

# Sentiment Analysis of 'Aetna', 'Obamacare', and 'Single Payer'

Zachary Schultz DS 710

## Questions:

Health Insurance. Simply reading those two words can cause visceral reactions in some. Public sentiment towards private insurers seems tumultuous at the moment. Attitudes towards the health insurance marketplace are valuable to insurers, politicians, and marketers. This industry, more than most, is under increasing public scrutiny and government intervention. Significant percentages of Americans would vote to introduce a public option and do away with the entire industry. If the marketplace doesn't improve, voters may do exactly that. The boiling climate in this sector pushed me to perform a sentiment analysis on tweets containing 'Aetna', 'Obamacare', and 'single payer'. I investigated the following questions:

1. Are tweets with the word 'Aetna' more negative than tweets containing 'Obamacare' or 'Single Payer'?
2. Are tweets with 'Aetna' more likely to contain profanity than tweets with 'Obamacare' or 'Single Payer'?

## Analysis:

Tweets were gathered from the twitter REST API. This tool searched and gathered thousands of tweets containing the keywords "Aetna", "Obamacare", and "single payer". It was performed during the month of April, 2017. There were additional searches on other private insurers and other disliked topics for comparison's sake. Most tweets were gathered between 4/9 and 4/22. A tool called Textblob was used to determine the polarity and subjectivity of the tweets. Polarity ranks tweets on a scale of -1 (for negative sentiment tweets) to positive 1. The more negative a tweet is, the closer to -1 the polarity will be. Subjectivity of tweets is on the scale of 0 to 1, with 0 being objective and 1 equating to subjective. I also gathered the percentage of keyword tweets that contained profanity. For the results, I compared the polarity and profanity in the questions above.

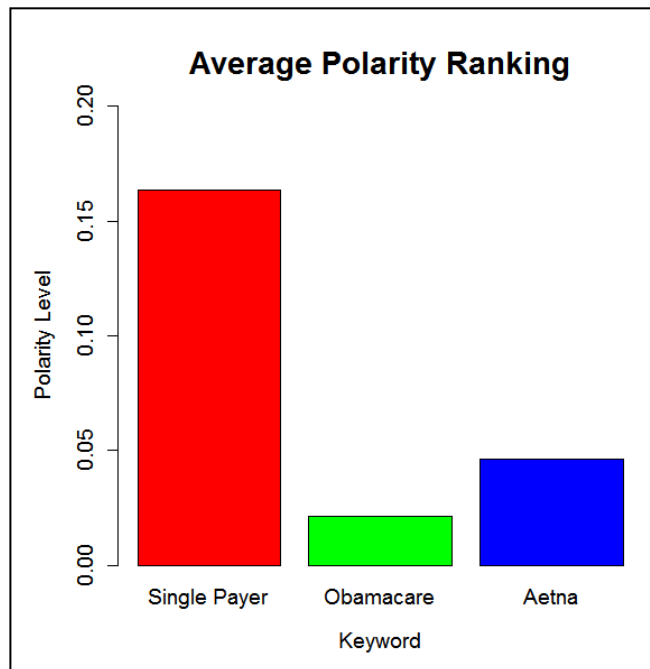
## Results:

### 1. Question 1 results:

The first comparison of 'Aetna' tweets and 'obamacare' tweets showed significant difference in polarity. (For the sake of brevity, when using a term like 'aetna tweets', I am referring to any tweet containing the word 'aetna') Aetna had over double the polarity ranking. This suggests that, of the collected tweets, Aetna tweets appear to have a more positive sentiment than Obamacare tweets. However, most notable (as shown by the graph on the left), is the average polarity of tweets containing 'single payer'. These tweets were significantly greater than the mean polarity of Aetna tweets. The sample mean polarity of tweets containing 'single payer' was almost 4 times those containing 'aetna'. This suggests single payer may have be viewed in a more positive light than the current marketplace and further research.

### 2. Question 2 results:

Each set of tweets was searched for profanity. 8.77% of 'single payer' tweets, 3.27% of Obamacare tweets, and 3.32% of Aetna Tweets contained profanity. At first, the 8.77% number was alarming, but the large majority (181/237) of the offending single payer tweets were from a heavily retweeted tweet containing the word 'hate'. Had that tweet been removed from the dataset, the profanity rate for single payer drops to 2.22%.



### Other Interesting Findings:

For the sake of comparison, I analyzed tweets with other private insurers as well as tweets containing the word 'homework' and 'trump'. Of the collected tweets, 'Single Payer' has notably higher polarity and subjectivity scores than private insurers. 'Cigna' has the best polarity of the private insurers studied. 'Homework' tweets had a humorously high profanity rate.

<b>Tweets Containing Keyword</b>	<b>Polarity</b>	<b>Subjectivity</b>	<b>N= # of tweets</b>	<b>Tweets w/cuss words</b>	<b>Profanity Frequency</b>
'Single Payer'	0.163	0.421	2700	237	8.77%
'Obamacare'	0.021	0.253	2686	88	3.27%
'Aetna'	0.046	0.263	1657	55	3.32%
'Cigna'	0.120	0.263	1809	41	2.26%
'United Health'	-0.084	0.369	375	4	1.06%
'Anthem Insurance'	0.0371	0.172	133	2	1.50%
'Trump'	0.003	0.315	2654	114	4.30%
'Homework'	0.073	0.291	2686	286	10.7%

### Conclusions:

Hypothesis tests were performed at the  $\alpha=.05$  level on question 1. The results suggest that 'Aetna tweets' are significantly more negative than 'Single Payer tweets'. However, another hypothesis test suggested that 'Aetna tweets' have a higher average polarity than 'Obamacare tweets'.

A two proportion test was performed at the  $\alpha=.05$  level and determined there is not a significant difference in the proportion of profane tweets for 'Aetna' and 'Obamacare' groups. A test was also performed on the profanity frequency of 'Single Payer tweets' and 'Aetna Tweets'. The hypothesis test suggested that profanity frequency is higher for 'Single Payer tweets' than 'aetna' tweets.

I found the most interesting piece in the research to be how much greater the polarity of 'Single Payer tweets' was to any of the private insurer group tweets or 'Obamacare tweets'. This would be something worthy of further research. For the parties interested in these sentiments, further data collection would be instrumental as legislation is being voted on in coming weeks. I'm also curious as to the polarity of tweets that intersect the different groups. For example, what would the polarity of tweets that contain both 'aetna' and 'obamacare' be?