# 

# 

# Customer Service and Product Rating Recommendations for Amazon

Bishoy Soliman Hanna, Ananya Gade, Dylan Hoffman, Akshay Shinde, Vagmi Bhagavathula

Business Intelligence

5 December 2023

## 

[**Customer Service and Product Rating Recommendations for Amazon 1**](#_edpztrppvl4)

[Introduction: 3](#_hmf9naykodh2)

[Understanding Business Intelligence: 4](#_qigj67v1mt7x)

[Expected Value of BI for Amazon: 4](#_q3q3asmkpxmp)

[Proposed BI Solution Architecture for Amazon 5](#_mg4mepk3wgj2)

[Relevant Case Study: 6](#_h5ydgxj0u0rf)

[Datasets We Worked On: 6](#_i6ogm4blfb6x)

[Key Performance Indicators: 7](#_a9ekr143dm64)

[Total Revenue Per Category: 7](#_dpzczjt5ejpz)

[Product Performance: 7](#_38dzxdoywtl0)

[Inventory Stock Status: 9](#_mh0iu6hb77go)

[Total Sales by Discount: 9](#_fow256yunjoi)

[Dashboards: 10](#_c5xnghsez6ks)

[Recommendations: 12](#_jw2nxi87rqqd)

[Limitations: 12](#_h47gseilq9mg)

[Conclusion: 13](#_1wdohiul4jnq)

[References: 14](#_qhc1f3nvf0u)

## 

## Introduction:

Amazon is a multinational company in Seattle, Washington, specializing in e-commerce. Established in 1994, Amazon is known for several other components, such as cloud computing, artificial intelligence, and online streaming services. Their geographic market area has no range since Amazon’s main use of service exists in e-commerce, which is what our project’s scope lies. An enterprise expected to exceed $1.5 TRILLION in market cap, Amazon has cemented itself as the cream of the crop, having over 1.68 million employees. Amazon has over 40 departments with about 100 internal teams within those departments.

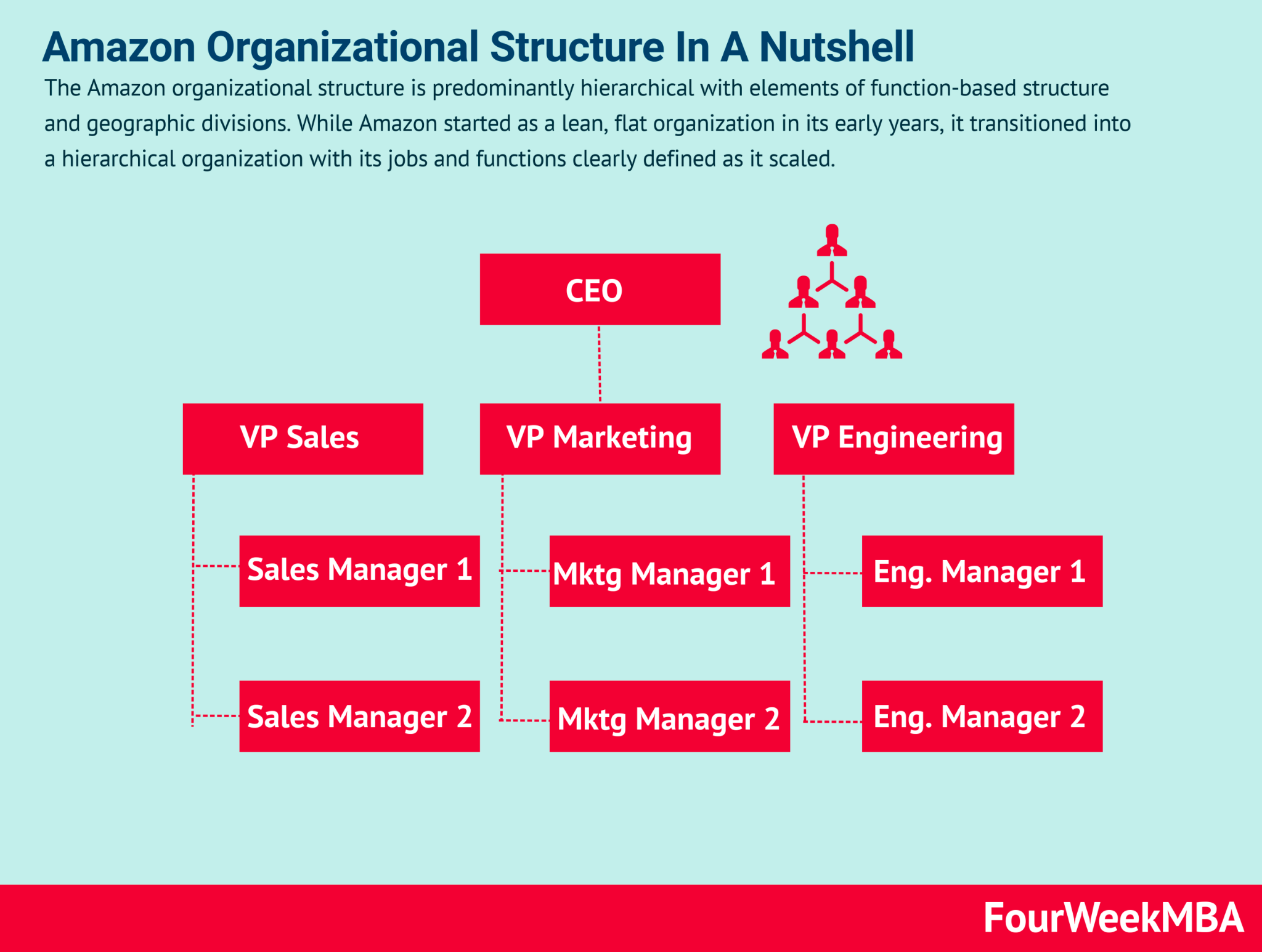


Figure 1: Amazon hierarchy example from FourWeekMBA (<https://fourweekmba.com/amazon-organizational-structure/>)

Given how robust the company is, it is safe to say Amazon does incorporate business intelligence processes in its decision-making. They do have business reports saved in the format of comma-separated values (CSV), that can easily be converted into Microsoft Excel. Amazon does indeed use data warehouses, but it operates under a cloud-computing platform known as Amazon Web Services. They have their data warehousing product known as Amazon Redshift. What makes Redshift different from other traditional data warehouses is the high-performance query processing and efficient storage, two aspects that traditional data warehouses fail to achieve consistently. Amazon does have a separate BI team, consisting of several feature engineers and managers who oversee dashboards as well as provide a series of recommendations. The BI team operates in a highly collaborative environment, employing agile methodologies and continuous integration/continuous deployment (CI/CD) practices to ensure that insights are delivered swiftly and efficiently. Managers and employees can generate dashboards in Amazon Web Services, thanks to its fast cloud-computing and flexibility to implement diverse aspects of their data sources.

## Understanding Business Intelligence:

Business Intelligence (BI) is a technology-driven process for analyzing data and presenting actionable information to aid in corporate decision-making. By aggregating, processing, and analyzing large datasets, BI systems provide historical, current, and predictive views of business operations. The value generated through BI is multifaceted, encompassing enhanced decision-making, improved operational efficiency, and a deeper understanding of market dynamics.

## Expected Value of BI for Amazon:

For Amazon, a leader in e-commerce and cloud computing, BI is not just a tool but a strategic asset. Amazon is known to use an array of other BI tools and technologies to handle its diverse data analytics needs. For instance, Amazon incorporates QuickSight, Tableau, and Looker that complement its analytics arsenal. The implementation of BI can lead to significant advancements in various aspects of the business:

* **Optimized Inventory Management**: By predicting purchasing trends and customer demand, BI can help Amazon maintain optimal stock levels, reducing holding costs and minimizing stockouts.
* **Customer Experience Personalization**: Analyzing customer data enables Amazon to offer personalized shopping experiences, recommendations, and targeted marketing campaigns, which can increase customer loyalty and sales.
* **Supply Chain Efficiency**: BI technologies can pinpoint inefficiencies and bottlenecks in the supply chain, allowing for streamlined logistics and cost savings.
* **Market Trend Analysis**: In the rapidly changing retail landscape, BI tools help Amazon stay ahead by identifying emerging market trends and consumer preferences.

## Proposed BI Solution Architecture for Amazon

The proposed BI architecture for Amazon is designed around the BI framework introduced in class, comprising the following components:

* **Operational Databases**: Amazon's operational databases are likely to include transaction processing systems that capture data from its e-commerce platform in real-time, facilitating immediate order processing and customer service interactions.
* **External Data Sources**: These may consist of market research data, social media trends, and economic indicators that provide insights into consumer behavior and market conditions.
* **Data Warehouse and Data Marts**: A centralized data warehouse, potentially utilizing Amazon Redshift, acts as the repository for integrated data from various sources. Data marts cater to specific analytical needs of different departments, such as sales, marketing, and logistics.
* **ETL Processes**: Amazon would employ robust ETL (Extract, Transform, Load) processes to ensure data quality and readiness for analysis. Tools like AWS Glue could automate the preparation and loading of data into the warehouse.
* **Data Analytics and Reporting Tools**: Utilizing tools like AWS QuickSight and Tableau, Amazon's BI system would offer powerful analytics and reporting capabilities. These tools can generate real-time dashboards and reports for performance monitoring.
* **Advanced Analytics**: Amazon could leverage machine learning and predictive analytics, using technologies like AWS SageMaker, to forecast trends and automate decision-making processes.

## Relevant Case Study:

A case study to consider is **Walmart's** BI implementation. Like Amazon, Walmart is a retail giant that has successfully adopted BI to enhance its operations. Walmart's BI system integrates data from its 11,000+ stores to inform inventory decisions and sales strategies. The company has seen benefits from its BI investments, including improved supply chain efficiency and increased sales due to better product positioning and stock levels.

## Datasets We Worked On:

For the scope of our project, our team has decided to focus on e-commerce operations, due to its immense data generation and foundation for BI applications. The dashboard will be designed to cater to various stakeholders, from operational managers to marketing strategists. This project aims to prototype a BI dashboard that harnesses this potential for Amazon. The objective is multifaceted: to enhance Amazon's ability to monitor key performance indicators, streamline its vast operations, and foster an environment of data-driven decision-making.

We have incorporated two different datasets for our dashboard. The first data source focuses on customer satisfaction, whereas the other data source addresses product sales and ratings. Both were separate datasets and not combined because neither dataset has a common column. The first dataset contains 80,000 rows [[1]](#footnote-0) in it in which we can see the different products as well as some reviews for the products, and the sales per product. When cleaning this dataset we needed to make sure that we separated the products into main and subcategories because of the size of the dataset, so that we could work with the data. In the smaller dataset we had an issue in that under the product column of the dataset the product had several different categories that they had separated by a “|” symbol. In Tableau, we fixed this by doing a separation by using that character to separate each different category. This allowed us to work with the dataset better so that we could pick by each category in the KPIs we were using, so we could get as specific as we needed to not make us have hundreds of the same name for the product show up within one graph. Besides that, we also removed all the nulls that were in the dataset.

## Key Performance Indicators:

### Total Revenue Per Category:

Our first Key performance indicator is looking at total revenue per category of goods. We see that the category with the highest revenue is TV, audio, & cameras with $6.1B. Following that we have Appliances at $936.1M and Stores at $228.2M. The category with the least revenue was Home, Kitchen, and Pets at $2.5M. Overall it seems people are inclined to purchase their electronics and tech goods at Amazon. This KPI allows Amazon to see what its customers are purchasing. Some managerial questions that can be answered are: “Where to keep a greater stock of product?”; “How to market to increase the revenue of low-performing products?”. This information would be mainly used by executives, marketing teams, or store owners.

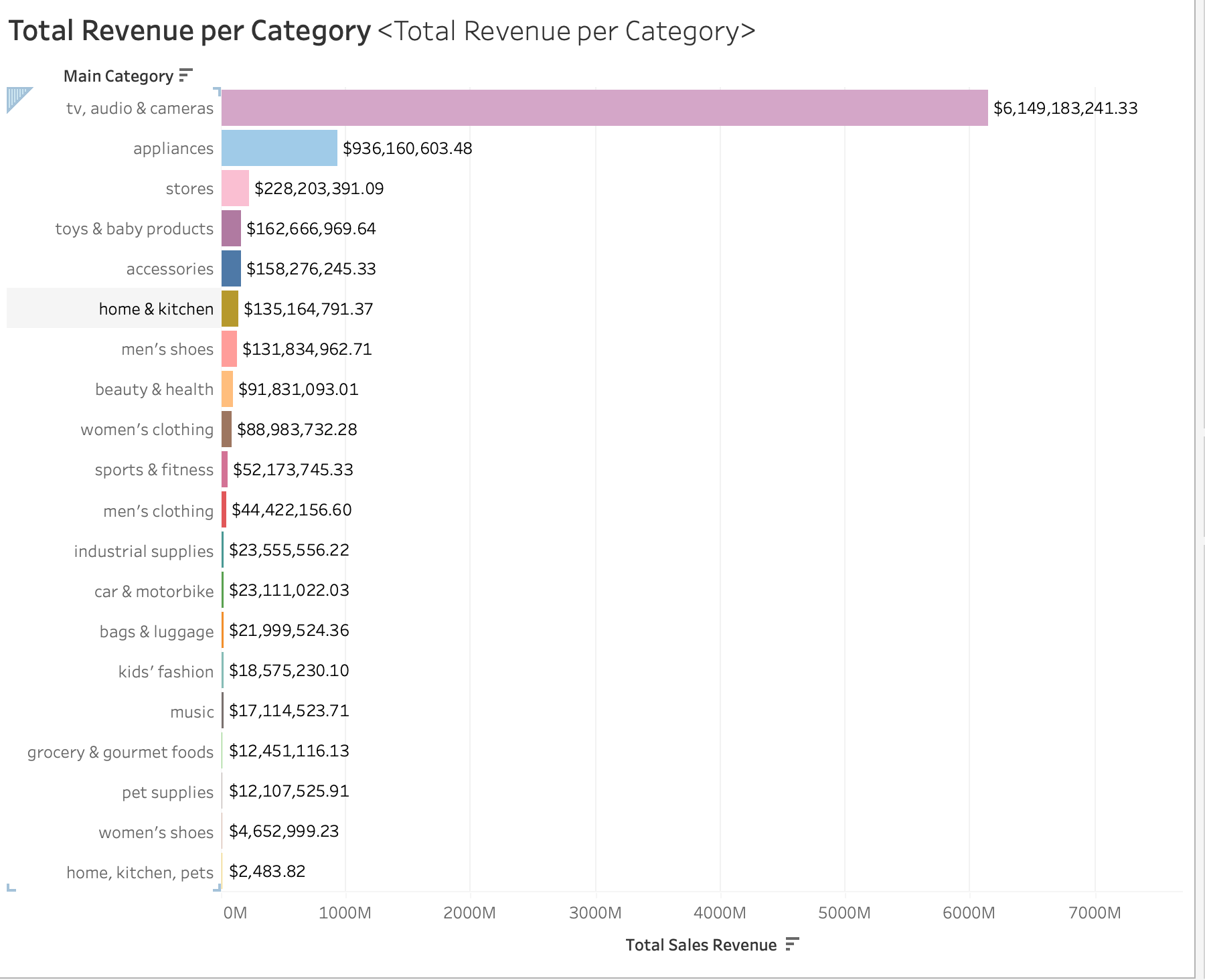


Figure 2: Total Revenue per Category

### Product Performance:

Now we are looking at the individual products sales in the category with the highest revenue. The Tv, Audio, & Cameras Category is split into 5 subcategories; All Electronics, Camera Accessories, Home Entertainment Systems, Speakers, and Televisions. The product that generated the greatest revenue overall was the SanDisk 400GB Class 10 MicroSD Card. The total sales for this product is 110.4M, and this MicroSD Card Currently retails for 57$ on Amazon. The product with the second greatest sales revenue is the SanDisk 512GB Memory Card, and the third was the SanDisk Extreme Pro 1TB Card. All the highest revenue sales fell under the Camera Accessories Category. This information can aid Amazon in managing its inventory, along with pricing to increase revenue. Here are the managerial questions that can be answered using this information: Wh

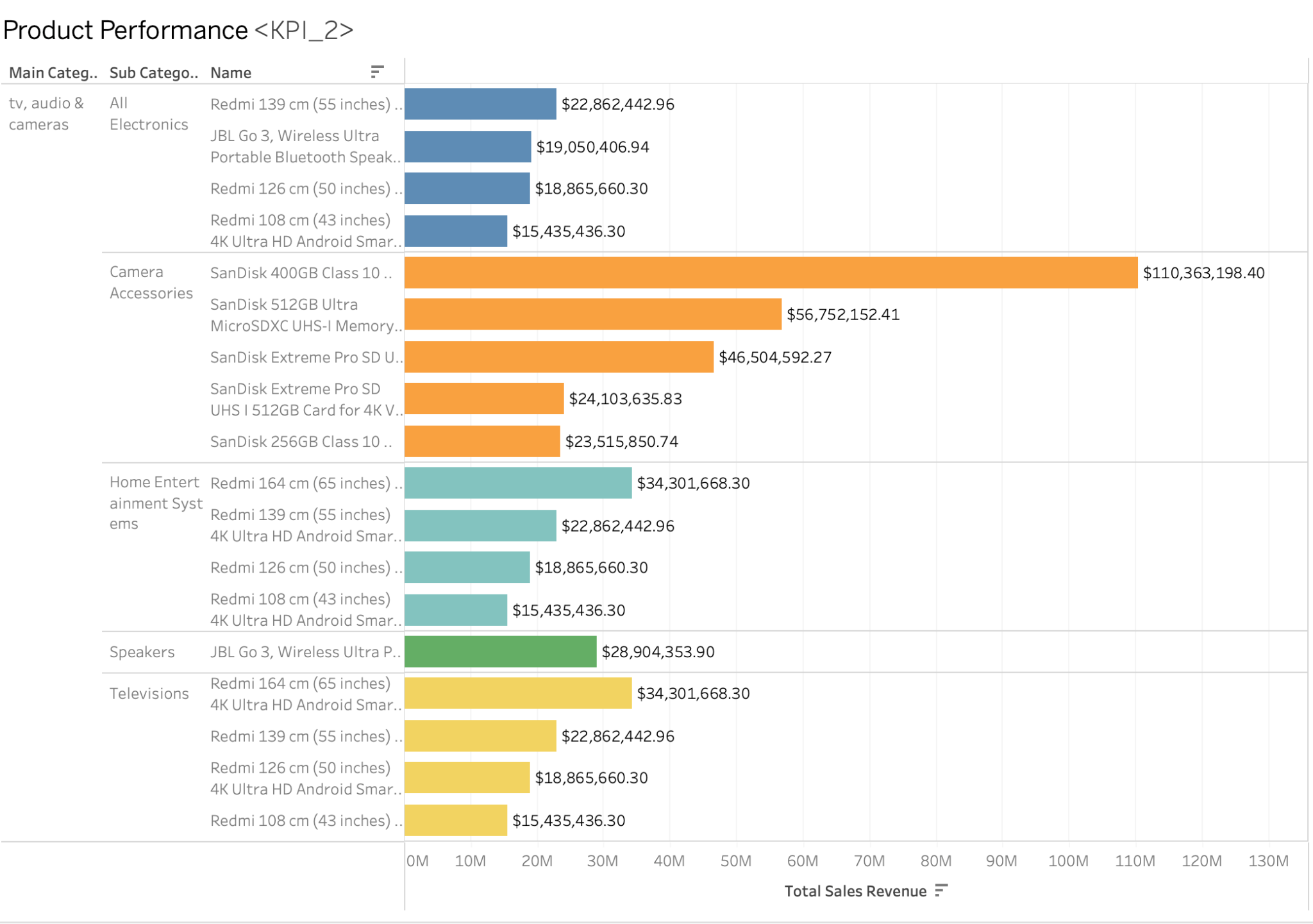


Figure 3: Product Performance

Customer Review and Ratings:

The main categories are broken down into subcategories with an average customer rating. Tablets have the highest overall rating of 4.6, followed by Power Accessories at 4.5 and Office Electronics at 4.45. The subcategory Cars Accessories have the lowest average rating of 3.8.

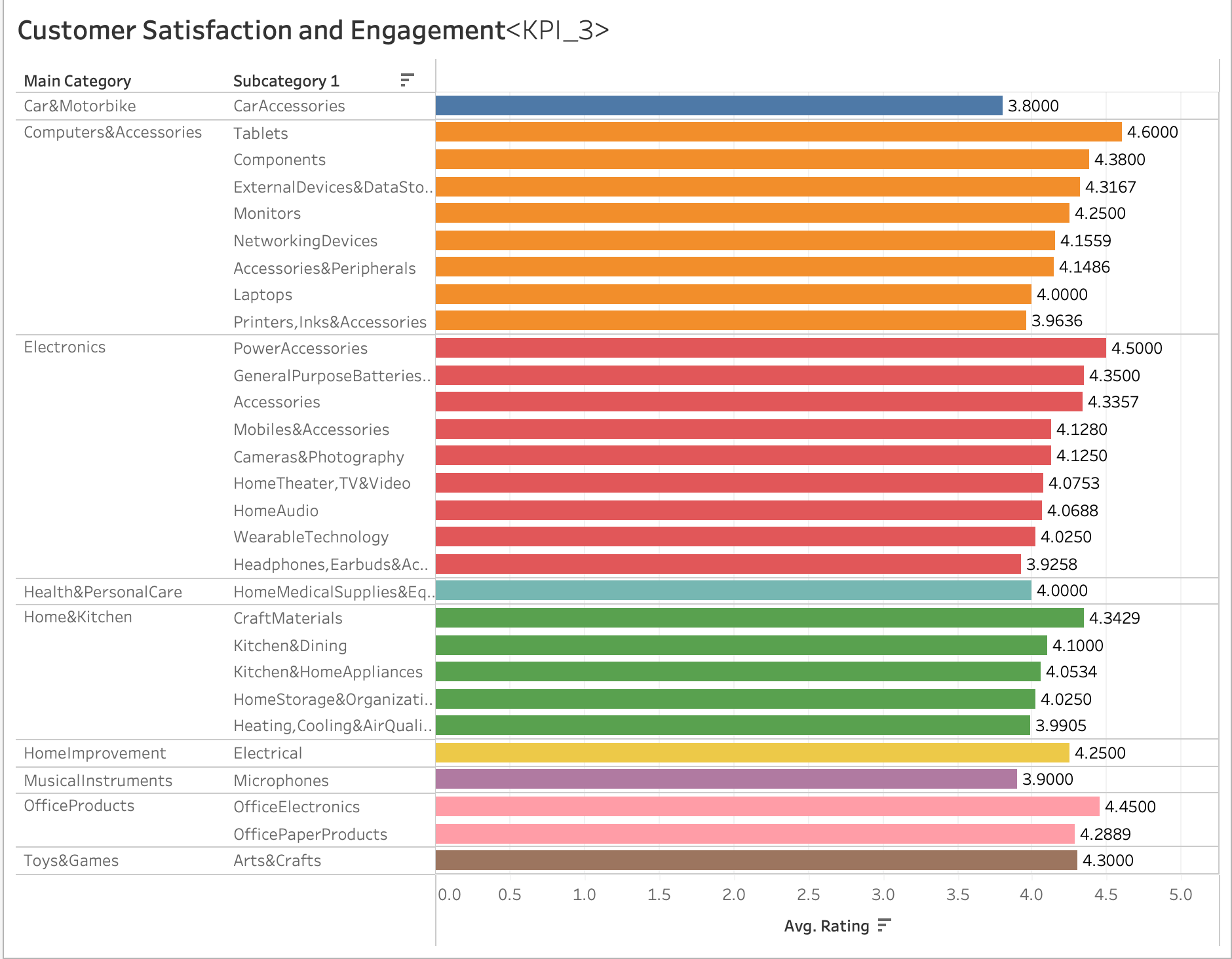


Figure 4: Customer Satisfaction and Engagement

### Inventory Stock Status:

The Inventory Stock Status shows the main category that's then broken up into multiple subcategories which are then broken down by the product itself. Each product's inventory is shown with either a low inventory symbol (orange) or a high inventory symbol (blue). Going through these we can see how each category varies in its distribution of inventory.

### Total Sales by Discount:

Looking at the Total Sales by Discount KPI, product sales are grouped based on the discount percentage of the sales. Products listed at a discount of 20%-30% created the most sales with $1.68 billion in sales. Overall marking a product at a discounted price between 10% and 60% yielded the majority of sales. Comparing this to the number of products in each discount group, 20-30% and 30-40% groups stand out with less over-products, but more overall sales. This shows that listing products at a 20-40% discount gives Amazon the best sales overall. This could be due to the marketing illusion that a product on sale at the same price as another product seems more desirable to consumers, leading to more sales.

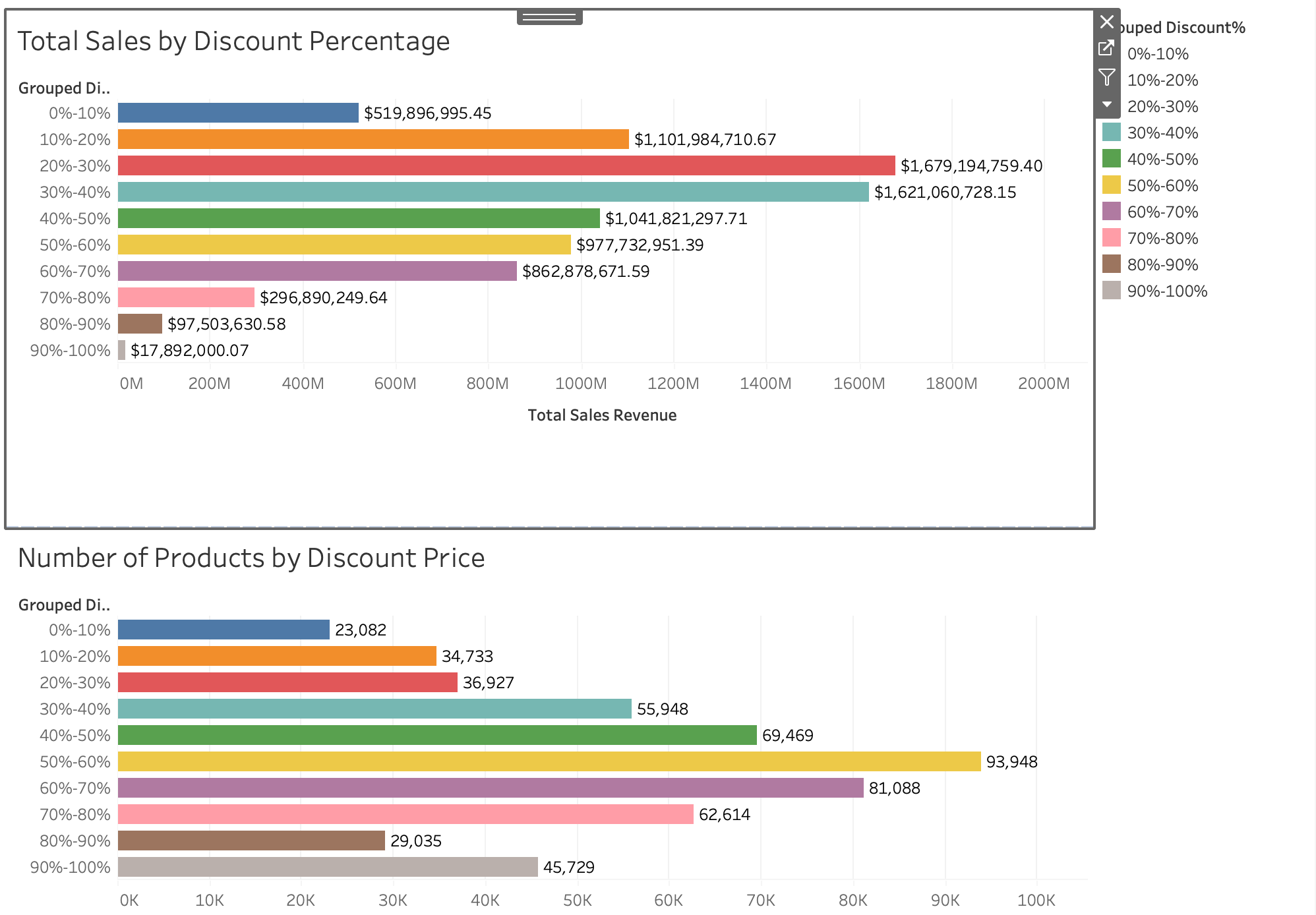


Figure 5: Total Sales By Discount Percentage and Number of Products by Discount Price

## Dashboards:

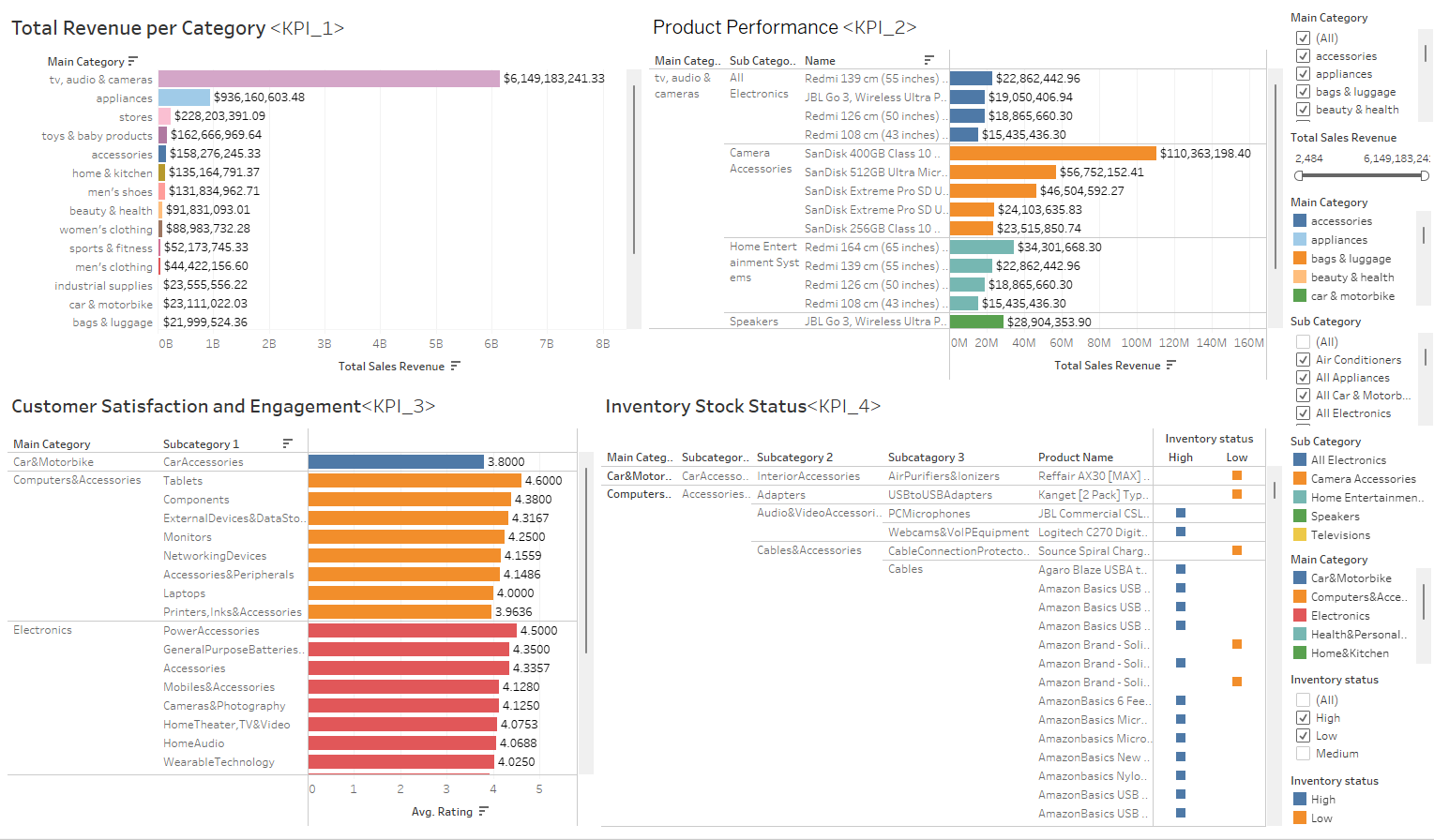


Figure 6: Dashboard 1

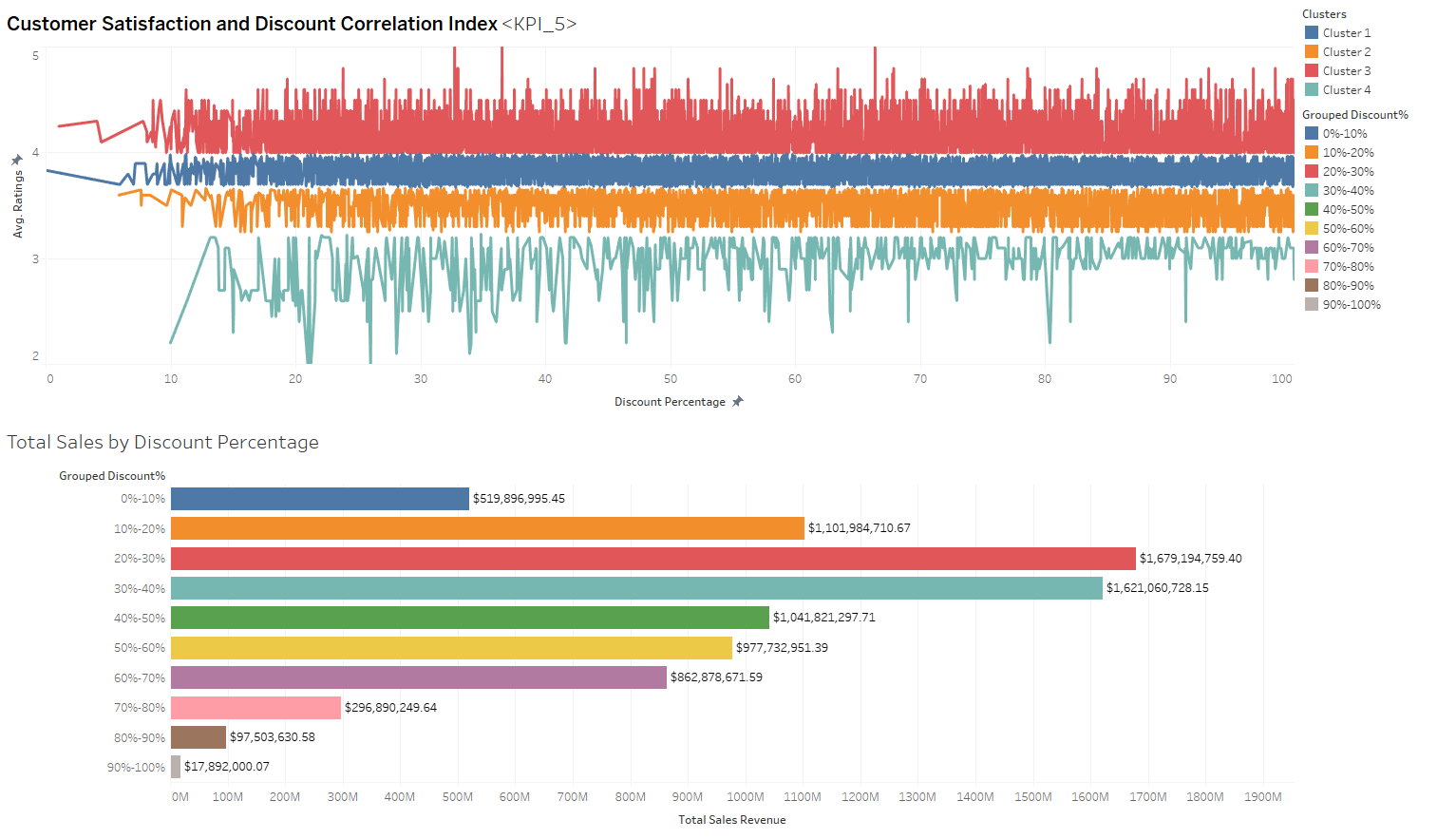


Figure 7: Dashboard 2

## 

## Recommendations:

Based on our dashboards, we can provide a series of recommendations. Our first recommendation is to combine additional data sources like customer demographics or economic indicators that line up with our proposed BI architecture solution. Next, we propose employing predictive modeling to forecast future trends in sales, customer satisfaction, and inventory needs. This way, Amazon can understand how well their models are performing on the products, which can allow them to evaluate if certain products are performing worse or better than the average. These machine learning algorithms can optimize customer personalization based on their buying and searching patterns. We recommend developing new KPIs that focus on emerging market trends like online engagement metrics, as a sort of baseline comparison relative to other major e-commerce companies. Applying some multiple logistic or linear regression models into another dashboard can be beneficial for understanding human tendencies.

## Limitations:

We encountered minor limitations in the study, some of which we were unable to overcome. One example was that in one dataset, we had to separate the product category into seven different categories to include them consistently in the dashboard. We also planned to use three different data sources, but the two we decided to measure already encompassed most of the third. The 3rd data source was created from a market location in India, considering the currency used was rupees. Thus, we decided to omit that dataset from the research. Lastly, as mentioned in the [Datasets](#_i6ogm4blfb6x) chapter, we were unable to merge the two datasets into one larger data source, since there was no common column to join on.

## Conclusion:

In conclusion, Amazon’s BI ecosystem transcends to more dynamic, robust BI platforms, such as AWS, Redshift, and tools part of their suites, which are more capable of handling large-scale data operations. Although they use MS Excel for their business reports, product performance, and customer service are effectively monitored on tools such as Tableau and Looker, providing advanced visualization capabilities that turn complex data sets into actionable insights. Our proposed BI architecture, revolving around operational databases and external data sources, is deemed to be successful based on the dashboards in Figures 6 and 7. Although we encountered some limitations, such as the inability to join the data sources, we were able to overcome them and provide a series of recommendations that Amazon may consider for improving their BI processes.

## References:

Hall, M. (2023, December 8). Amazon.com. Encyclopedia Britannica. <https://www.britannica.com/topic/Amazoncom>

Reuter, Dominick (30 July 2021). 1 out of every 153 American workers is an Amazon employee. Business Insider. <https://www.businessinsider.com/amazon-employees-number-1-of-153-us-workers-head-count-2021-7>

Parab, L. (11 March 2023). Amazon Products Sales Dataset 2023 (1) [Data set]. Kaggle. <https://www.kaggle.com/datasets/lokeshparab/amazon-products-dataset>

Karkavelraja, J. (2022). Amazon Sales Dataset (1) [Data set]. Kaggle. <https://www.kaggle.com/datasets/karkavelrajaj/amazon-sales-dataset>

DataFiniti (2018). Consumer Reviews of Amazon Products (1) [Sample Data set]. Kaggle. <https://www.kaggle.com/datasets/datafiniti/consumer-reviews-of-amazon-products?select=1429_1.csv>

1. [↑](#footnote-ref-0)