1. Background information

The analysis is conducted on a large-scale Amazon Reviews dataset, collected in 2023 by McAuley Lab. The metadata dataset includes rich features as:

Categories, average rating, rating number, features, description, price, store name, details, product ASIN

The product categories include Beauty, Fashion, Appliances, Tool & Home improvement, Amazon Home, Industrial & Scientific, Automotivem Health & Personal Care, and Handmade products.

Total size of sampled dataset: 212327.

1. Analysis Result
2. Comment count by categories

In descending order: Handmade (nearly 100000 comments), Fashion (nearly 60000 comments), Tool & Home improvement, Beauty, Appliances, Amazon Home, Industrial & Scientific, Automotive, Premium Beauty, Health & Personal care.

1. Categories such as Handmade, Fashion, and Tool & Home Improvement seem to attract a significant amount of engagement based on the high number of comments. This could indicate that these categories are of particular interest to the audience.
2. Market Trends: The order of categories by comment count can provide insights into current market trends and consumer preferences. For instance, if Handmade products are receiving the most comments, it might indicate a growing interest in artisanal or custom-made items.
3. Opportunities for Improvement: Categories with lower comment counts, such as Premium Beauty or Health & Personal Care, may present opportunities for increased engagement or targeted marketing efforts to boost audience interaction.
4. Average rating and price
5. Highest price category: Camera&Photo has the highest average price (over 400 USD), but also has a high rating (4.66).
6. Other high priced categories include:

Computers (approximately 300USD, rating 4.52).  
Appliances (approximately 186USD, rating 4.23).

1. Low price category:  
   The average prices of Digital Music and Video Games are the lowest (approximately 9.37 and 7.25USD, respectively). Among them, Digital Music has the lowest rating (3.0), while Video Games has the highest rating (5.0, but only includes one product).
2. Rating Trend:  
   The average rating for most categories is in the range of 4.0-4.5, which is generally high.
   1. Categories with a rating below 4.0:  
      Toys & Games (3.96)  
      Digital Music (3.0, lowest rating).
   2. The relationship between price and rating:  
      High priced categories often come with high ratings, such as Camera&Photo and Cell Phones&Accessories. The performance of low-priced classification ratings varies greatly, for example, Digital Music has low ratings for low prices, while Video Games has high ratings for low prices.
3. Price distribution under 100USD
4. The price distribution of most categories is concentrated in the range of 0 to 50USD, with a few categories having a higher price distribution (close to 100USD).
5. Books and Digital Music have the most concentrated and low prices, almost all below 20USD.
6. The price distribution of Appliances, Camera&Photo, and Musical Instruments is relatively wide, with some products approaching 100USD.

Business owner should price a product within the market average range. This can help a business position itself competitively. If a product is priced significantly higher than the market average, it may deter price-sensitive consumers and lead to a perception of being overpriced. On the other hand, pricing below the market average might raise concerns about quality or value. This can also facilitate easier market entry and adoption, especially for new products or brands trying to establish themselves in the market.

1. Rating distribution
2. Overall rating distribution: The ratings of most classified products are concentrated in the range of 4.0-5.0, with relatively high ratings. Outliers are mainly concentrated in the low scoring areas (1.0-3.0), indicating the presence of a small number of low scoring products in these classifications.
3. High rated categories: The ratings for Amazon Devices, Movies&TV, and Video Games are very concentrated and close to 5.0, indicating a high level of user satisfaction with the products.
4. Categories with lower ratings: Except fot the one product in Digital Music receiving the lowest rating (3.0), Toys&Games also has a low rating, with a median close to 4.0 and a high number of low rating outliers.
5. The wide distribution of ratings for Appliances and Automotive indicates significant differences in user satisfaction for products within these categories.
6. Top 10 stores ranked by product number

In descending order: Generic (self-operated small shops without labeled brand), Whirlpool (approximately 5000 products), GE (approximately 3000 products), YELLOBIRD ART & DESIGN (approximately 3000 products), SAMSUNG, UpStart Components, FRIGIDAIRE, Upstar Battery, LG, Lone Start Art.

1. Stores ranked by rating score
   1. The store with the highest rating: CERAMICHE D'ARTE PARRINI is the highest rated store with an average rating of 4.89 and a product count of 81.
   2. Other high rated stores include:  
      Whitehouse Art (4.89, 61 items).  
      PetArt.us (4.88, 99 products).  
      TiaZoey (4.86, 63 items).  
      The ratings of these stores are close to perfect, indicating a very high level of user satisfaction with their products.
   3. The store with the lowest rating (filtered): The store with the lowest rating in the table is QNAVIC, with an average rating of only 3.15 and a product quantity of 60.
   4. Other stores with relatively low ratings include:  
      OdrillionGems (3.52, 74 items).  
      Forum Novelties (3.59, 61 items).  
      These stores have relatively low ratings, which may indicate issues with product quality or inadequate user experience.
   5. The relationship between rating and product quantity:  
      The number of products in high rated stores is mostly in the range of 50-124. Stores with a larger quantity of products may not necessarily have higher ratings, but stores with lower ratings typically have fewer products.

The result suggests that stores with a smaller product selection might struggle to meet diverse customer needs or preferences, leading to lower ratings. Providing a limited range of products could result in customers seeking alternatives elsewhere, impacting their overall satisfaction with the store. Nonetheless, businesses should pay attention to customer reviews, address any issues raised, and continuously improve their offerings to maintain high ratings, regardless of their product quantity.

1. Relationship between product information integrity and rating
   1. Products with complete information have higher ratings:  
      Products with feature highlights, product description and images have the highest average rating (4.36).
   2. Products with images but without feature highlights, product description, and videos have the lowest rating (4.18)
   3. Having product descriptions and images are important criteria for high ratings.
   4. The impact of the video is relatively small: Adding videos has limited improvement on ratings, and the overall rating has not changed much.
2. The relationship between the number of ratings and ratings
   1. Products with fewer ratings have greater fluctuations in ratings:  
      The distribution of ratings within the range of<10 and 10-50 is relatively wide, with product ratings ranging from 1.0 to 5.0, indicating that products with fewer ratings have unstable ratings and are easily affected by extreme ratings.
   2. Products with more ratings have more concentrated ratings:  
      The distribution of ratings in the 100-500, 500-1000, and 1000+ ranges is relatively narrow, with product ratings concentrated in the 4.0-5.0 range, indicating that products with more ratings tend to have more stable ratings.
   3. Distribution of low rated products:  
      In the range with fewer ratings (especially<10 and 10-50), there are more low rated products (outliers with ratings below 3.0).  
      As the number of ratings increases, the number of low rated products gradually decreases, indicating that products with more ratings are more representative of overall user satisfaction.

Consumers often perceive products with a larger number of ratings as more reliable indicators of quality and user satisfaction. A higher volume of ratings can instill greater trust in potential buyers, as they are based on a more extensive and diverse set of experiences and opinions.

Suggestions:

1. Actively encouraging and accumulating ratings can help to establish credibility and trust among consumers. Products with a substantial number of ratings are more likely to attract potential customers and inspire confidence in their purchase decisions.
2. Understanding the dynamics between the number of ratings and rating stability guide businesses in managing customer feedback effectively. Monitoring and responding to feedback for products with fewer ratings can help mitigate fluctuations and improve overall rating consistency.
3. Encouraging customers to leave reviews and ratings (e.g. through offering coupons if the customer leaves comments) can help increase the sample size and enhance the reliability of product ratings, providing a more accurate reflection of user satisfaction levels.