Replication of ACSI index values for US airlines using sentiment analysis of Twitter data

# Concept

ACSI index ranks customer satisfaction performance based on Customer Expectations, Perceived Quality and Perceived Value. These translate into Customer Complaints and Customer Satisfaction[[1]](#footnote-1).

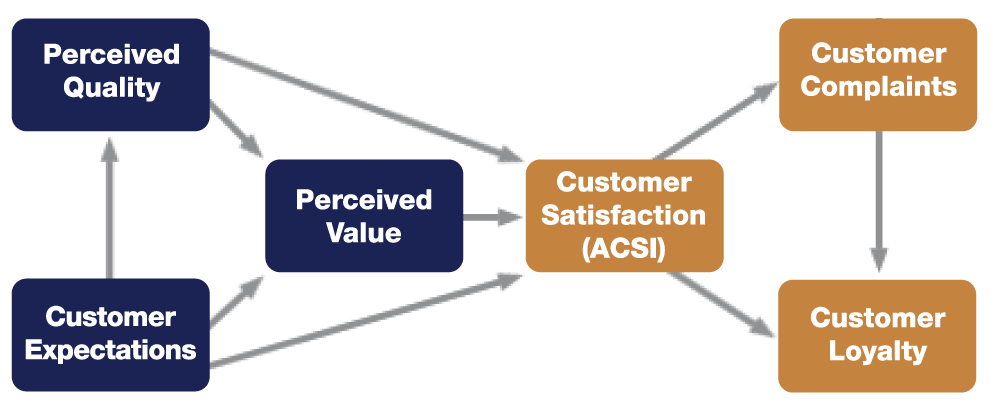


Figure - ACSI rank method

This project seeks out to replicate the index results, not using traditional customer satisfaction assessment methods such as focus groups, interviews and questionnaires, which are time consuming, expensive and have limited scalability, but using sentiment analysis on Twitter data.

R offers 3 lexicons using which one can perform sentiment analysis. “Bing” and “Afinn” provide numerical sentiment values to data (Afinn being more granular than Bing). These lexicons can be used to categorise tweets as “good” or “bad”, to a certain degree of polarity.

The “NRC” lexicon is different, categorising words into emotions rather than sentiments.

# Method 1 – Using the Bing lexicon

The Bing lexicon categorises words as having positive or negative sentiment, to which the scores 1 and -1 could be assigned.

We then analysed the tweets in two ways:

1. Total of Positive and negative for each airline: Simply the points for each airline were summed up and the airlines were ranked up. Surprisingly this gave a ranking similar to the ACSI ranking, by an average rank offset of 1.5. However, as 10 positive words in one tweet does not have the same effect as 10 tweets with one positive word each, we decided to perform the analysis again, on a polarity by tweet basis
2. Creating a polarity metric and assigning a polarity for each tweet. Then a set of metrics representing polarities of different thresholds (1, 2 and 3) were calculated for each airline, and the airlines were ranked for each polarity threshold metric. This produced an average rank offset of 0.88 and a standard deviation of 0.92. Upon further thought, we thought that a simpler way of assessing holistic polarity is to average polarities of all tweets for each airline, thus ranking airlines by average tweet polarity. This produced the same mean and standard deviation results, showing that the sentiment ranking method plays minimal difference

Other techniques employed involve removing retweets to keep original information, as well as removing all tweets sourcing from airline accounts, as they do not reflect consumer sentiment and skew the results to being more positive.

Finally, we went into each lexicon to identify words that were contributing highly, but instead were out of context. For this reason, we removed the word “trump” from the analysis, as it was a top 20 contributor to the FlyFrontier sentiment, but is obviously out of context.

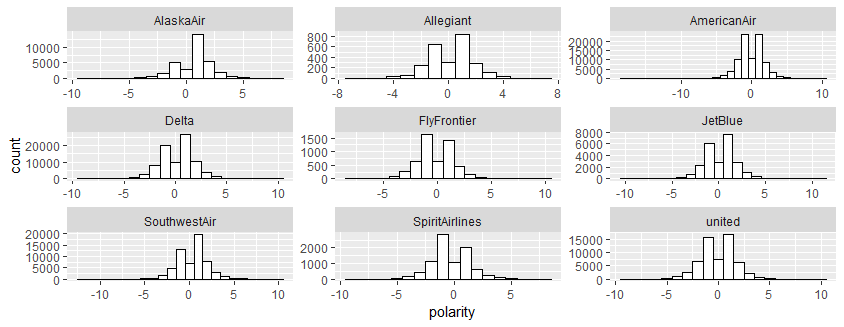


Figure - Tweet polarity histogram, by airline

In this chart we can start seeing that some Airlines like AlaskaAir and Southwest have clearly higher polarities, while Spirit is clearly lagging.

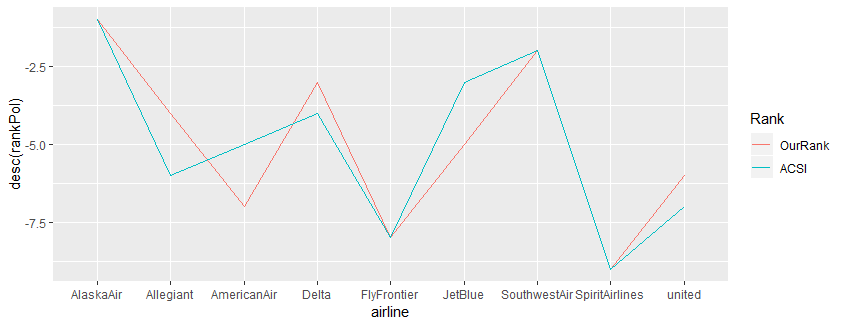


Figure - ACSI vs sentiment analysis rank

In this graph we can see that our rank is closely associated with the ACSI rank. For the first two and last two ranks, where the average polarities had a clear difference with the rest, the results match.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| airline | avgPol | rankPol | acsi\_rank | rank\_diff |
| AlaskaAir | 0.701131 | 1 | 1 | 0 |
| SouthwestAir | 0.354022 | 2 | 2 | 0 |
| Delta | 0.227481 | 3 | 4 | 1 |
| Allegiant | 0.178805 | 4 | 6 | 2 |
| JetBlue | 0.150674 | 5 | 3 | 2 |
| united | -0.01806 | 6 | 7 | 1 |
| AmericanAir | -0.08446 | 7 | 5 | 2 |
| FlyFrontier | -0.18428 | 8 | 8 | 0 |
| SpiritAirlines | -0.33204 | 9 | 9 | 0 |

Figure - Sentiment analysis results using Bing lexicon

# Method 2 – Using the Afinn lexicon

The Afinn lexicon does not give a positive and negative sentiment, but instead a sentiment score, from -5 to +5 in integers. Again, we calculated the polarity of each tweet by summing the sentiment scores of its words, and then calculating the average polarity per tweet of each airline. Again, we removed all retweets, all tweets by airline accounts, and all words that were out of context, such as “no” and “united”.

The mean difference in rank using the afinn lexicon was better by 25% compared to Bing (0.66 isntead of 0.88 mean rank difference), while the standard deviation was also better (0.70 compared to 0.92). Therefore, we can see that granularity in scores provided slightly better results.

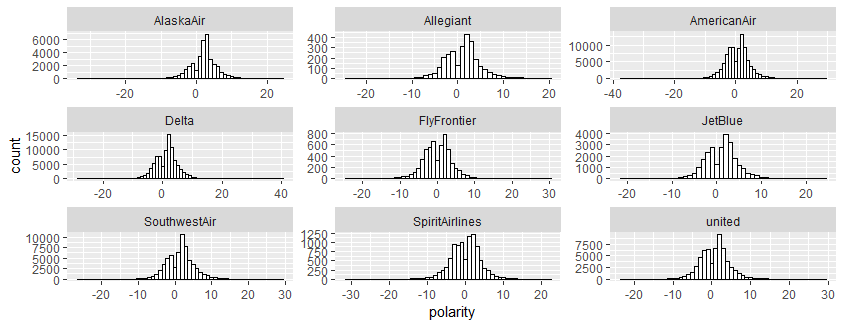


Figure - Tweet polarity histogram using Afinn lexicon

As we can see from the histograms, the results are similar to Bing lexicon, in the sense that airlines such as Alaska and Southwest clearly show more positive results, while Spirit, even though it shows it might leaning to a positive polarity average, its polarity is clearly lower than the former two airlines.

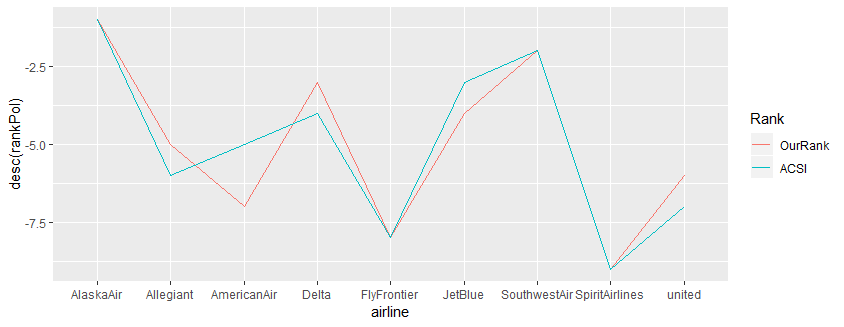


Figure - ACSI vs sentiment analysis rank using Afinn lexicon

This rank comparison for the Afinn lexicon results is similar to the same graph obtained for the Bing lexicon, with some minor differences, where some airlines’ sentiment analysis rank moved closer to their ACSI rank. Again, the results for the first and last two ranks match those of ACSI.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| airline | avgPol | rankPol | acsi\_rank | rank\_diff |
| AlaskaAir | 2.209336 | 1 | 1 | 0 |
| SouthwestAir | 1.512982 | 2 | 2 | 0 |
| Delta | 1.231541 | 3 | 4 | 1 |
| JetBlue | 1.136214 | 4 | 3 | 1 |
| Allegiant | 1.031094 | 5 | 6 | 1 |
| united | 0.686554 | 6 | 7 | 1 |
| AmericanAir | 0.522334 | 7 | 5 | 2 |
| FlyFrontier | 0.020775 | 8 | 8 | 0 |
| SpiritAirlines | -0.07177 | 9 | 9 | 0 |

Figure - Sentiment analysis results using Afinn lexicon

As we can see from the results, the only airline with a negative polarity is Spirit, compared to 4 airlines which had negative average polarities using Bing. These shows us that Afinn skews sentiment to more positive values compared to Bing.

1. <https://www.theacsi.org/about-acsi/the-science-of-customer-satisfaction> [↑](#footnote-ref-1)