1. Background information:

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

ratings, review titles, review content, helpfulness votes, whether the purchase is verified, review timestamp.

The product categories include Beauty, Fashion, Appliances, and Handmade products. Total size of sampled dataset: 50000.

1. Analysis Result
2. Rating Result
3. In the rating distribution, 5-star ratings account for the absolute majority, with a much higher number of reviews than other ratings. 5-star ratings account for 64.7%, which is close to 2/3 of the total reviews, indicating that the overall evaluation of the product is very good.
4. 4-star ratings account for 11.3%. 3-star ratings account for 7%. 2-star ratings account for 5.2%. 1-star ratings account for 11.8%.
5. Generally, the distribution of ratings biased towards high scores.

The result shows that the overall user experience is good. But marketers should notice there is 17% of ratings are 1- and 2-star ratings, indicating that there are a small number of users who are not satisfied with the products.

1. Keyword in review
2. High frequency: “great”, ”like”, ”love”, ”good”, “fit”, “well”.

This result is consistent with the rating result, indicating that the overall evaluation of the products on Amazon is very good.

1. Relatively high frequency: “price”, ”product”, ”quality”, ”looks”.

Provide information of what are the main concern of costumers.

1. Sentiment analysis for all: categorized to “Positive”, “Negative”, and “Neutral”
   1. The majority of comments are positive, approximately 7 times the negatives.
   2. Neutral comments are slightly fewer than negative ones.

The overall satisfaction of users with the product is relatively high, with more positive feedback in the reviews. Neutral or negative reviews may reflect a minority of users' dissatisfaction or uncertain attitude towards the product.

1. Sentiment analysis for Long and Short Review respectively
   1. The average rating of long comments is slightly lower than that of short comments. This indicates that long comments may be more inclined to express detailed opinions, which may contain more criticism or dissatisfaction. Short comments may be more inclined to express satisfaction directly (such as "Great product!")
   2. Generally, positive sentiment dominates the review no matter it is long or short.This result is consistent with the rating result, indicating that the overall evaluation of the products on Amazon is very good.
2. Top ten most reviewed products

Their ASIN’s are “B01DLEL4EM”, “B00LGEKOMS”, “B000AST3AK”, “B000GAWSDG”, “B00UXG4WRB”, “B01BIDPHYQ”, “B003N1ZSYG”, “B002JAKRAM”, “B01DP1WKU”, “B01ALBMIEI”.

People may search these ASIN’s in Amazon to see the product details.

1. Daily review count from 2006 to 2024
   1. In 2006-2011: less than 5 reviews per day. This may be due to the small market size of related products or the limited user base of the platform.
   2. In 2012-2020: The daily review count has been through a rapid growth, reaching more than 40 reviews per day. This fact reflects that: the sales volume of the product has increased, user engagement has increased, and the platform has expanded in scale and attracted more users.
   3. Peak in 2020: The pandemic shifted the sales online and brought more costumers to the platform.
   4. In 2021-2024: The daily review count has gradually decreased. This may be due to the data being incomplete, resulting in missing records of the latest comments.

The increase in the number of comments may be directly related to the popularity of products, the development of platforms, and the improvement of user activity. E-commerce business owners should pay attention to such index.

1. Monthly review count (all years combined)

Peak number of comments:

1. January: May be related to the New Year
2. March: May be related to Easter
3. November and December: may be related to Halloween and Christmas  
   Consumers may purchase products more frequently during holidays, and they tend to leave reviews on products after the holidays.

Low period of comment quantity: May and June  
The decrease in the number of comments may be related to seasonal demand decline, such as a decrease in consumer purchasing and commenting behavior in non holiday months.

1. Helpful votes
   1. The relationship between comment length and useful voting:
      1. Long comments are more likely to receive highly useful votes. Although the number of long comments is relatively small, they are more likely to receive higher useful votes (over 100 votes), indicating that users may consider long comments more valuable.
      2. When the comment length exceeds 2000 characters, there is a significant increase in the number of useful votes.
      3. Outliers: One comment is excessively long (over 4000 characters) and has received extremely high useful votes (over 800 votes), possibly because it provides detailed evaluations or profound insights.
   2. The relationship between rating and average useful voting:
      1. Extreme ratings (1 star and 5 stars) are more likely to receive highly useful votes. The average useful votes for 1-star and 5-star reviews are the highest, indicating that users are more inclined to believe that these reviews have reference value. For example, 1-star reviews may reveal serious issues with the product and attract more user attention. A 5-star review may provide a detailed positive evaluation of the product and is also likely to receive more useful votes.
      2. Comments with intermediate ratings have fewer useful votes: The average useful votes for 2-star, 3-star, and 4-star reviews are relatively low, possibly because these rated reviews have less reference value for users' purchasing decisions. For example, 3-star reviews are usually neutral, with neither specific advantages nor clear disadvantages, and users may not place much importance on these reviews.
2. Average rating and purchase verification
   1. Verified purchase rating is higher: The average rating given by verified purchase users is higher than that of unverified purchase users. This may be because unverified purchase reviews may include low rating comments given by competitors due to vicious competition.
   2. The difference in ratings is relatively small: The average rating of both types of reviews is close to 4.0, indicating that users are generally satisfied with the product, regardless of whether the reviews are based on verified purchases.
3. Weighted average rating

The weighted scoring formula is: weighted\_rating = rating \* (1 + helpful\_vote)

rating: Refers to the user's original evaluation of the product (1-5 stars).

helpful\_vote: amplifies the impact of useful comments on ratings.

1. Weighted ratings tend to enhance the contribution of comments that have both high ratings and more useful votes to the overall rating.
2. Weighted ratings can better reflect users' trust in comments:

Unlike simple average ratings, weighted ratings comprehensively consider the usefulness of ratings and reviews, making products with higher review quality rank higher. This approach is more suitable for the sorting logic of recommendation systems or product display pages.

1. Top 10 products by weighted average rating

Their ASIN’s are “B09H2VFWX4”, “B08R1VNVJY”, “B09CKQVH7N”, “B00SZASXD8”, “B087Q27ZKD”, “B08GZKN5RF”, “B0711NNMLP”, “B01CZT54OU”, “B01H4MA31C”, “B0009MD9Q6”.

People may search these ASIN’s in Amazon to see the product details.

1. Mismatched reviews
   1. Mismatched comments refer to comments with high ratings (4 stars or above) but negative sentiment analysis polarity (sentiment polarity<0). This type of comment is contradictory and may reflect a discrepancy between the content of the comment and the rating.
   2. The highest number of mismatched comments is 5-star rating. Possible reasons: the user expressed overall satisfaction with the product in the rating, but specific dissatisfaction was mentioned in the comments (such as logistics or after-sales issues).