## **Amazon Product Performance Analysis**

Junior Data Analyst at RetailTech Insight

**Project Summary**

As part of the analytics team at RetailTech Insights, I was tasked with analyzing Amazon product and customer review data. The goal was to generate insights that can guide:

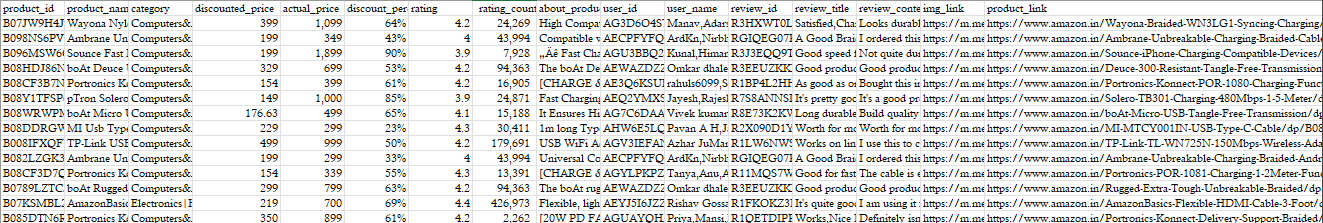
* Product improvement
* Marketing strategies
* Customer engagement

**Dataset Description**

Total Records: 1,465 rows

Total Columns: 16

Each row = 1 product

Data includes: product names, categories, prices, discounts, ratings, reviews, etc.

*Overview of raw dataset*

**Tools Used**

Microsoft Excel for:

* Data cleaning
* Calculated columns
* Pivot Tables & Charts
* Dashboard

**Data Cleaning (in Excel)**

* Changed data types to appropriate formats (e.g., number, text)
* Split the category column for consistency
* Removed 64 duplicate rows
* Corrected spelling errors
* Applied Proper Case formatting for readability
* Removed blank rows using filters

**Project Objectives**

To analyze product and review data to generate insights that can guide product improvement, marketing strategies, and customer engagement.

**Questions and Analysis**

**1. What is the average discount percentage by product category?**

Rows: category

Values: discount\_percentage → Average

Insight: Identify which categories give the highest average discounts.

A screenshot of a computer

AI-generated content may be incorrect.

**2. How many products are listed under each category?**

Rows: category

Values: product\_name → Count (or count of products)

Insight: Understand product spread across categories.

**3. What is the total number of reviews per category?**

Rows: category

Values: rating\_count → Sum

Insight: See which categories drive the most customer interaction.

**A screenshot of a computer

AI-generated content may be incorrect.**

**4. Which products have the highest average ratings?**

Rows: product\_name

Values: rating → Max or Average

Sort Descending

A screenshot of a computer

AI-generated content may be incorrect.Insight: Surface top-performing products.

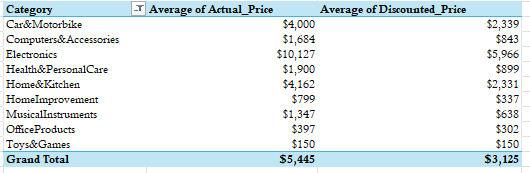
**5. What is the average actual price vs discounted price by category?**

Rows: category

Values: actual\_price → Average

Values: discounted\_price → Average

Insight: Compare real prices with offered deals across categories.



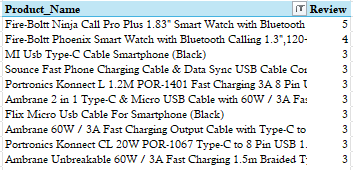
**6. Which products have the highest number of reviews?**

Rows: product\_name

Values: rating\_count → Max

Sort Descending

Insight: Identify popular products.



**7. How many products have a discount of 50% or more?**

Method:

Added a calculated column:

=COUNTIF(F:F, ">=50%")

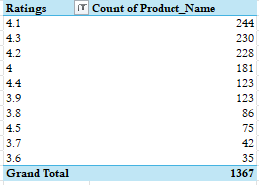
Insight: Know how many heavily discounted products exist.

**8. What is the distribution of product ratings?**

Rows: rating

Values: product\_name → Count

Optional: Create a bar chart

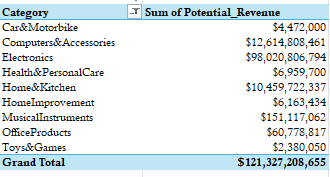
Insight: See how ratings are spread (e.g., how many 3.0, 4.5, 5.0)

**9. What is the total potential revenue by category?**

Calculated Column: = actual\_price \* rating\_count

Rows: category

Values: new column → Sum

Insight: Estimate revenue potential by category.

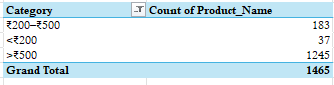
**10. Number of unique products per price range bucket**

Added a column for price bucket:

=IF(actual\_price<200,"<₹200",IF(actual\_price<=500,"₹200–₹500",">₹500"))

Rows: new bucket column

Values: product\_name → Count

Insight: Segment product pricing.

**11. How does the rating relate to the level of discount?**

Insert a Scatter Plot:

Select: rating and discount\_percentage

Insert → Chart → Scatter Plot

Insight: Look for any visible trend (e.g., higher ratings = less discount?).

**12. How many products have fewer than 1,000 reviews?**

Calculated Column:

=COUNTIF(H:H, “<1000”)

Insight: Know how many low-engagement products exist.

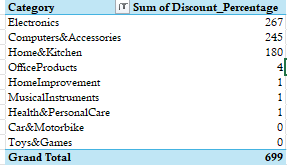
**13. Which categories have products with the highest discounts?**

Pivot Setup:

Rows: category

Values: discount\_percentage → Max

Sort Descending

Insight: Discover which categories offer the biggest discounts.

**14. Top 5 products by rating and number of reviews combined**

Calculated Column:

= rating + rating\_count

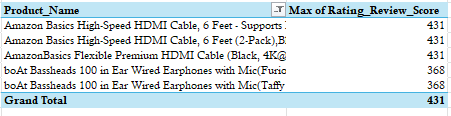
Then:

Pivot Table with product name

Values: New column → Max

Sort descending → Show Top 5

Insight: Find star products.



**KPIs Displayed on Dashboard**

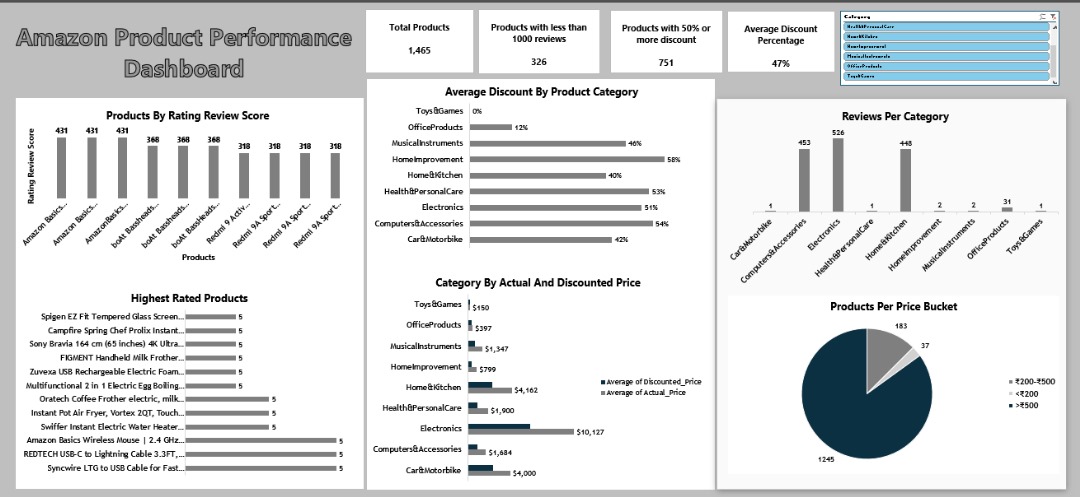
Avg Discount Average discount across all products

Total Products Total products analyzed

Products with 50% or more discount

Products with less than 1000 reviews

**Dashboard Overview**



**Insights Summary**

* Categories like Home Improvement and Health & Personal care offer the highest discounts
* Products in Electronics category get the most reviews.
* High-rated products don’t always need high discounts.
* Majority of products fall in the >₹500 price range.
* Customer engagement varies strongly across categories.

**Recommendations**

* Promote top-rated products with more reviews
* Improve engagement for low-review products with potential
* Optimize pricing for low revenue categories
* Highlight high-discount categories during marketing campaigns