PROJECT IX

June 26, 2025

AMAZON SALES ANALYSIS

Ananya Jha

OVERVIEW

In the age of digital retail, platforms like Amazon generate massive amounts of customer behavior and sales data. Understanding these data patterns is essential for businesses to optimize pricing, increase customer satisfaction, and identify growth categories. This project explores an e-commerce dataset resembling Amazon's sales ecosystem to uncover which product categories are performing best, how pricing impacts customer satisfaction, and what role discount strategies play in shaping product success.

The dataset includes details of over 1400 products, covering features such as category, original and discounted prices, ratings, rating count, and user reviews. The goal of the project is to simulate the role of a business analyst working for an online marketplace—extracting actionable insights that help improve revenue, retention, and customer experience.

PROJECT OBJECTIVE

- Identify top-performing product categories based on estimated revenue.
- Understand the relationship between discounts and customer ratings.
- Explore customer preferences across categories based on rating volume and satisfaction.
- Provide visual dashboards to assist sales and marketing teams in decision-making.

TOOLS USED

1. Microsoft Excel - for initial exploration and data cleaning
2. Power BI - for interactive data visualization

STAKEHOLDERS

1. Sales Team – to identify which product categories drive the most revenue
2. Marketing Team – to analyze the impact of discounts on customer engagement
3. Category Managers – to optimize product listings, pricing, and performance
4. Customer Experience Analysts – to monitor satisfaction across product types

DATA PREPARATION

- Loaded the raw dataset containing 1466 rows and 16 columns.
- Dropped non-analytical columns: about product, user id, user name, review id, review title, review content, img link, and product link.
- Created a new calculated column:

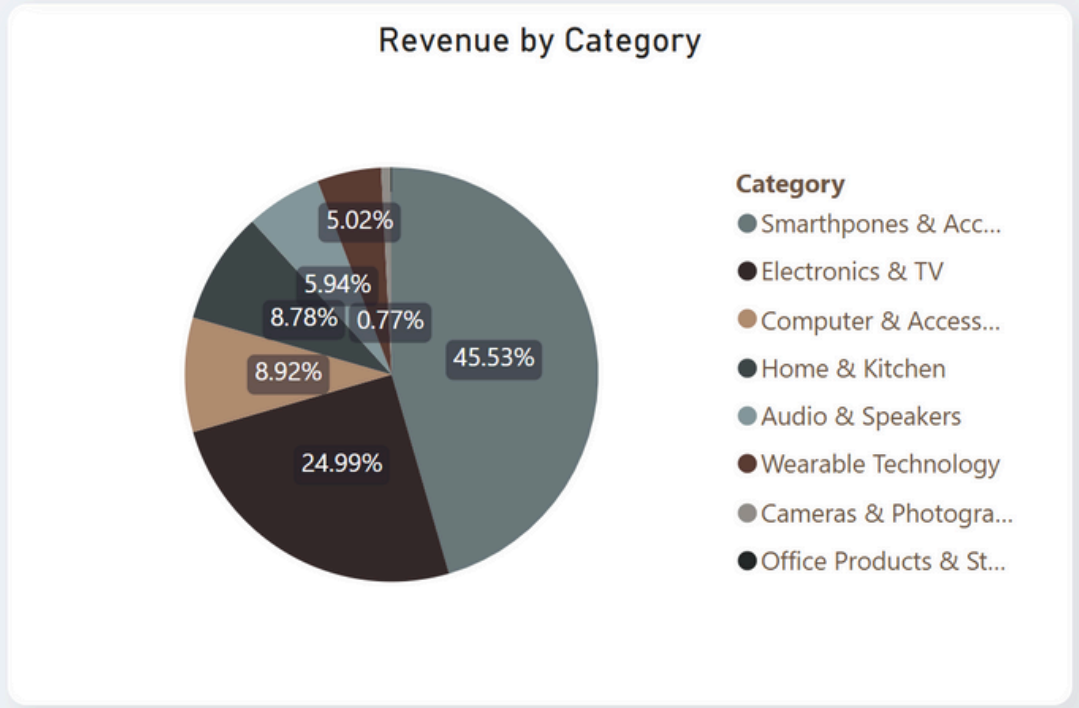
Estimated Revenue = Discounted Price × Rating Count

- Cleaned the discount percentage, actual price, and rating columns for data consistency.
- Filtered out rows with missing or zero ratings to ensure quality insights.
- Categorized product types for easier aggregation in dashboards.
- Imported final dataset into Power BI and built visual dashboards.

KEY VISUALIZATIONS AND INSIGHTS - 1

Revenue % by Category

- Tool: Power BI
- Type: Pie Chart



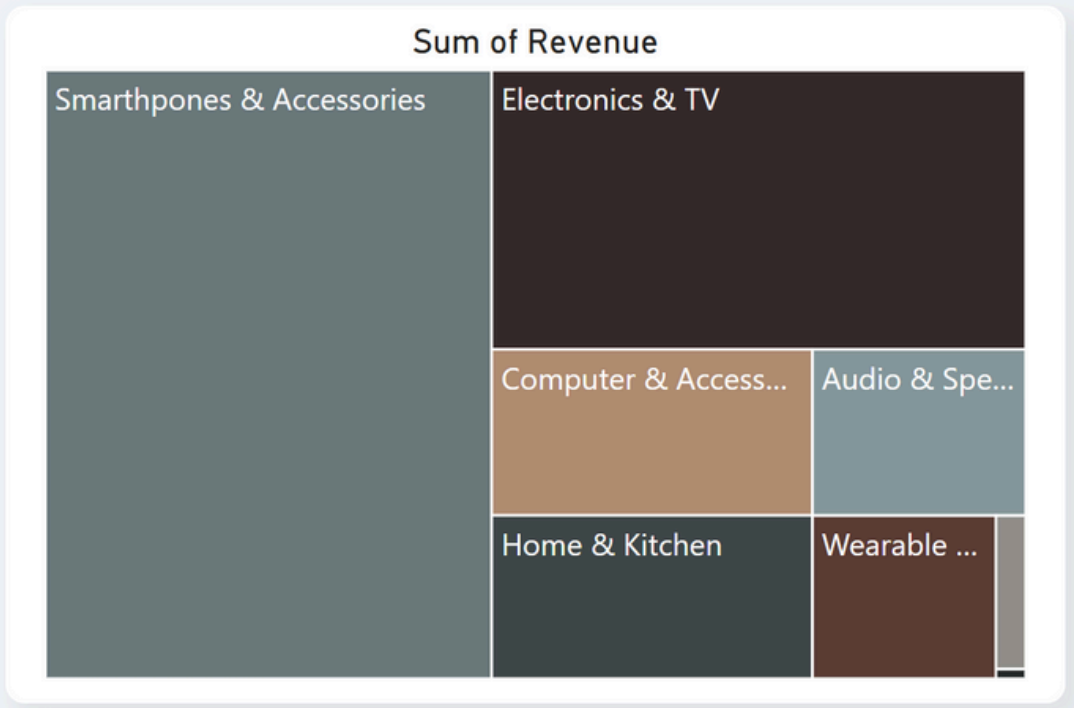
Insight:

- Smartphones & Accessories dominate revenue share at 45.53%, making them the platform's primary growth engine.
- Office products & Stationery contribute just 0.77%, indicating a potentially weak-performing segment.
- Other notable contributors include Electronics & TV (25%) and Computers & Accessories (8.9%), showing diversified tech-focused sales.

KEY VISUALIZATIONS AND INSIGHTS - 2

Net Revenue by Category

- Tool: Power BI
- Type: Tree Map



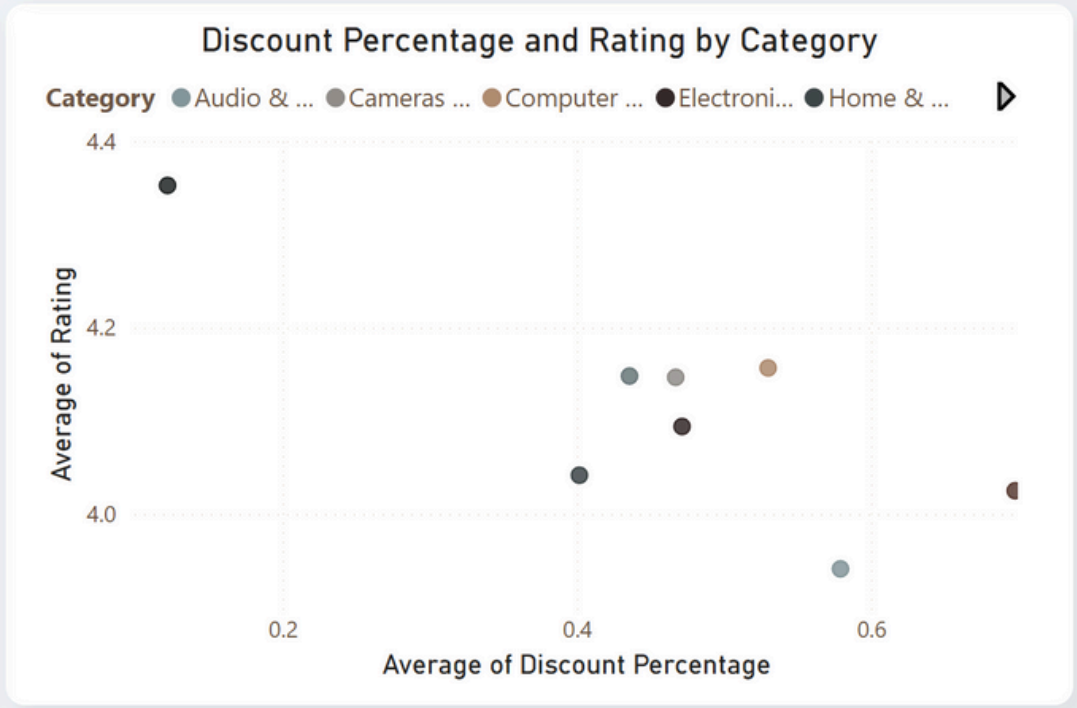
Insight:

- The tree map visually confirms that Smartphones & Accessories occupy the largest revenue area, reinforcing their dominance.
- Mid-sized boxes like Computers & Accessories and Electronics & TV show secondary contributors with healthy revenues.
- Smaller segments (e.g., Office Products & Stationery) are visually negligible, suggesting low financial impact.

KEY VISUALIZATIONS AND INSIGHTS - 3

Rating vs Discount %

- Tool: Power BI
- Type: Scatter Plot



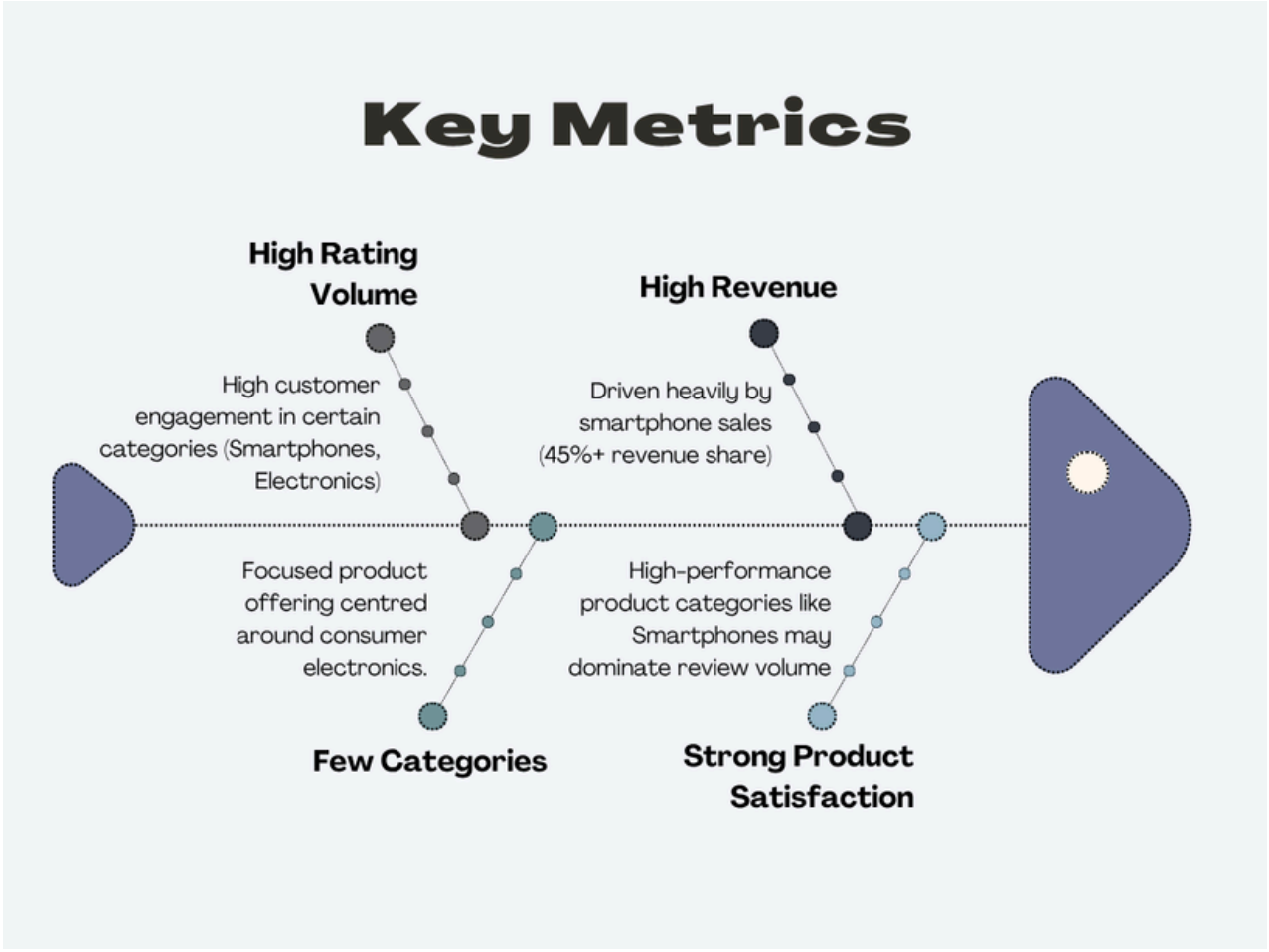
Insight:

- Office Products & Stationery have the highest rating (4.2) with the lowest discount, showing value-driven customer satisfaction.
- Wearable Technology show the highest discount (70%) but only moderate rating (4.03), suggesting aggressive promotions with limited perception gain.
- Audio & Speakers perform poorly, with the lowest rating (3.94) despite significant discounting (58%), indicating possible quality or expectation gaps.

ROOT CAUSE ANALYSIS

Insight 1: Key Metrics Root Cause

- Tool: Canva



Actionable Fixes:

- Expand category offerings
- Use customer insights from high-rated products
- Introduce category performance scorecards

ROOT CAUSE ANALYSIS

Insight 2: Scatter Plot - Rating vs Discount %

- Tool: Canva

"Audio products have the lowest average rating (3.94) despite offering a higher discount (0.58)"

5 WHYS ROOT CAUSE CHAIN

WHY 1:
Because customers are not satisfied with the product quality.

WHY 2:
Because the products may have performance issues or fail to meet expectations (e.g., sound quality, durability).

WHY 3:
Because sellers may be overpricing low-quality items and compensating through higher discounts.

WHY 4:
Because of poor supplier standards or lack of quality control in this category.

WHY 5:
Because the platform may prioritize sales volume over product reviews and feedback in discount campaigns.

Root Cause: Customers unsatisfied with audio products due to overpricing of low-quality items and poor supplier standards.

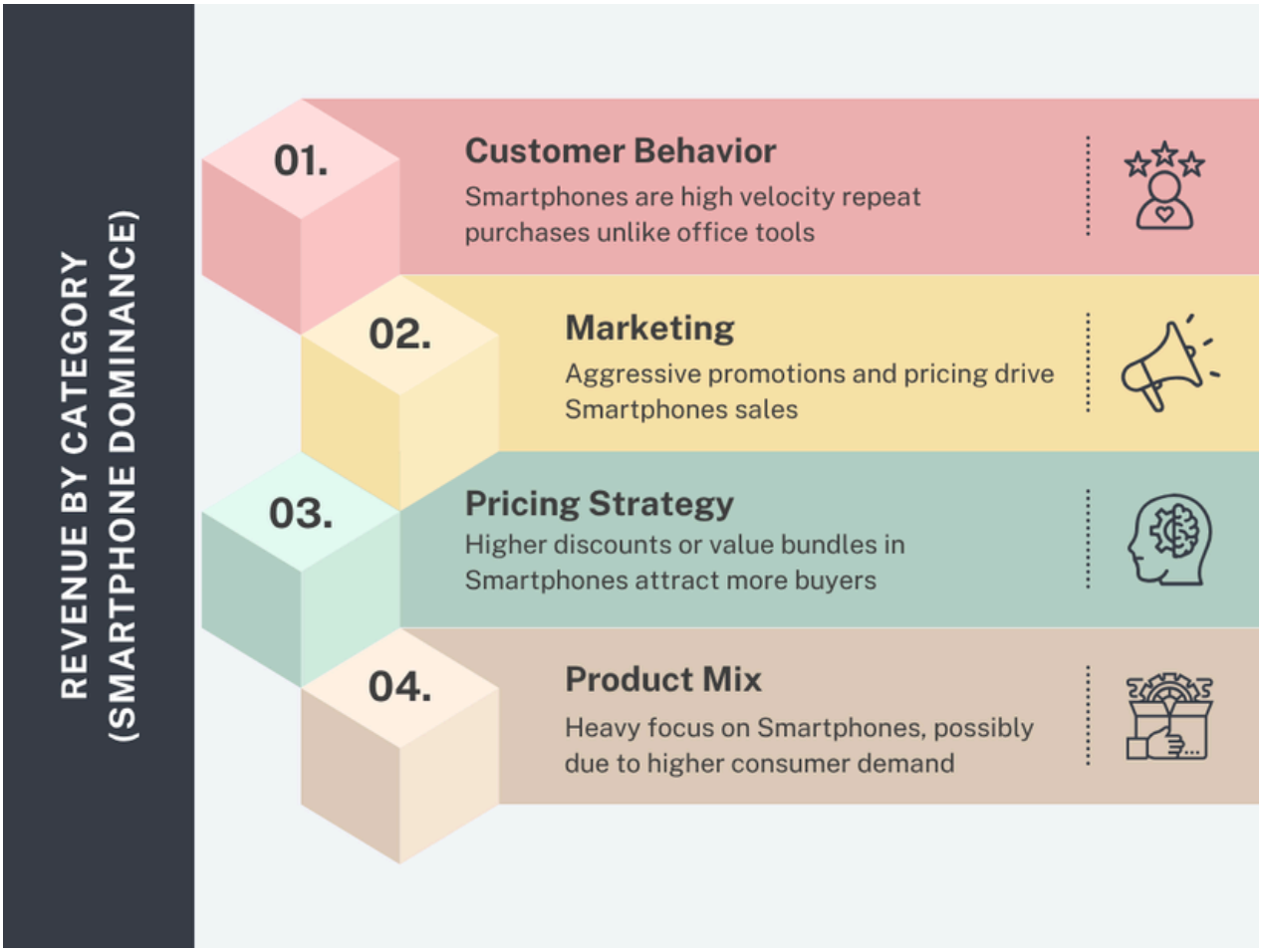
Solution:

- Audit audio product listings
- Set stricter seller standards
- Bundle offers with verified high-rated products

ROOT CAUSE ANALYSIS

Insight 3: Overdependence on Smartphone Sales

- Tool: Canva



Actionable Fixes:

- Set category revenue targets
- Highlight non-smartphone products in homepages, flash deals, and ad slots
- Launch curated bundles or lifestyle kits

CONCLUSION

This project provided a deep dive into sales performance, customer satisfaction, and pricing dynamics across key product categories in an e-commerce environment modelled on Amazon. Using visual analytics in Power BI, the analysis uncovered a clear overdependence on smartphone sales, inconsistencies in how discounts influence customer satisfaction, and strong but uneven product ratings across categories.

The analysis revealed a significant over-reliance on the smartphone category, which alone accounts for nearly half of the total estimated revenue. While this highlights its commercial strength, it also poses a strategic vulnerability if demand in that category declines or becomes saturated. On the other end of the spectrum, categories like Office Products, despite having the highest customer satisfaction, are underperforming in terms of visibility and revenue. This imbalance suggests that high-quality offerings may be overlooked due to weak promotional strategies or limited discoverability.

By transforming raw product-level data into a structured and visual narrative, the project fulfilled a key objective of simulating the real responsibilities of a Business Analyst. The findings are not just data points—they serve as decision-making inputs for sales, marketing, and category management teams to act on. This project reaffirms that data must not be viewed in isolation but as part of a larger business ecosystem, where every insight should lead to a tangible, actionable improvement.

CLOSING STATEMENT

This analysis strengthens the understanding that meaningful revenue growth in e-commerce lies not only in product pricing or discounting, but in balancing customer experience, category focus, and strategic execution across the board.

SELF-REFLECTION

This project has been a powerful milestone in my journey toward becoming a data-literate, decision-oriented business analyst. It pushed me to understand and apply analytical thinking to a real-world-like dataset where the "right answers" are not obvious.

Working with Power BI gave me a clearer sense of how business dashboards are built, customized, and used in professional environments. I learned how to create visuals that don't just "look good," but serve a purpose—whether it's highlighting a dominant category, spotting weak performers, or discovering broken relationships between expectations and results. The scatter plot analysis especially changed how I think about discounting: I used to assume more discount = happier customers. But this project taught me that over-discounting can harm perceived value and that customer experience is built on much more than price.

I also developed a stronger appreciation for the importance of data cleaning and modeling. Even a small typo like "Electronic & TV" versus "Electronics & TV" impacted my entire dashboard until corrected. This experience reminded me that attention to detail, and ownership of the data lifecycle, is just as important as analysis itself.

Most importantly, this project taught me how to tell stories with data—stories that persuade, inform, and lead to decisions. It felt like solving a business puzzle using evidence, logic, and curiosity. As I move forward in my analyst journey, this project will serve as a strong foundational example of what I can build, question, and improve using only tools, time, and structured thinking. I'm proud of how far I've come, and excited about how much further I can go.