Seboca AI

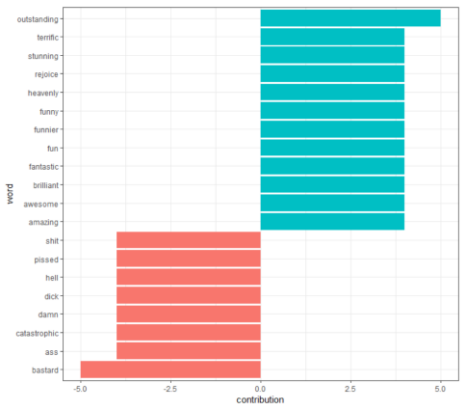
YELP REVIEWS

Sentiment Analysis

Overview

Instead of reading millions of reviews, we can already get a good feel for the content by analyzing it with techniques such as wordclouds, sentiment scores and bigram counts. The yelp data set is a collection of much more data than we have actually considered in this example but even with this limited amount of data we can generate business relevant insights. Some of the key findings are that the sentiment scores are not in line with the star ratings users assign to rate restaurants. The wordcloud we created for the reviews of the poweruser told us instantly that he or she is writing about food quality whereas the best average star rated reviews are talking more about experience and customer service. Beyond the word cloud, bigrams are extremely helpful because they put neutral words such as "price" in relation which in this case is needed to understand whether is has a positive or negative sentiment. The word tree graph shows that price is used with words like "reasonable" and "fair" and is thereby positive.

Sentiment Contribution



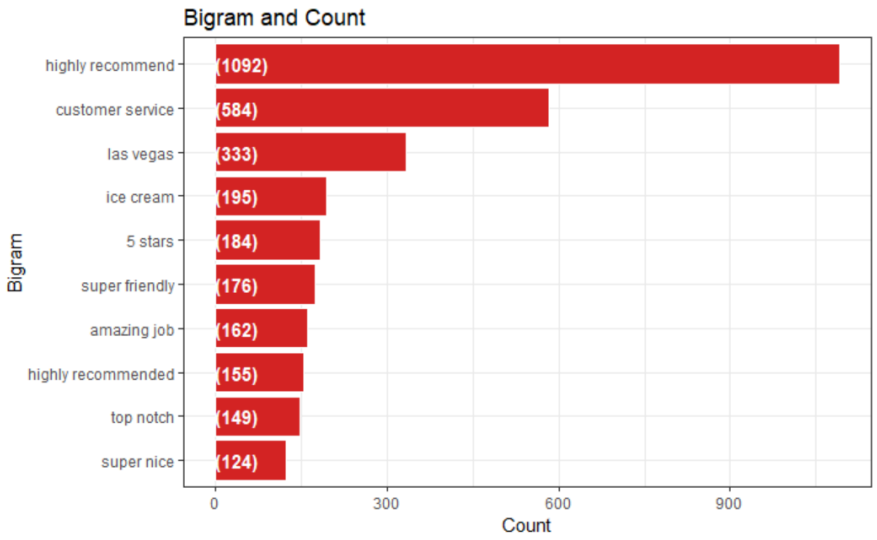
Sentiment Score

The idea behind sentiment analysis lies in counting the frequency of words used in the analysed text and evaluating their sentiment. In this analysis we used the "AFINN" lexicon which scores close to 2500 words with a value between -5 (negative) and 5 (positive).

Wordcloud Best Rated

Bigram Trees

Bigram Counts



Get in touch hello@seboca.com