


Geekster

Amazon Sales EDA

Project by Ramtarun



1- Data Collection

 The dataset we'll be using has information on over 1000 products sold by Amazon, like their names, categories, prices, ratings, and reviews.

product_id - Product ID

product_name - Name of the Product

category - Category of the Product

discounted_price - Discounted Price of the Product

actual_price - Actual Price of the Product

discount_percentage - Percentage of Discount for the Product

rating - Rating of the Product

rating_count - Number of people who voted for the Amazon rating

about_product - Description about the Product

user_id - ID of the user who wrote review for the Product

user_name - Name of the user who wrote review for the Product

review_id - ID of the user review

review_title - Short review

review_content - Long review

img_link - Image Link of the Product

product_link - Official Website Link of the Product

2 -Data preparation

Before we dive into the data analysis and visualization, we need to make sure our dataset is clean and properly formatted. This process is called data preparation, and it involves several steps:

2.1 Data Inspection: We'll start by inspecting the dataset to see if there are any missing values, duplicates , or inconsistent data. We'll also check if the data types are correct and make sure the dataset is ready for analysis.

2.2 Data Cleaning: Next, we'll clean the dataset by removing or correcting any errors, inconsistencies, or irrelevant information. This will make the dataset more reliable and accurate.

2.3 Data Transformation: After cleaning the dataset, we may need to transform the data to make it more useful for analysis. This can include scaling, normalization, or feature engineering.

2.4 Data Saving: Once we've prepared the data, we'll save it in a new file to avoid overwriting the original dataset. This way, we can always go back to the original dataset if we need to.

By following these steps, we can ensure that our data is clean, accurate, and ready for analysis!

Overall, the data can help businesses identify the most popular subcategories and adjust their product offerings to meet customer demand. By focusing on these subcategories, businesses could increase their sales and improve their competitiveness in the market.

Summary and Key Insights

- Discounts: Discounts between 15-25% are most effective at driving customer engagement without sacrificing revenue.
- Category Performance: Electronics, especially smartphones, are the top-selling and most reviewed products.
- Customer Priorities: Smartphone buyers prioritize functionality, camera quality, and battery life.

These insights can guide businesses in tailoring their product offerings and marketing strategies to better align with customer preferences and increase sales.