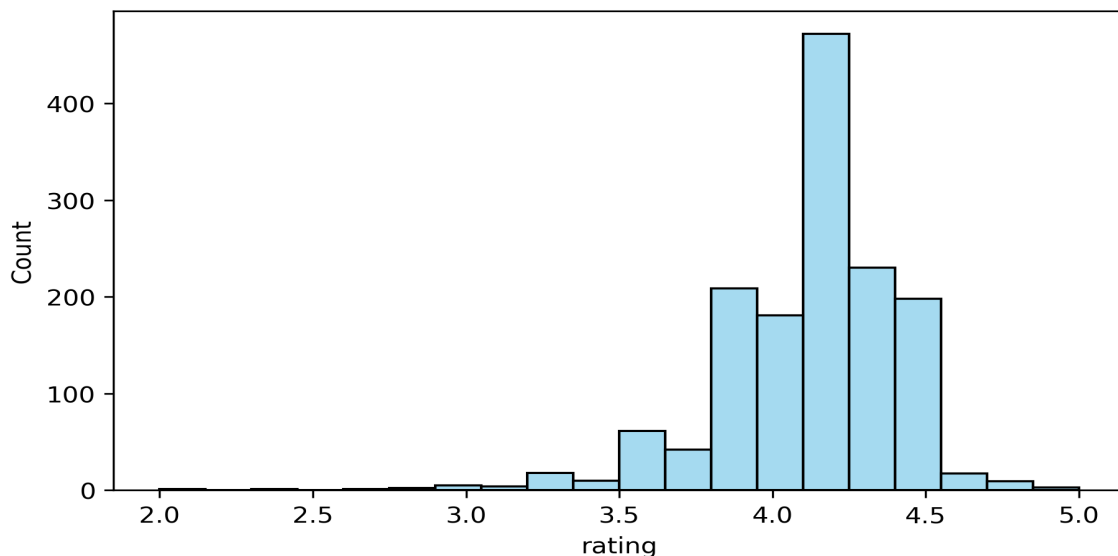# Amazon Products Exploratory Data Analysis

Data Summary

The dataset contains 1465 products with 17 features including product details, pricing information, and customer ratings. The data spans across multiple categories, with a focus on technology products.

Rating Distribution

The histogram below shows the distribution of product ratings:

Observation: Most product ratings fall between 3.5 and 4.5, indicating a generally positive customer experience. Very few products are rated below 3, which may suggest that low-rated products are removed or not reviewed often.
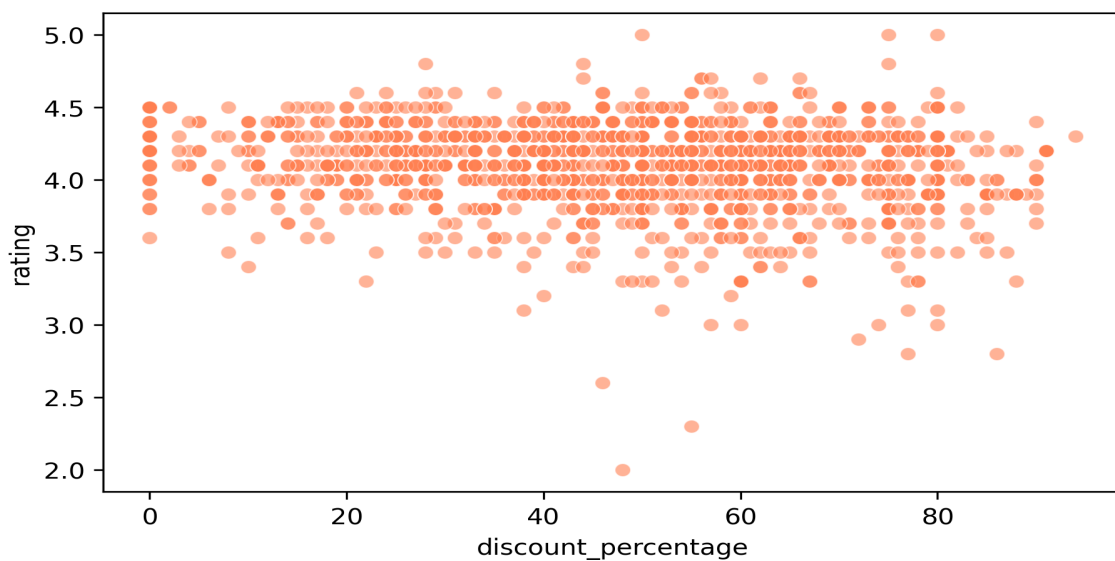
Discount Percentage Distribution

The boxplot below shows the spread of discount percentages offered:
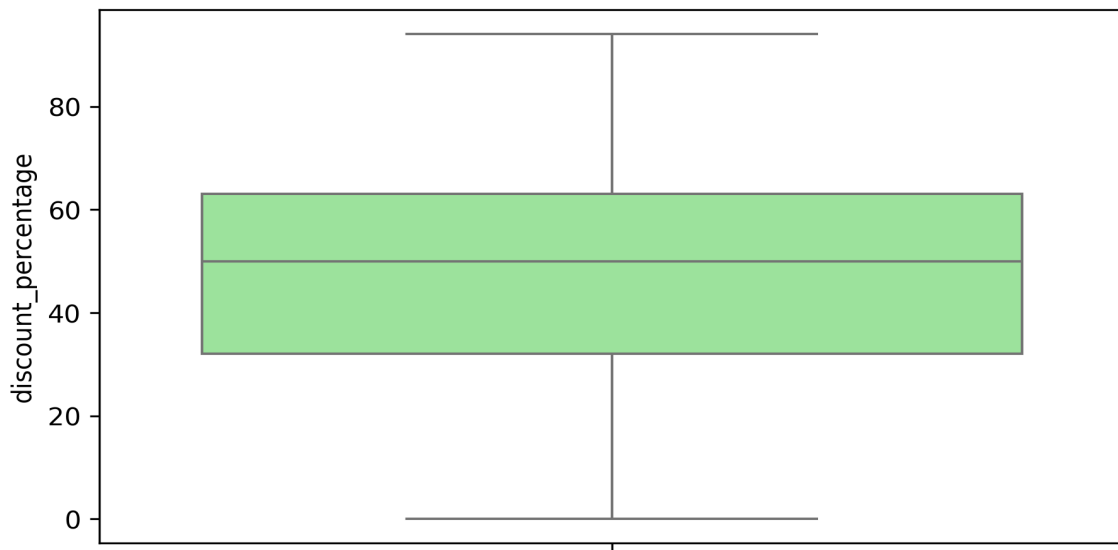
Discount vs Rating Analysis

This scatter plot explores the relationship between discount percentage and product rating:

Observation: No strong correlation between higher discounts and higher ratings. This suggests that price cuts do not necessarily improve product satisfaction, and product quality or brand reputation likely plays a larger role.

Actual Price vs Discounted Price

This scatter plot shows the relationship between original and discounted prices:
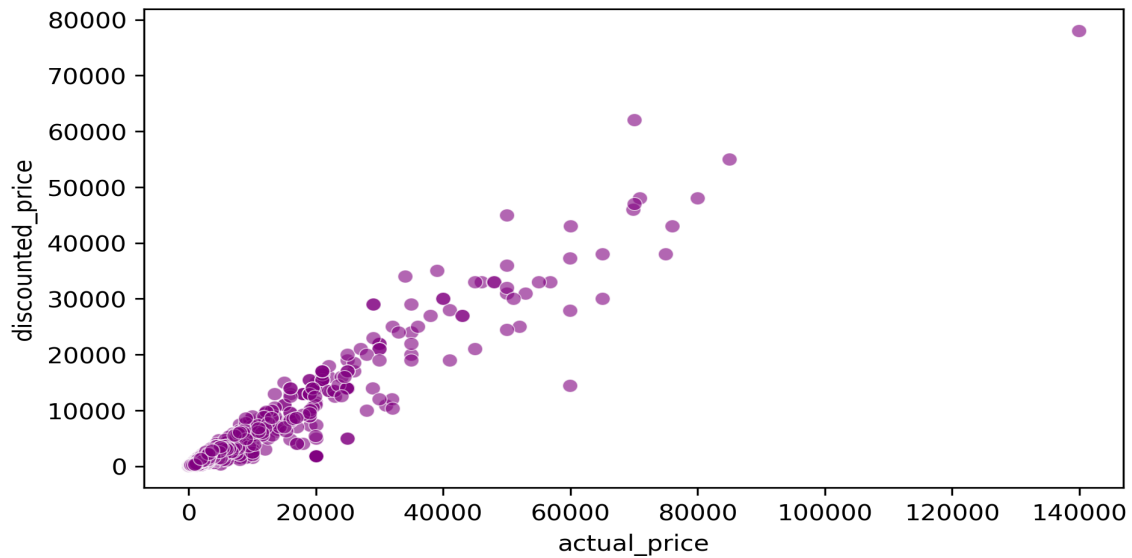
Observation: As expected, higher actual prices generally lead to higher discounted prices, but a few high-ticket items are deeply discounted. This helps identify overpriced products or luxury items on sale.

Top Product Categories

The bar chart below shows the most common product categories in the dataset:
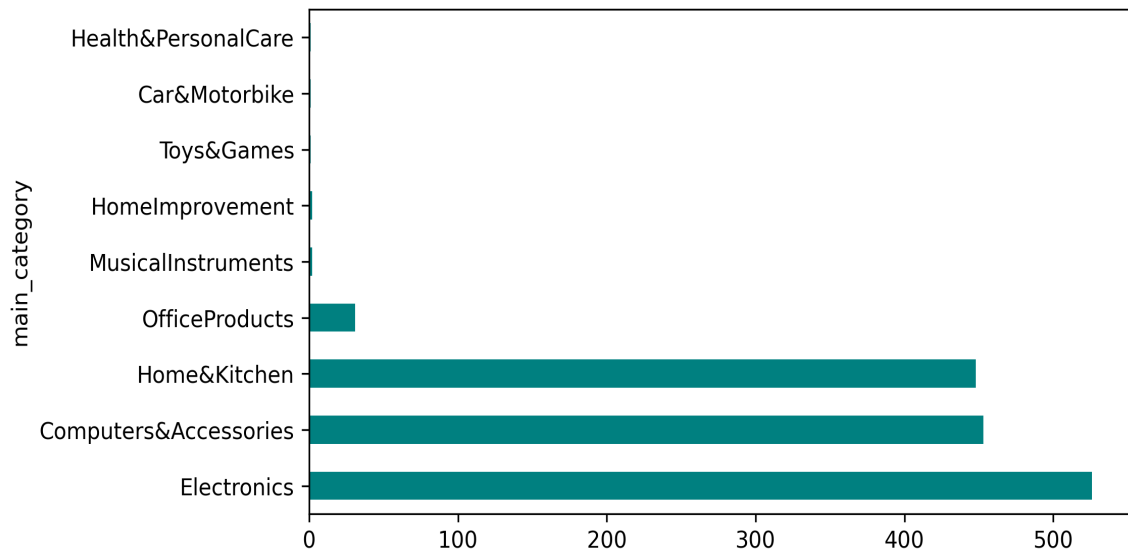
Observation: The Computers & Accessories category dominates the dataset. This indicates the dataset likely focuses on tech products, especially mobile and USB-related accessories.

Correlation Analysis

The heatmap below shows correlations between numerical variables:
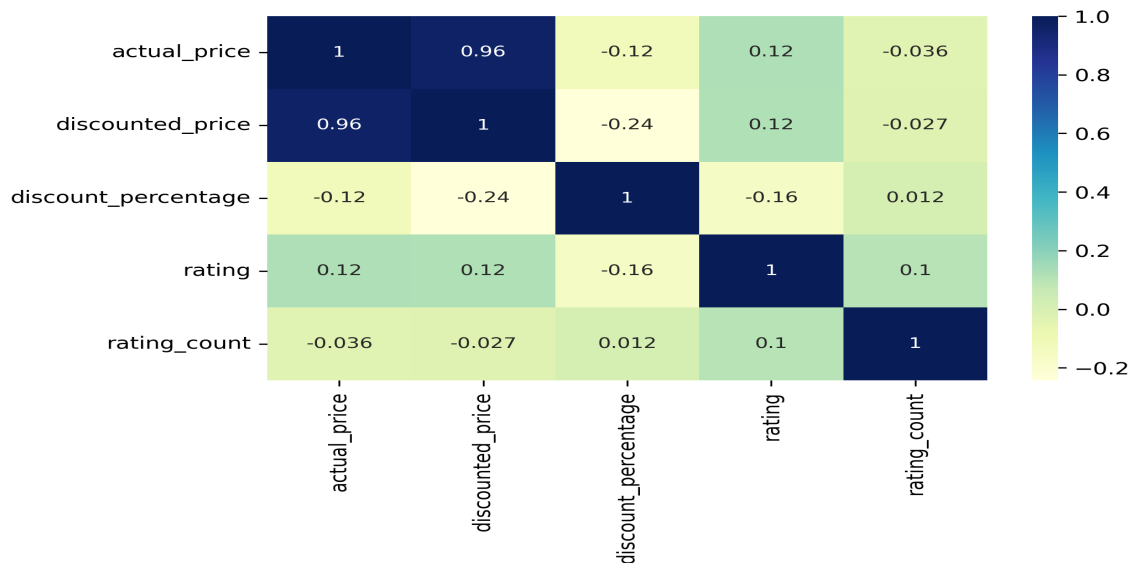
Observation: Actual_price and discounted_price show a strong positive correlation (as expected). Discount_percentage has a moderate negative correlation with discounted_price and actual_price. Rating and rating_count have very low correlation with price or discount, showing customer satisfaction is not

strongly price-driven.

Key Findings and Insights

• Most Amazon products in the dataset maintain high ratings between 3.5-4.5 stars.

• Discount percentage doesn't correlate strongly with customer satisfaction or rating.

• Some heavily discounted products (over 70%) still have low ratings, suggesting discounts are used to move underperforming inventory.

• Technology products, particularly computer and mobile accessories, dominate the dataset.

• Price and rating correlation is minimal, indicating customers value factors beyond just price.

• The dataset includes significant price outliers, particularly in the higher price ranges.