

PYTHON DATA ANALYSIS PROJECT



# DISCOUNT IMPACT ON AMAZON SALES & CUSTOMER ANALYSIS



PREPARE BY  
TAARA NGUYEN

DATE PREPARED  
17/03/2025

# BACKGROUND

Amazon, a leading e-commerce platform, offering an extensive range of products alongside frequent discounts and promotions. However, despite the implementation of significant discounts across various categories, data indicates that higher discount levels do not consistently lead to improved customer satisfaction. Certain products with substantial discounts still receive lower ratings, raising critical questions regarding customer perception, product quality, and purchasing behavior.

In addition, Amazon is focused on optimizing its discount strategies for high-value items to ensure customers are satisfied with their purchases. The company also seeks to gain a deeper understanding of which product categories are most responsive to discounts and how customer feedback can be leveraged to drive sales growth and improve the overall shopping experience.

## EXECUTIVE SUMMARY

- Objective: Examine the distribution of discounts to gain insights into their effects on product ratings, sales performance, and customer satisfaction
- Key insights:
  - Relationship between discounts and ratings.
  - Other factors impacting rating besides discounts.
  - Primary driven deterrents to growth in sales and customer satisfaction of high-ticket products.
  - Most beneficial categories from discount promotion and their top 10 best-selling products as well as least-selling products.
  - Frequent concerns identified in products with low ratings.
- Recommendation:
  - Improve revenue and rating in underperforming products.
  - Leverage promotional strategies to high-rated products.

## DATA PREPARATION AND METHODOLOGY

Code Link: <https://github.com/taaranguyen/Amazon-data-analysis-project>



# ANALYSIS

## 1. RELATIONSHIP BETWEEN DISCOUNTS AND RATINGS

### 1. 1. Discount and Rating Distribution across Products

Amazon product's ratings range from a minimum value of 2 to a maximum value of 5, with **2** representing the **lowest** possible rating and **5** the **highest**.

The majority of ratings fall within the range of 3.75 to 4.5, indicating a concentration of customer feedback in this mid-to-high rating spectrum (figure 1).

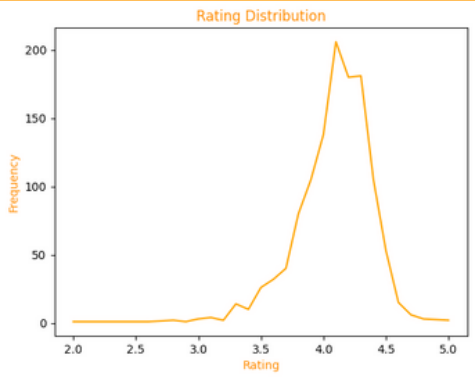


Figure 1. Rating Distribution

Out of 1,256 products, 1,211 are eligible for **discount** promotions ranging from **2%** to **94%**.

The most frequent discounts fall between 30% to 60% (figure 2).

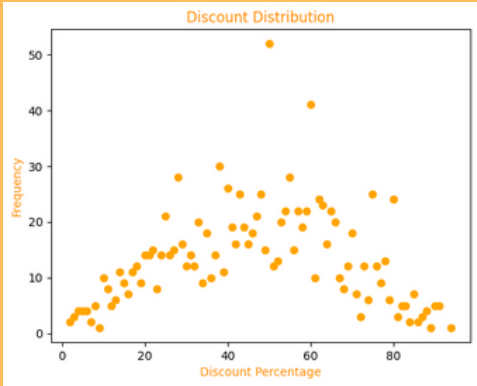









Figure 2. Discount Distribution

The total number of products in the high discount range is more than three times greater than in the low discount range, and only marginally exceeds the number in the medium discount range as shown in figure 3.

Discount percentages are divided into 3 different ranges for analysis:

- Low:  
- Med:   
- High:  

Discount Range	Total Product
Low Discount	178
Med Discount	444
High Discount	589

Figure 3. Table of Total Product in each Discount Range

# 1.2. Rating Distribution across Discount Ranges

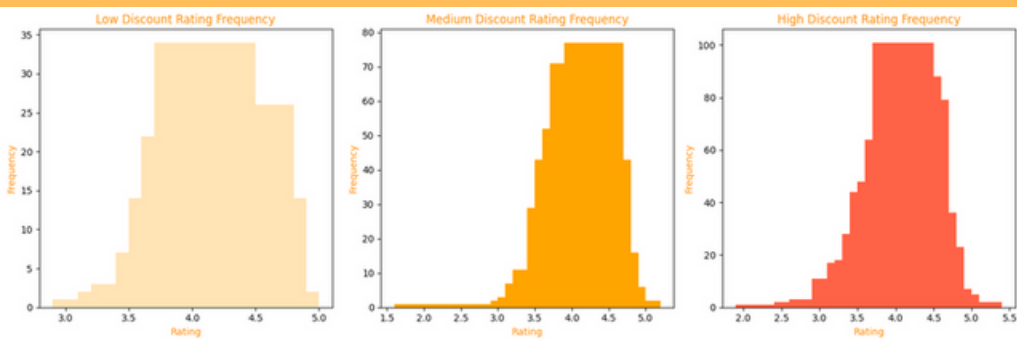


Figure 4. Rating Frequency in Low, Med and High Discount Level

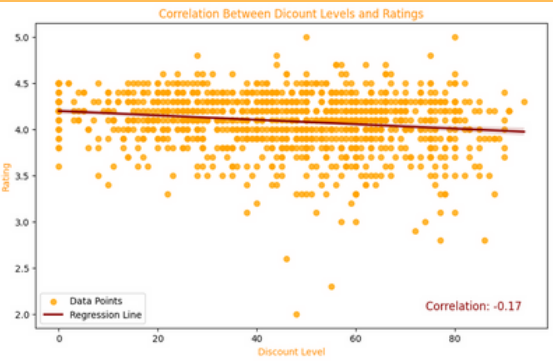
Discount Range	Low Discount	Med Discount	High Discount
Lowest Rating	3.3	2.0	2.3
Highest Rating	4.6	4.8	5.0
Average Rating	4.15	4.10	4.05

Figure 5. Table of Lowest, Highest & Average Rating across Discount Range

Based on this finding (figure 4 & figure 5), there appears to be a slight negative trend between discount levels and average ratings, although it is also evident that higher discounts tend to correlate with higher ratings.

To assess the relationship between them, a correlation analysis was conducted.

Figure 6. Correlation Between Discount Levels and Ratings



The resulting correlation coefficient,  $R = -0.1467$ , indicates a **weak negative linear relationship** between **discount levels** and **ratings**. This implies that as the discount percentage slightly increases, the ratings tend to decrease, though the relationship is weak. Hence, discounts do not appear to be a critical factor in determining customer satisfaction (figure6).

## 2. OTHER POTENTIALLY AFFECTING ASPECTS

### 2.1. Discounted Price Distribution across Products

Discounted prices should also be taken into account, as they could play a crucial role in shaping customer ratings. In some cases, customers may have higher expectations for more expensive products. As a result, if the product falls short of these expectations, customers may assign lower ratings.

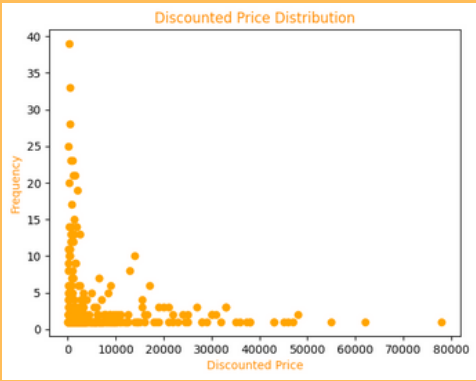


Figure 7. Discounted Price Distribution

The prices after discount range from ₹39.00 to ₹77,990.00 (figure 7).

Discounted prices are divided into 3 different ranges for analysis:

Low

₹500

Med

₹500 – ₹1000

High

₹1000

Price Range	Low Price	Med Price	High Price
Total Product	418	218	575

Figure 8. Table of Total Product in each Discounted Price Range

The total number of products in the high-price range appears to be more than twice that of the medium-price range, with only a small difference compared to the low-price range (figure 8).

## 2.2. Rating Distribution across Discounted Price Ranges



Figure 9. Rating Frequency in Low, Med and High Price Range

Price Range	Low Price	Med Price	High Price
Lowest Rating	2.6	2.3	2.0
Highest Rating	5.0	4.7	4.8
Average Rating	4.06	4.09	4.10

Figure 10. Table of Lowest, Highest & Average Rating across Price Ranges

This suggests an upward trend between discounted price ranges and average ratings. However, higher discounted prices are also associated with the lowest ratings. To further investigate this, a correlation analysis was conducted to assess their relationship (figure 9 & figure 10).

The correlation coefficient of **0.1342** between **discounted prices** and **ratings** indicates a **weak positive linear relationship**. This suggests that while there is a slight tendency for ratings to increase with higher discounted prices, the relationship is not strong enough to draw definitive conclusions. Thus, customer satisfaction does not appear to be significantly impacted by Discounted Prices (figure 11).

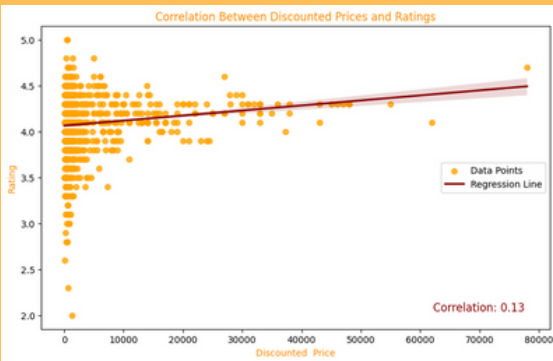


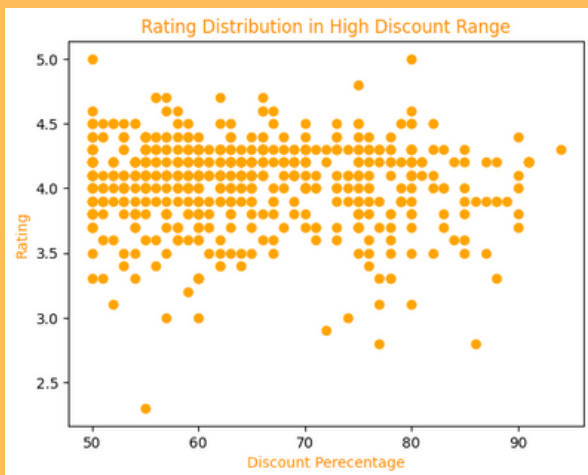
Figure 11. Correlation Between Discounted Prices and Ratings

## 2.3. Most Common Feedback for Products with High Discounts but Low Ratings

Figure 12 displays the distribution of ratings across 589 products in the high discount range, with a clear visualization of the most frequent ratings falling between 3.5 and 4.5.

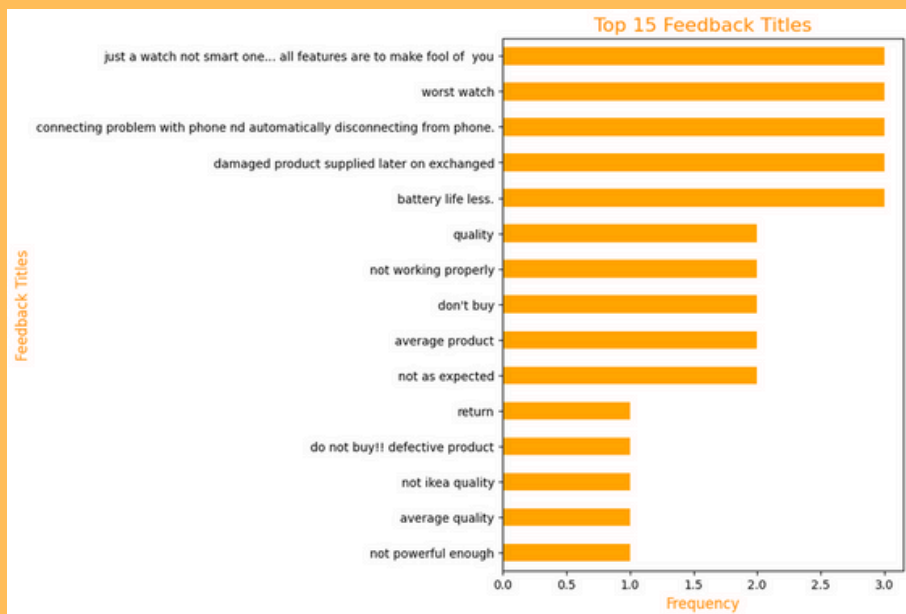
Figure 12. Rating Distribution in High Discount Range

Nevertheless, there are several abnormally low ratings, which prompt significant questions regarding customer perception, product quality, and purchasing behavior (figure 12). Consequently, the reviews with low ratings are analyzed to identify the most common feedback, helping to pinpoint the underlying issues.



As shown in Figure 13, **poor quality** and **malfunctioning products** appear to be the **primary factors** influencing customer satisfaction, such as disconnection and battery issue.

Figure 13. Top 15 Feedback Titles



### 3. STRATEGIES FOR HIGH TICKET ITEMS

#### 3.1. Sales Performance and Rating Distribution in High-Ticket Products

After applying discounts, prices range from ₹39.00 to ₹77,990.00, with high-ticket items defined as **products priced over ₹1,000.00**.

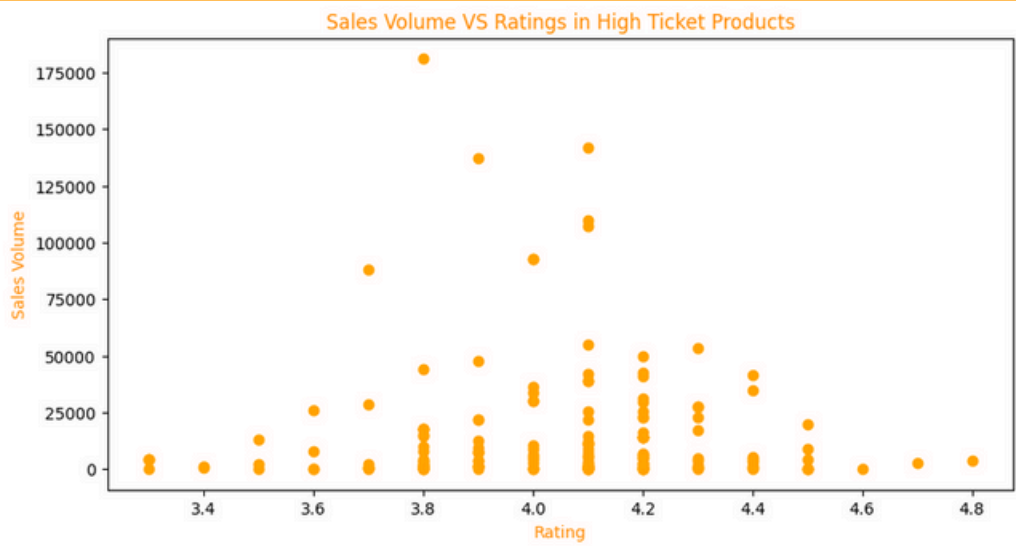


Figure 14. Sales Volume VS Ratings in High-Ticket Products

Figure 14 illustrates ratings in this segment range from 3.3 to 4.8, with the highest frequency occurring between 3.8 and 4.3.

Its total sales range from just a few units to over **175,000** units sold.





## 4. CATEGORY-SPECIFIC INSIGHTS

### 4.1. Average Discount VS Average Rating Across Categories

A total of 9 categories exists, of which only 8 are eligible for discounts. Average discount distributed across them ranging from **21.28%** to **57.50%**.

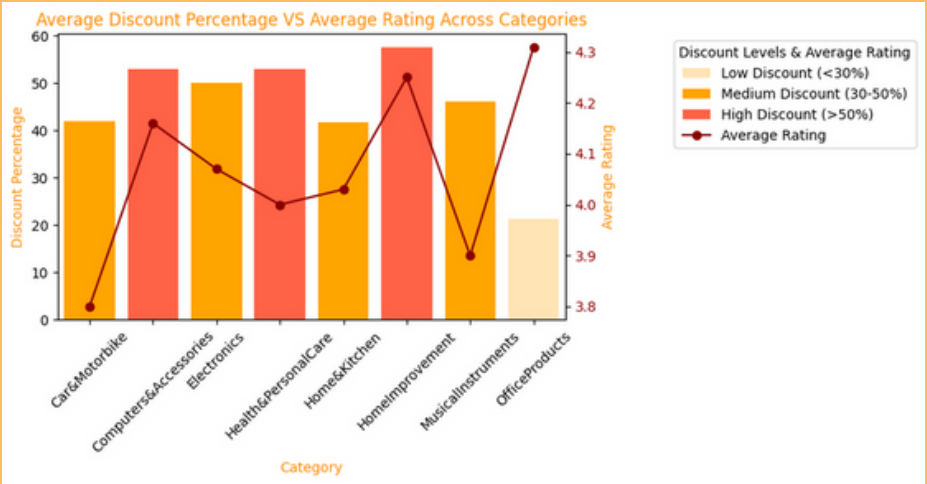


Figure 17. Average Discount Percentage VS Average Rating Across Categories

Home Improvement has the highest average discount, with a total of 2 products, while Office Products has the lowest average discount, with a total of 18 products (figure 18).

Although average discount of **Electronics** lies within the medium discount range, it has the **most total products** and **sales volume** among categories.

Figure 18. Total Products per Main Category

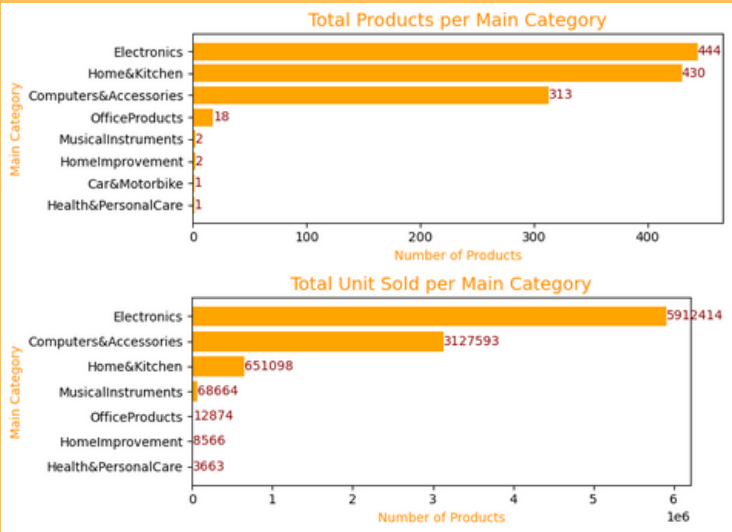


Figure 19. Total Unit Sold per Main Category

## 4.2. Most Beneficial Categories from Discounts

The most advantageous product categories are those that feature products with significant discounts (high discount range) and also generate high total sales.

After analysis, three **categories** emerge as the **most beneficial** including **Electronics**, **Computers & Accessories** and **Home & Kitchen** respectively (figure 20 & figure 21).

Category	Total Sales	Average Discount
Electronics	5912414	66.73%
Computers & Accessories	3127593	65.04%
Home & Kitchen	651098	59.54%

Figure 20. Table of Total Sales and Average Discount of Most Advantageous Categories

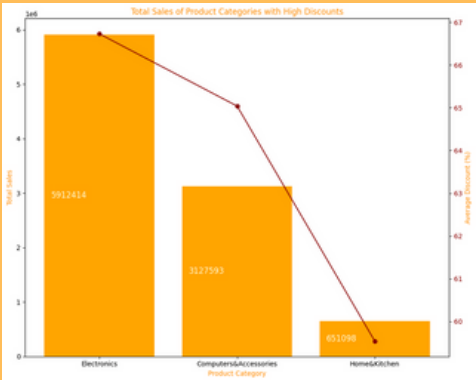


Figure 21. Total Sales of Product Categories with High Discount

Additionally, the **top 10 products** with the **highest sales** and the top 10 products with the **lowest sales** are identified to enhance inventory management and marketing strategies (figure 22).

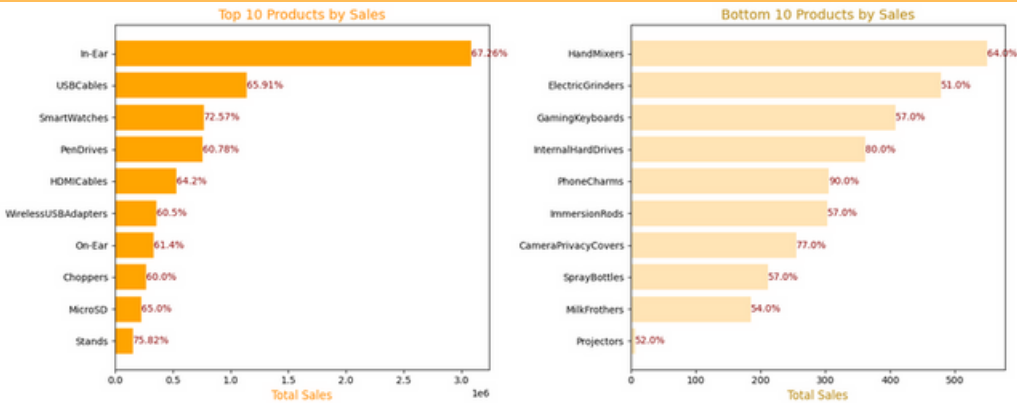


Figure 22. Top 10 and Bottom 10 Products by Sales

The inventory of the top 10 products shall be prioritized for restocking in anticipation of upcoming discount promotions, while the sales performance of the bottom 10 products could be improved by offering higher discounts.

## 5. CUSTOMER FEEDBACK IN LOW-RATED PRODUCTS

The low-rated products are primarily found in the Electronics and Home & Kitchen categories, including items such as Remote Controls, Smart Watches, Hand Blenders, Juicer Mixer Grinders, Fan Heaters, Vacuum Sealers, and Water Purifier Accessories.

Customers frequently raise concerns about product quality, with the **majority of feedback** indicating that the products are “**not working properly**”.

Figure 23. Customer Common Feedback in Low-Rated Products

# KEY FINDINGS AND RECOMMENDATIONS



While discounts and prices exhibit a limited influence on customer satisfaction, their impact is relatively minor compared to the substantial effect of unreliable quality and malfunctioning products on customer ratings which explains for the concern of the low rating appearance despite high discounts on products.



For high-ticket items, offering budget-friendly prices combined with high-quality products will strengthen total sales and improve product ratings.



Electronics, Computers & Accessories and Home & Kitchen are the 3 most beneficial from discount promotion since their total sales rank highest across 8 discount-applicable categories. In which, an enormous gap in sales between top 10 with the most sales and bottom 10 with the least sales is clearly noticeable. The top 10 products which should be significantly replenished in preparation for the upcoming discount promotion, while the bottom 10 products with weakest sales volume should either upgrade their quality nor have their discount percentages increased further to drive sales performance.



Regarding low-rated products, defective and malfunctioning items pose a critical issue. To address this, several approaches can be implemented, including enhancing quality control processes, conducting root cause analysis, and offering clear return and refund policies.