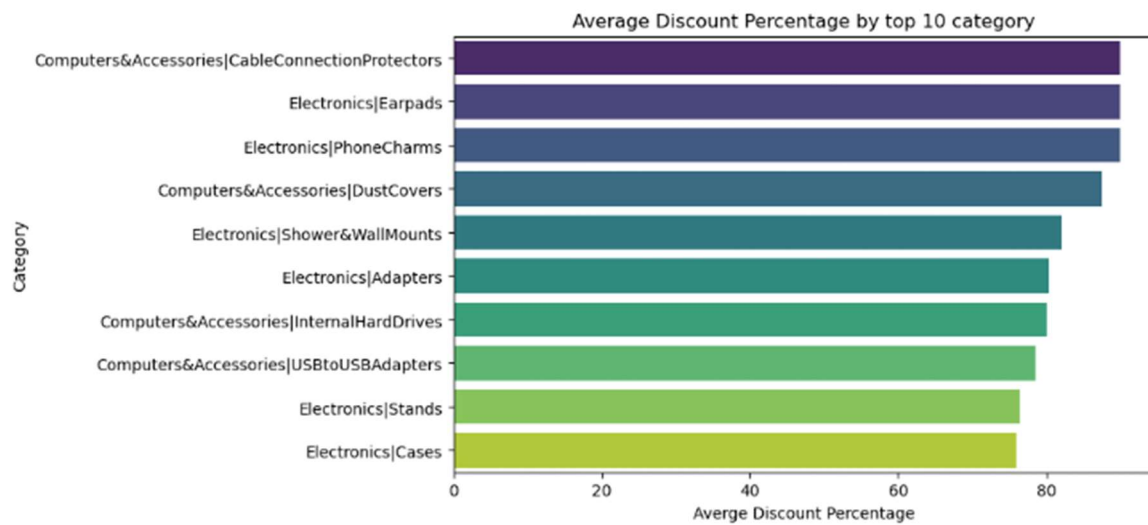
- High discounts are heavily concentrated in electronics-related accessories.
- Categories such as Cable Connection Protectors, Earpads, and Phone Charms have average discounts exceeding 75%. This indicates a highly competitive or oversaturated segment where sellers may rely on steep discounts to boost visibility and sales.

Rating distribution (histogram)

Distribution of Rating Across All Products
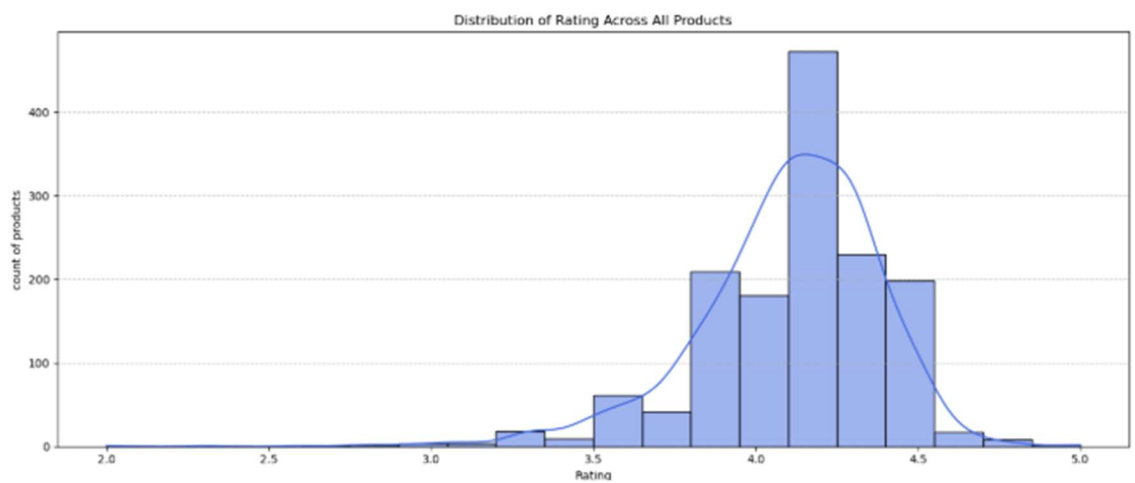
- Most products have high ratings (4.0 - 4.5)

Top 10 selling categories
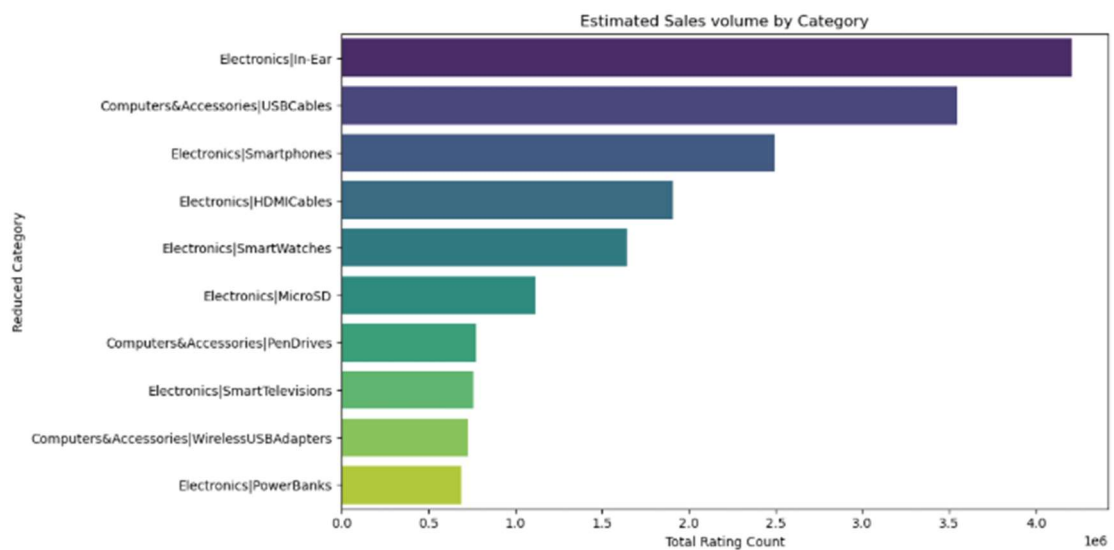
The Electronics | In-Ear category has the highest estimated sales volume, suggesting that earphones/headphones are one of the most popular product types on Amazon.

Least-rated products

The product "Khaitan ORFin Fan Heater for Home and Kitchen-KO 2215" holds the lowest average customer rating in the dataset, with an approximate rating of 2.1.

Least-sold products

The products "NGI Store Pet Hair Remover for Washing Machines" and "Khaitan ORFin Fan Heater – KO 2215" recorded the lowest sales volume, with only 2 customer ratings each.

Heatmap: Rating × Discount %

High-rated categories like Cord Management, Memory, and Power LAN Adapters receive significant discounts, suggesting competitive pricing or stock clearance strategies. In contrast, categories like Painting Materials and Streaming Clients maintain high ratings with minimal or no discounts.

8. Key Business Takeaways

- Electronics and tech accessories lead in both sales and ratings
- Discounts do not strongly drive sales volume
- Most products maintain high ratings (4.0+)
- Several low-rated products suggest quality or listing issues
- Some products have very low sales and visibility despite belonging to popular categories

9. Conclusion

This analysis helps Amazon sellers and e-commerce managers gain insight into what drives product performance. By examining sales patterns, rating behaviour, and discount strategies, this project offers a foundation for better inventory decisions and marketing prioritization.