

Amazon Product Analysis

By Shefa Alkilany - Data Analyst



Project Overview:

This project analyzes Amazon product data to uncover patterns in ratings, discounts, and customer engagement across multiple product categories. The goal is to identify actionable insights that can inform pricing strategies, product positioning, and marketing efforts.

Using **SQL**, the raw dataset was cleaned, filtered, and aggregated to create meaningful metrics such as discount ranges, rating segments, and category-level averages. The processed data was then visualized in **Tableau** to build an interactive dashboard.

The dashboard provides:

- **Key KPIs** such as total product count, average rating, average discount, maximum discount, and average number of reviews.
- **Category-Level Analysis** showing variations in average ratings and discounts between product categories.
- **Customer Engagement Segmentation** based on the number of ratings, highlighting the distribution of customer interaction across products.
- **Discount Range vs. Average Rating** analysis, revealing the relationship between pricing strategies and customer satisfaction.
- **Product Distribution by Category**, offering a quick view of category size and market representation.

From the analysis, several insights emerged:

Products with **very high discounts (76–100%)** tend to have lower average ratings, while low-discount products achieve higher ratings.

Most customer engagement is concentrated in a smaller set of highly-reviewed products, while a large portion of the catalog receives minimal reviews.

This project demonstrates skills in **data extraction, transformation, and visualization** while emphasizing the ability to translate raw data into business-relevant insights.

Data Source & Description

Dataset Origin:

Amazon Sales Dataset from Kaggle

(<https://www.kaggle.com/datasets/karkavelrajaj/amazon-sales-dataset>)

Dataset Size:

Approximately 13,000 records with detailed sales data

Time Range:

Sales data spans from 2013 to 2018

Main Columns:

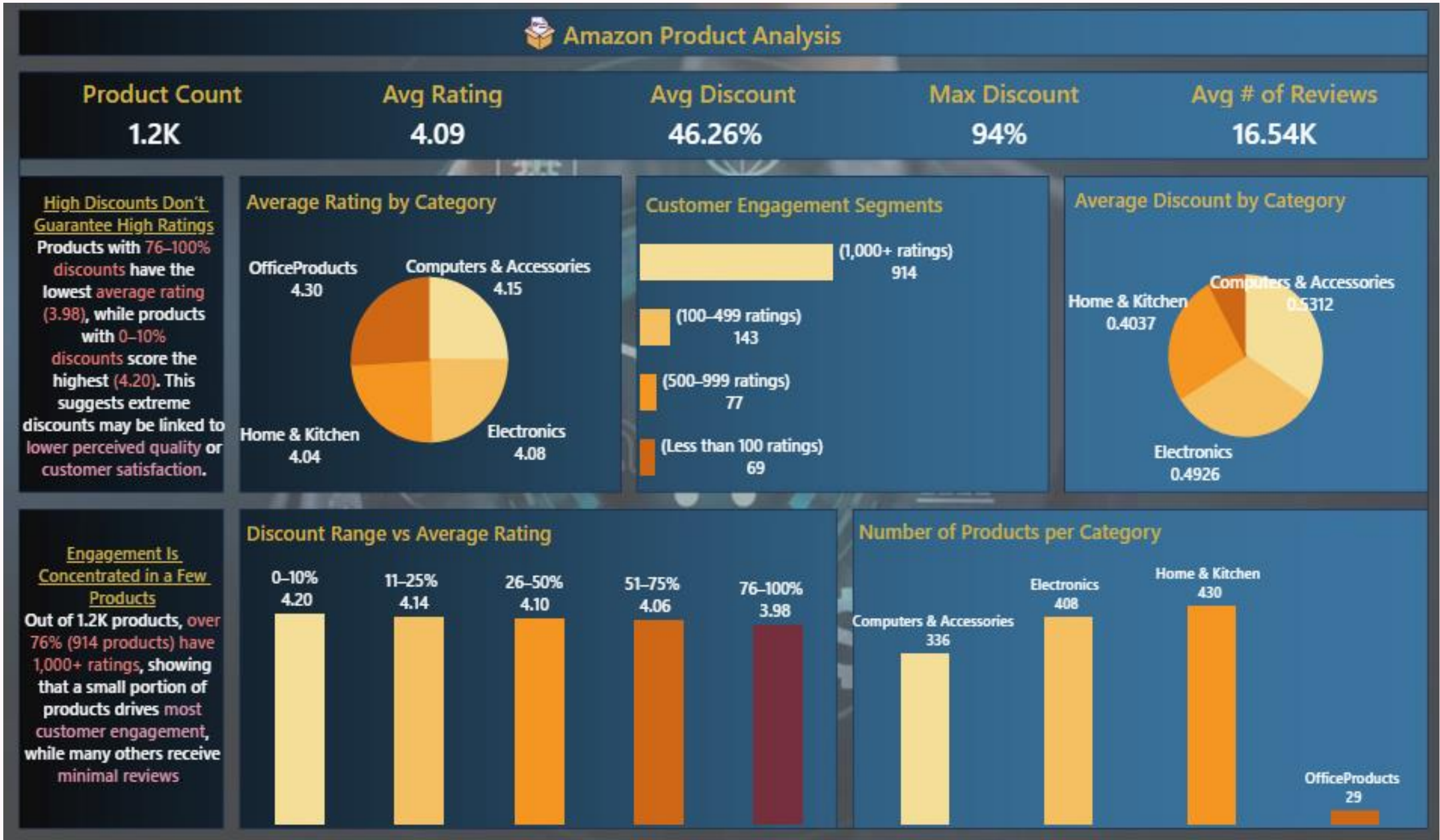
- Order ID — Unique identifier for each order
- Product Name — Name of the product sold
- Category — Product category (e.g., Electronics, Clothing)
- Quantity — Number of units sold
- Price — Sale price per unit
- Discount Percentage
- Discounted Price

This dataset provides a comprehensive overview of Amazon's product sales, enabling analysis of sales trends, customer preferences, and revenue patterns.



Visualizations & Dashboard

The interactive dashboard summarizes key insights from the Amazon Sales dataset, focusing on product performance, pricing, discounts, customer engagement, and category-level analysis.



- Key Features:**
- **Summary Metrics:** Shows total product count (1.2K), average rating (4.09), average discount (46.26%), maximum discount (94%), and average number of reviews (16.54K).
 - **Category Insights:**
 - *Average Rating by Category* highlights Electronics, Home & Kitchen, Office Products, and Computers & Accessories.
 - *Average Discount by Category* compares discount strategies across categories.
 - **Customer Engagement:** Segments products by number of reviews, showing that most engagement is concentrated in a small set of products.
 - **Discount vs. Ratings:** Reveals that high discounts (76–100%) often correlate with lower ratings, while low discounts (0–10%) score the highest.
 - **Product Distribution:** Displays the number of products per category, with Electronics and Home & Kitchen leading in count.
 - **Purpose:** This dashboard enables quick identification of top-performing categories, the effect of discounts on ratings, and customer engagement patterns, supporting data-driven decisions on pricing and product strategy.

Data Cleaning & Preparation

Why Clean the Data?

To ensure data accuracy, consistency, and relevance for meaningful analysis.

Steps Taken:

- Converted actual_price and discounted_price columns to USD format
- Transformed discount_percentage from string format (e.g., "15%") to decimal (0.15)
- Removed unnecessary columns such as about_product, user_id, review_id, and others
- Deleted duplicate records to avoid redundancy
- Fixed data entry errors for consistency
- Removed records irrelevant to analysis goals
- Shortened long product names for better readability

Result:

Cleaned dataset with **1,203 records** ready for SQL and Tableau analysis



Analysis & SQL Queries

SQL queries were used to explore, summarize, and extract insights from the cleaned dataset. The queries focused on identifying sales trends, top-performing products, and customer purchasing patterns.

1. Top 5 Most Reviewed Products Overall

Purpose: To find the 5 products with the most reviews, showing which items attract the most customer attention.

product_name	rating_count	rating	category
AmazonBasics Flexible Pre...	426973	4.4	Electronics
Amazon Basics High-Speed ...	426973	4.4	Electronics
boAt Bassheads 100 in Ear...	363713	4.1	Electronics
boAt BassHeads 100 in-Ear...	363711	4.1	Electronics
Redmi 9 Activ (Carbon Bla...	313836	4.1	Electronics

2. Bottom 5 Low-Rated Products with High Review Counts

Purpose: To identify 5 products with many reviews but low ratings, highlighting items with potential quality or satisfaction issues.

product_name	rating	rating_count	category
Eureka Forbes car Vac 100...	2.5	992	Home & Kitchen
MR. BRAND Portable USB Ju...	2.8	109	Home & Kitchen
IONIX Tap filter Multilay...	2.9	159	Home & Kitchen
7SEVEN Compatible LG TV R...	3	103	Electronics
Amazon Brand - Solimo 65W...	3	16500	Computers & Accessories

3. Average Discount by Category

Purpose: To calculate the average discount in each category, showing which categories offer the biggest price reductions.

category	avg_rating
OfficeProducts	4.3
Computers & Accessories	4.15
Electronics	4.08
Home & Kitchen	4.04

4. Bottom 10 Least Reviewed Products Overall

Purpose: To find the 10 products with the fewest reviews, showing items with low customer engagement.

product_name	rating_count	rating	category
Khaitan ORFin Fan heater ...	2	2	Home & Kitchen
NGI Store 2 Pieces Pet Ha...	2	3.1	Home & Kitchen
VAPJA?? Portable Mini Jui...	4	3.6	Home & Kitchen
Longway Blaze 2 Rod Quart...	4	3.7	Home & Kitchen
Lifelong LLQH925 Dyno Qua...	4	3.8	Home & Kitchen

5. Maximum Discount Given

Purpose: To identify the highest discount offered on any product, showing the maximum price reduction available.

Maximum Discount Given
0.94

Conclusion & Recommendations

Key Learnings:

The dataset analysis showed **1,200+ products** across four main categories, with an overall average rating of **4.09**.

- **High discounts (76–100%) do not guarantee better ratings** — instead, products with minimal discounts (0–10%) scored the highest average ratings (4.20).
- **Customer engagement is highly concentrated:** over 75% of products have more than 1,000 reviews, while many others receive very few, indicating a small group of products drives most attention.
- **Electronics and Home & Kitchen** dominate product counts and discounts, but Computers & Accessories and Office Products tend to achieve slightly higher average ratings.

Recommendations:

- **Optimize Discount Strategies:** Avoid excessive discounts that may reduce perceived quality. Focus instead on moderate, well-targeted discounts.
- **Boost Low-Engagement Products:** Increase visibility of products with few reviews through marketing campaigns, bundles, or review incentives.
- **Leverage High-Rating Categories:** Promote Office Products and Computers & Accessories more strongly since they consistently score well.
- **Focus on Top Performers:** Prioritize products that already receive high engagement (thousands of reviews), as they drive most customer trust and revenue.
- **Continuous Monitoring:** Regularly track sales, ratings, and reviews in dashboards to adapt strategies based on real-time performance.

Technical Stack & Files

Tools used

Excel → Data cleaning & preparation

SQL (MySQL) → Data querying & analysis

Tableau → Visualization & dashboard creation

Project Files & Links

Tableau Dashboard Link

[Click here](#)