amazon

Python Data Analysis

Impact of Discounts on Amazon Sales & Customer Analysis

GitHub: https://github.com/wendyluong/Amazon-Project.git
**By: Thu Huyen(Wendy) Luong**

Executive summary:

Amazon employs widespread discounting to boost sales and ratings, yet data analysis reveals that higher discounts do not consistently enhance customer satisfaction. In fact, higher discount percentages show a negative correlation with sales growth and satisfaction increases, suggesting that aggressive discounting may not be an effective long-term strategy. Among product categories, Computers & Accessories benefit the most from deep discounts, experiencing a significant increase in rating count, whereas other categories do not exhibit the same trend. This highlights the necessity of a category-specific discounting strategy. Despite variations in discounting, customer satisfaction remains strong, with 75% of ratings between 4-5 stars. Keyword analysis indicates that quality, rather than pricing, is the primary driver of satisfaction. Additionally, a large proportion of reviews come from non-members, presenting an opportunity for Amazon to enhance customer engagement and loyalty programs. Strengthening membership incentives and personalized experiences could improve retention and reduce dependence on discounts as the primary sales driver.

Key finding:

- Diminishing Returns on Discounts: Moderate discounts (20%-40%) optimize ratings and engagement, while deep discounts (60%-80%) correlate with lower ratings and reduced rating counts, indicating a decline in perceived product quality.
- Computers & Accessories Benefit Most from Deep Discounts: This category sees a notable increase in rating count when heavily discounted, unlike other product categories.
- Customer Satisfaction is Quality-Driven: Only 0.4% of ratings fall below 3 stars, with positive reviews emphasizing "good product," "good quality," and "easy to use," proving that product performance outweighs pricing in driving satisfaction.
- Opportunity to Strengthen Customer Engagement: A large proportion of reviews come from non-members, highlighting the potential to enhance loyalty programs and customer retention strategies to reduce reliance on discounting.

1.Company overview

Distribution of Products by Main Categories
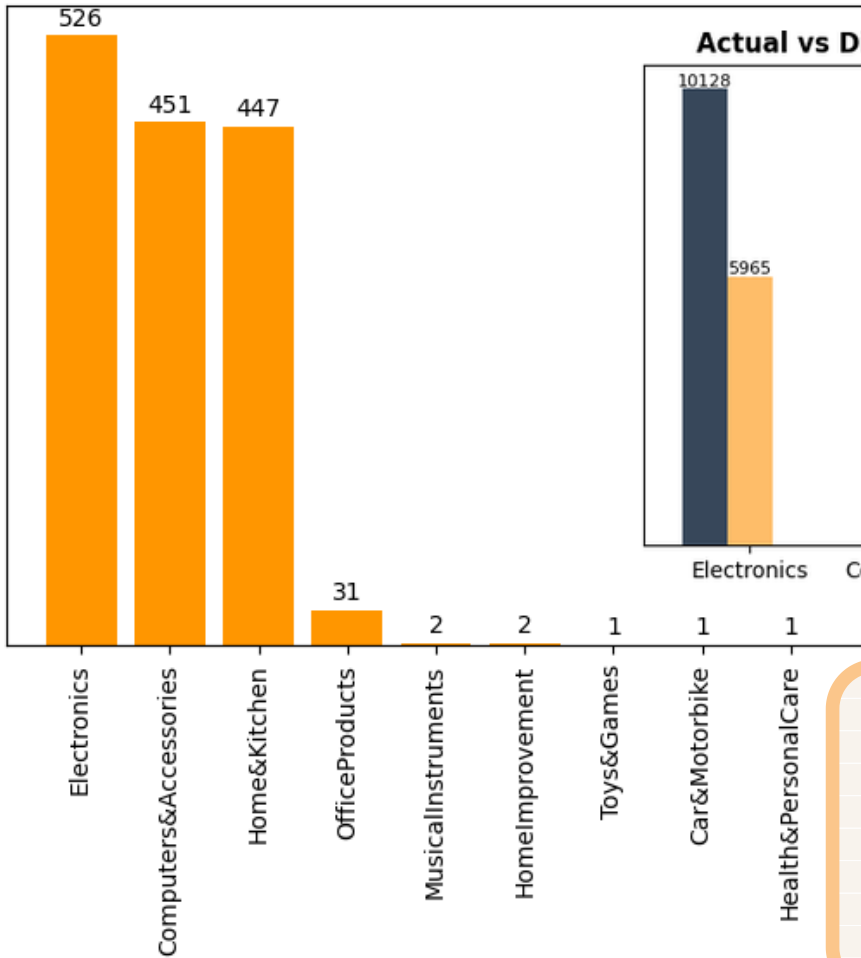
Actual vs Discounted Price among main categories

- Actual price
- Discounted price
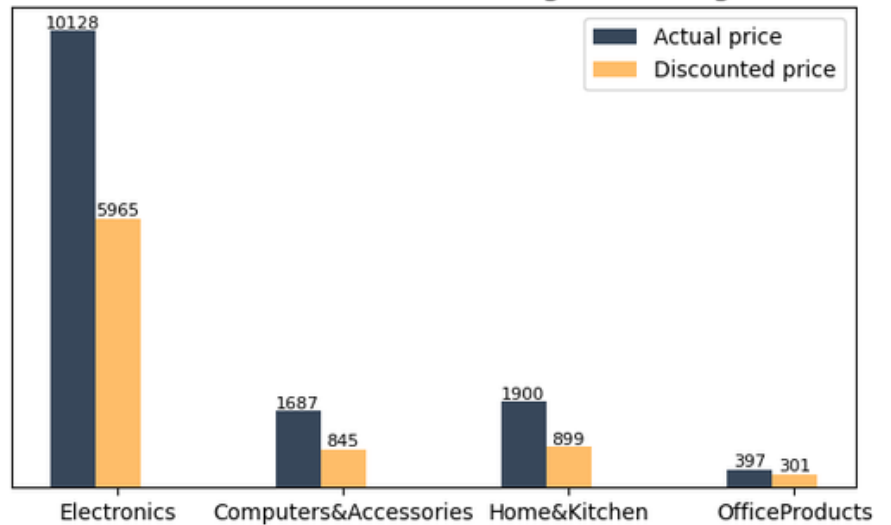
- Amazon's product distribution is highly imbalanced.
- Having the highest actual and discounted prices, Electronics remains the top-selling category, indicating strong consumer demand and willingness to pay premium prices.

2.Discount Effectiveness

Discount Distribution with Average Rating
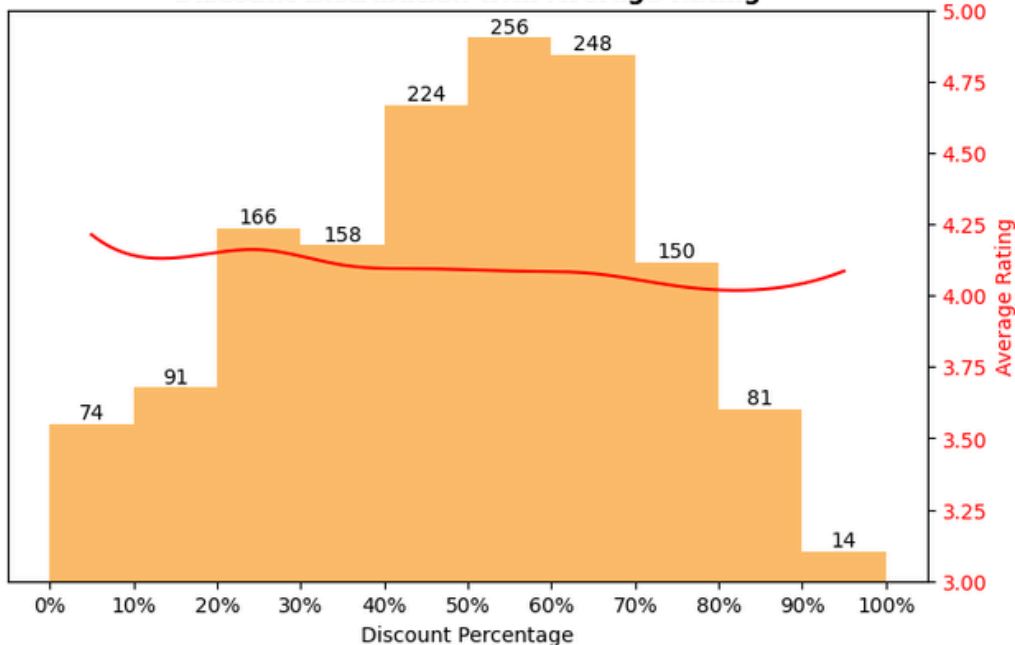
- Most products are discounted between 40%-70%, showing that Amazon strategically offers mid-to-high discounts to drive sales.
- Heavily discounted products may have marginally lower perceived rating, the difference is not dramatic.

Average Actual Price and Discount Amount by Discount Percentage Bins

- Higher discount percentages are applied to lower actual price products ➡ discount amounts remain relatively similar across mid-range categories.
- Discount amounts peak at >80%, but low distribution suggests it's mainly for clearance ➡ profit maintainace.

Impact of Discounts on Log Rating Counts
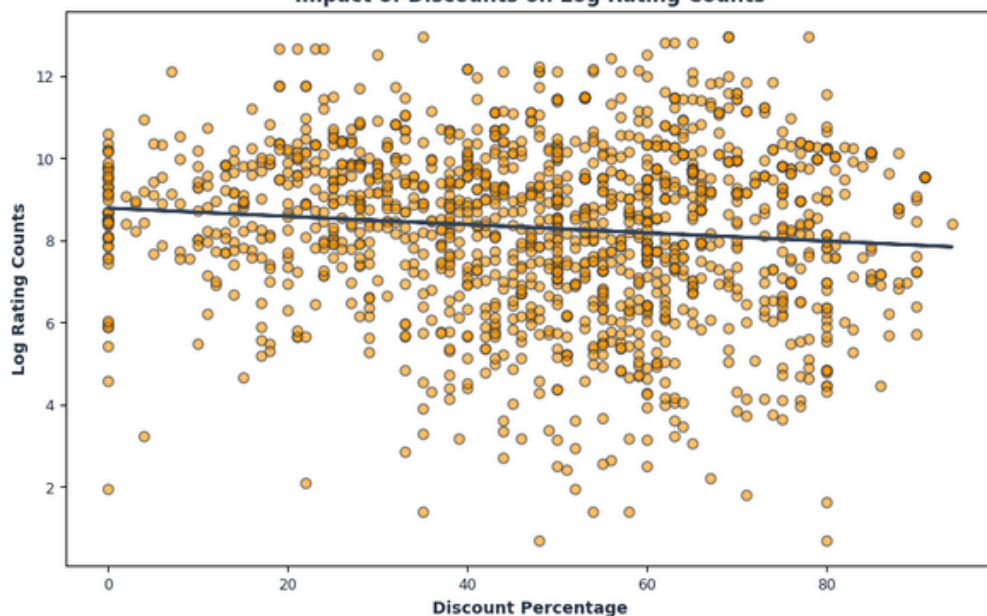
R-Squared	Coef.	P-Value
1.1%	-0.0100	0

- Significant negative statistical relationship between discount percentage and log rating, discount percentage and rating.
- Higher discounts tend to have side effect on rating and rating count.
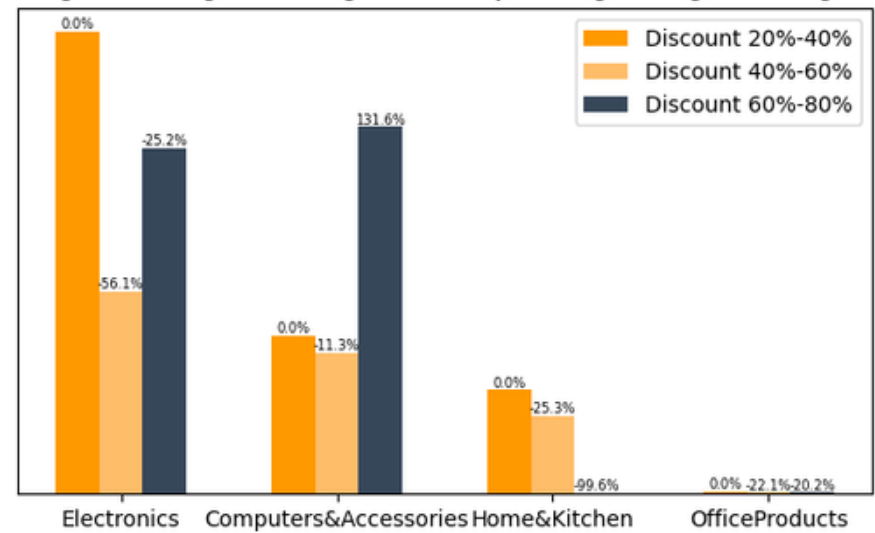
Impact of Discounts on Rating
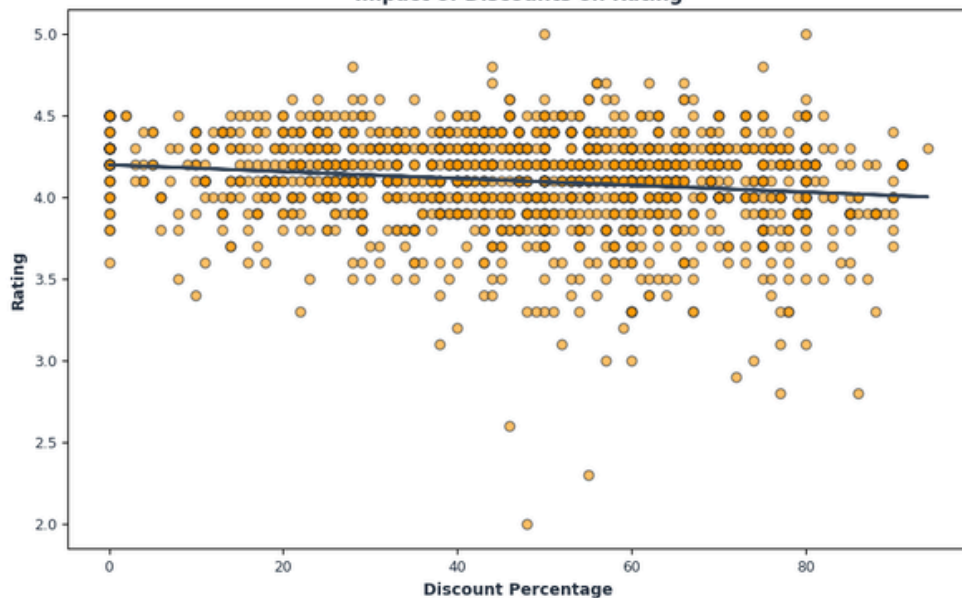
R-Squared	Coef.	P-Value
2.4%	-0.0021	0

Discounts do not effectively boost sales or increase customer satisfaction.

➡ Ineffective discount strategy

Rating count changes according to discount percentage among main categories

Legend:
- Discount 20%-40%
- Discount 40%-60%
- Discount 60%-80%

Bar values:
- Electronics: 0.0%, -56.1%, -25.2%
- Computers&Accessories: 0.0%, -11.3%, 131.6%
- Home&Kitchen: 0.0%, -25.3%, -99.6%
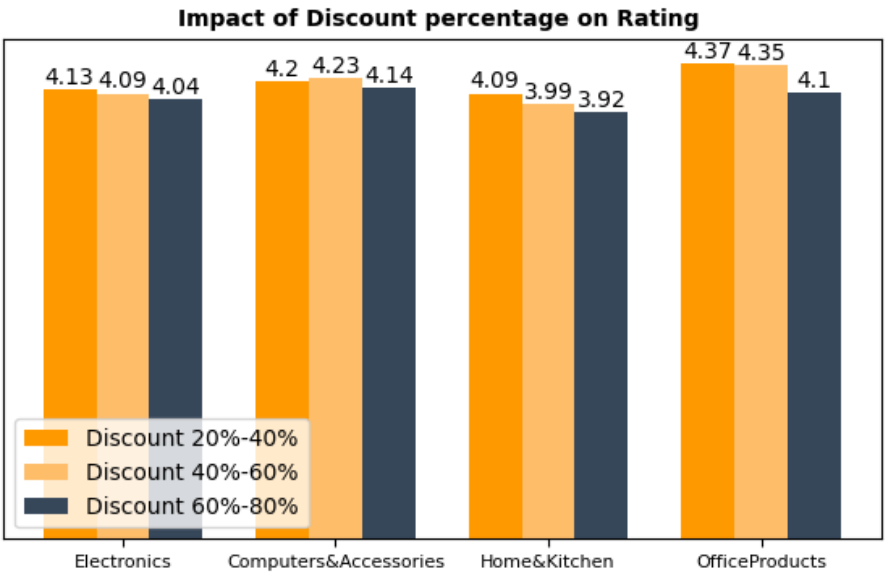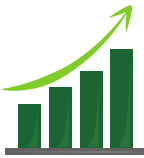- OfficeProducts: 0.0%, -22.1%, -20.2%

amazon

Extensive discounts (60%-80%) have extreme effects:

- Sharp increase in rating count for Computers & Accessories.
- Decline in rating count for other categories
- Decrease rating for all category, while it takes roughly 40% total discount distribution.

- Office Products maintain the highest rating despite having the lowest discount levels.
- A discount range of 20%-40% tends to provide the best impact on both rating count and rating, except for Computers & Accessories.

Impact of Discount percentage on Rating

Bar values:
- Electronics: 4.13, 4.09, 4.04
- Computers&Accessories: 4.2, 4.23, 4.14
- Home&Kitchen: 4.09, 3.99, 3.92
- OfficeProducts: 4.37, 4.35, 4.1

Legend:
- Discount 20%-40%
- Discount 40%-60%
- Discount 60%-80%

3.Rating and Sentiment analysis

Top 20 Keyphrase Frequencies in Review Content

Rating Distribution
- <3: 0.4%
- 3-4: 23.8%
- 4-5: 75.8%
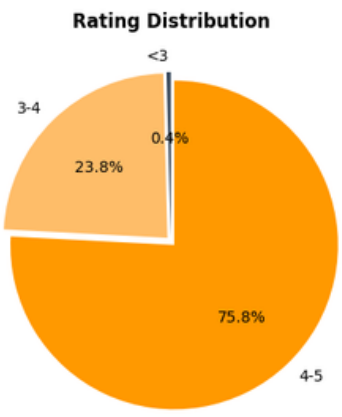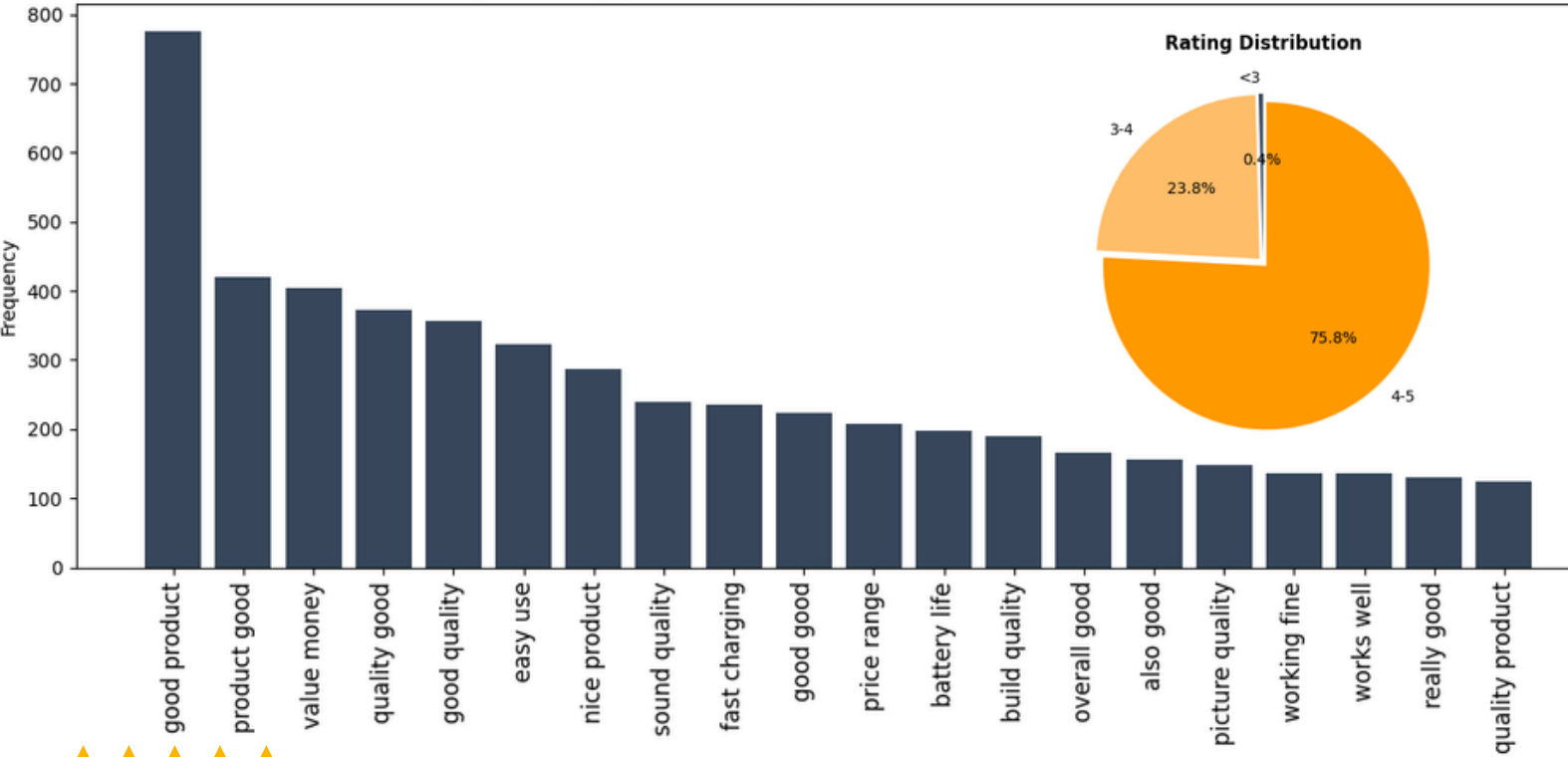
The majority of ratings (75.8%) fall within the 4-5 range, indicating strong customer satisfaction.
The primary feedback focuses on product quality. Only two phrases ("value money" and "price range") relate to pricing, indicating that customers prioritize product performance over price considerations.
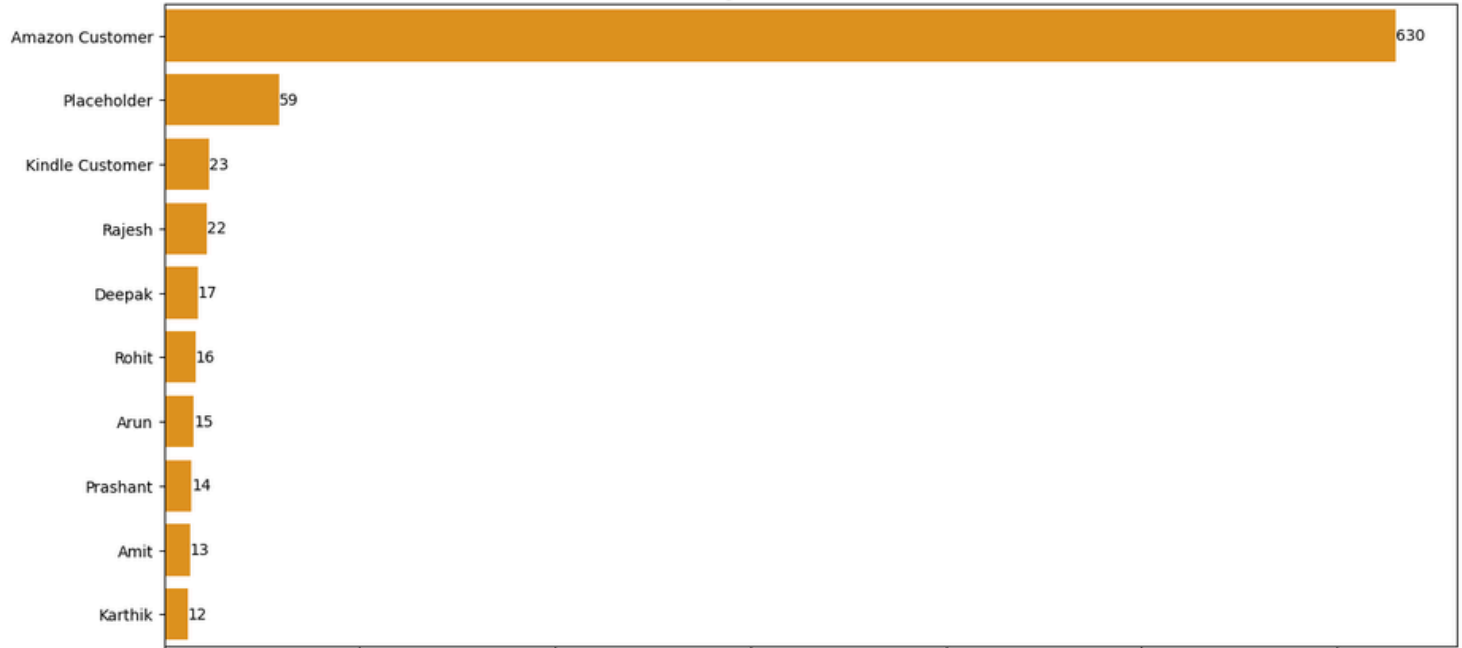
Wordcloud for review under 3 rating

- Common complaints focus on functionality, durability, and defects, with frequent mentions of words like "replacement," "return," "leaking," and "working."

- Customers are more dissatisfied with product performance than with pricing.

	discounted_price	actual_price	discount_percentage	rating	actual_discount_value	log_rating_count
discounted_price	1	0.96	-0.24	0.12	0.76	0.074
actual_price	0.96	1	-0.12	0.12	0.91	0.057
discount_percentage	-0.24	-0.12	1	-0.16	0.087	-0.11
rating	0.12	0.12	-0.16	1	0.11	0.24
actual_discount_value	0.76	0.91	0.087	0.11	1	0.022
log_rating_count	0.074	0.057	-0.11	0.24	0.022	1

- Discounts (both percentage and absolute value) show weak correlations with rating and rating count, suggesting that discounting alone does not strongly influence customer satisfaction or engagement.
- Other factors might be product quality, market competition, and delivery/service.

Top 10 Reviewers

Reviewer	Count
Amazon Customer	630
Placeholder	59
Kindle Customer	23
Rajesh	22
Deepak	17
Rohit	16
Arun	15
Prashant	14
Amit	13
Karthik	12

- Anonymous and non-member reviews make up the majority, indicating potential gaps in reviewer accountability and engagement.
- At least 8 identified customers have reviewed more than 10 products, suggesting a group of frequent reviewers

RECOMMENDATION

1. Optimize Pricing Strategy

- Reduce discount levels across all categories except Computers & Accessories
- Implement a tiered discount system, offering additional price reductions exclusively for loyalty customers to increase engagement and retention.
- Lowering discounts across most categories will help maintain product value perception and increase revenue margins while ensuring competitiveness.

2. Enhance Product Quality to Boost Sales & Ratings

- Prioritize quality improvements based on negative review patterns, such as durability, product descriptions, and customer service experiences.
- Higher prices should be justified with better product quality, ensuring that customers associate premium pricing with a superior shopping experience.

3. Improve Non-Member Experience

- Offering limited-time trial discounts or exclusive first-time purchase deals to encourage non-members to join loyalty programs.
- Introduce alternative incentives for non-members, such as free shipping or bundled offers, to prevent customer churn due to perceived unfair pricing differences.

4. Address Imbalance in Product Categories

- If the current imbalance in product categories is unintentional, consider expanding the variety and selection in underrepresented categories to diversify revenue streams and reduce reliance on a few dominant segments.
- Conduct demand forecasting and market analysis to identify potential growth opportunities in low-representation categories.