

The problem with online reviews and ratings

User ratings and product quality

Consumers use online user ratings because they assume these provide a good indication of product or service quality. For example, you would expect a laptop with an average rating of four out of five stars to be objectively better than a laptop with an average rating of three out of five stars, 100% of the time.

In order to test this assumption, one [researcher team](#) put together an impressive dataset comprising of 344,157 Amazon.com ratings for 1,272 products, in 120 product categories. For each product, they obtained objective quality scores from the website [Consumer Reports](#). They also collected data on prices, brand image measures, and two independent sources of resale values in the market for second hand or used goods.

The researchers found that average user ratings correlated poorly with the scores from Consumer Reports. For example, when the difference in average user rating between pairs of products was larger than one star, the item with the higher user rating was rated more favourably by Consumer Reports only about two-thirds of the time. In other words, if you were comparing a laptop with an average rating of four out of five stars, with another laptop with an average rating of three out of five stars, the first laptop would only be objectively better 65% (not 100%) of the time. This is a far cry from a sure difference in quality. Moreover, the average user ratings did not predict resale value in the used-product marketplace.

The reasons online ratings don't reflect the real thing

There are several reasons why average user ratings may not predict objective quality measures. User reviews may include a broader range of criteria than those Consumer Reports does, such as subjective aspects of the use experience (like aesthetics, popularity, emotional benefits).

Many reviews are also based on small samples. As any statistics teacher will tell you, all things being equal, the average user rating should be more informative as sample size increases relative to variability. Indeed, in the [online rating study](#), the correlation between average user rating and Consumer Reports scores was higher when the sample size was large. Unfortunately, average user ratings are often based on [small samples and high variability](#).

Online reviews are based on a biased subset of those who actually purchased the product. In general, reviews are left by those that “brag” or “moan” about their product experience, often resulting in a [two mode distribution of ratings](#).

This is where the average does not give a good indication of the true population average. For example, in one comprehensive dataset for a large private label retailer, the percentage of buyers who left a review was just [1.5%](#). This means that 98.5% of the people eligible to leave a review chose not to do so.

Many groups also now actively seek to manipulate average ratings. This can be done in the form of fake reviews. For [example](#), businesses (or their agents) may post fictitious favourable reviews for their own products and/or post fictitious negative reviews for the products of their competitors. [According to one study](#), roughly 16% of restaurant reviews on the website Yelp were suspicious or fake.

Websites like Yelp.com and Amazon.com try to mitigate such ingenuity. For example, one of the [Ivanka Trump collection's shoes](#) has an average rating of four and a half out of five stars despite hundreds of (presumably fake) one-star reviews.

What you can actually tell from online reviews

There is a way to use the information from reviews and ratings despite all of these potential pitfalls. First, look for products with a high average user rating, many reviews, and not a lot of variance in the rating scores. Beware placing too much faith in average ratings that are based on few reviews and with high variance in the ratings. You can also consider online reviews in light of additional sources that provide objective product evaluations, from technical experts. Sources of this kind of information include Consumer Reports, Choice, Consumers Union, Which? and CNET.

Where possible, you can consider employing technology designed to help you navigate the bias in online reviews. Examples include [Fakespot](#) and [ReviewMeta](#). For example, ReviewMeta scans all reviews from a product's online listing page, and then provides an adjusted average rating. This adjusted rating accounts for all sorts of suspicious activities such as a high proportion of reviews from users with unverified purchases.

So, the next time you're evaluating products online, feel free to start with the average user rating, but be wary of making your final judgement based only on this cue.