Project: Is talking to Alexa enough? Or do we need more?

Data analysis

Previous data analysis has shown that reviewers, as shown in Figure 1, when they are not as satisfied with the device given the lower star rating, would write more. Interestingly, the two screen-less devices, with their relatively higher mean star ratings, have shorter reviews than the three devices with screen.

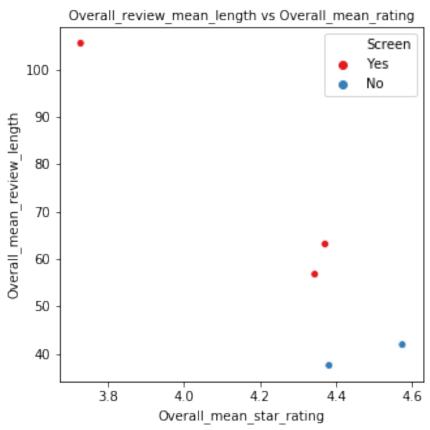


Figure 1. Mean review length vs. mean star rating of selected devices.

Further examination of the data shows that the reviews for all the five selected devices are highly skewed with mostly 5-star reviews. The number of reviews and their relative proportion for each star-rating group are plotted in Figures 2 and 3, respectively. Figure 2 shows that 5-star reviews made up for about 70% of the reviews for all devices but Echo Show 1^{st} Generation, from which $\sim 50\%$ are 5-star reviews.

Review count vs Star rating

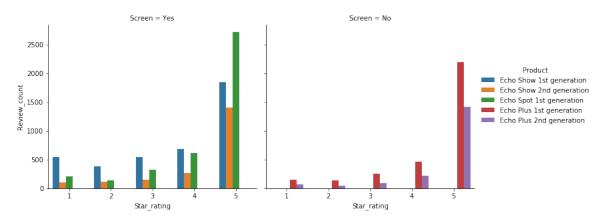


Figure 2. Number of reviews from each star rating group.

Review count percentage vs Star rating

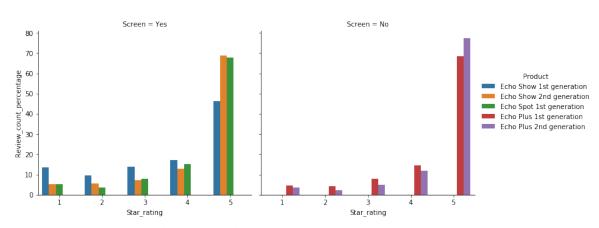


Figure 3. Relative proportion of review count from each star rating group.

The overall high mean star ratings can probably be explained by the skewed reviews. The above analysis also suggests that reviewer's satisfaction might correlate with how much they write. Understanding the relationship between the reviewer's sentiment and the review content would be of interest. However, being able to extract information regardless of the review's length would be important. The next steps will include examining the reviews further by word cloud analysis, which would help identifying features that are liked and disliked by the reviewers, and by performing a sentiment analysis, which would allow us to measure and analyze the sentiments expressed quantitatively.

Word cloud analysis

Monogram, bigram and trigram word cloud analyses were performed on both review titles and review contents of each star-rating group from all selected Echo devices. Aiming to identify the most liked and disliked features, the following analysis will focus on the 5-star and 1-star

reviews. Word clouds were generated based on monogram, bigram and trigram frequencies, which are shown respectively in Figures 4 and 5 from 5-star and 1-star reviews. In each figure, the 20 most frequent words/phrase are shown and the word size based on the computed frequency.

The word cloud analyses from the 5-star reviews reveal some of the most popular features (Figure 4). Reviewers, who gave 5-star reviews to Echo Show 1st and 2nd Gen, did mention the screen. When Echo Show was released, Amazon promoted the device's capability of making video calls via the Drop-in feature. Amazon also advertised Echo Show as a companion device in the kitchen. Both of these were present in the word clouds. Interestingly, reviewers of Echo Show 2nd Generation have mentioned "sound" a lot, which is similar to the reviews for the screen-less speakers. Reviews of Echo Spot, which is advertised as a smart alarm clock, show buyers do enjoy using the device as alarm clock. Unsurprisingly, Amazon's smart voice assistant, Alexa, appears to be a prominent feature for all devices. Non-Amazon services that are compatible with Echo, such as Spotify, however, are not present, suggesting that these third-party services are overshadowed by the Amazon's services.



Figure 4. Word clouds from 5-star review contents.





Figure 5. Word clouds from 1-star review contents.

Word clouds from 1-star reviews, as shown in Figure 5, were generated to examine reviews at the other end of the spectrum. Interestingly, some features appear in both unigram word clouds from 5-star and 1-star reviews, such as Alexa, Amazon, music, screen (for Echo Shows and Echo Spot) and speaker (for Echo Pluses). This may suggest that 1-star reviewers also like these features or the 1-star reviewers actually dislike these features liked by many 5-star reviewers. The bigram and trigram word clouds do provide a bit more insights into reviews. For Echo Show devices, the reviews mentioned about YouTube videos and flash player, both of which is specific to screen-enabled devices. This likely relates to the rollout of Echo Show's own YouTube app at its release. The Echo Show 1st generation was first released with a YouTube app, however, the app was then blocked by Google for some time. The "Alexa app" and "phone" also appear here for multiple devices, which could reflect on interactions with the Echo devices via the Alexa app in the phone. This analysis also helps identifying couple device-specific issues. Some Echo Plus 2nd generation devices might have a problem with WiFi signal (from bigram word cloud). Some 1-star reviewers also reported screen issues, such as screen flickering and lines on the screen, for Echo Show 1st generation. These issues, while do not appear in the word clouds of the review contents, was present the word clouds generated from the review titles as shown in Figure 6.

The next step is to carry out sentiment analysis to quantify the sentiments expressed in these reviews. The goal would be to connect the results from sentiment analysis and the features/issues identified from the word cloud analysis, so we could profile and compare user experience of these Echo devices.



Echo Show 1st Gen







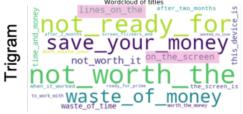


Figure 6. Word clouds from 1-star review titles of Echo Show 1st Generation. Words and phrases related to screen flickering is highlighted in pink.