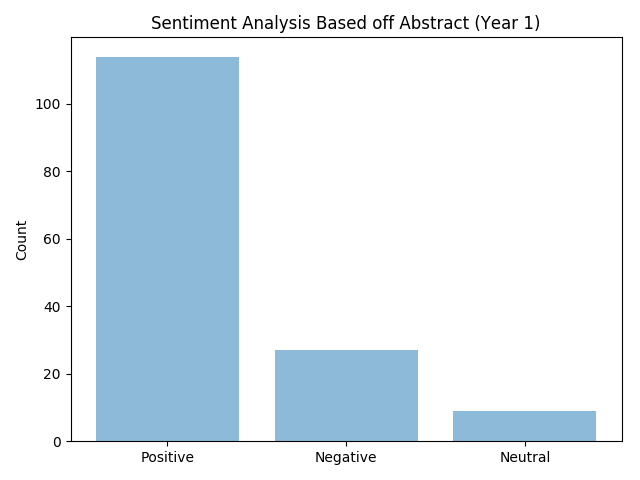
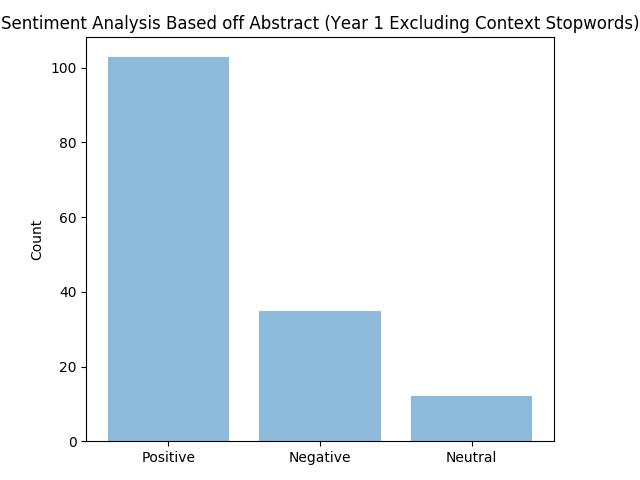
**Analysis**

Using the AFinn library, each word was analyzed and assigned a score before summing the scores for each abstract. Depending on the resulting sum, the abstract can be classified as positive, neutral or negative sentiment. Based off the algorithm, the following results were produced. Out of the 150 abstracts in the CSV file, there were 114 positive abstracts, 27 negative abstracts and 9 neutral abstracts which has been plotted below to better show the ratios of the 3 categories.



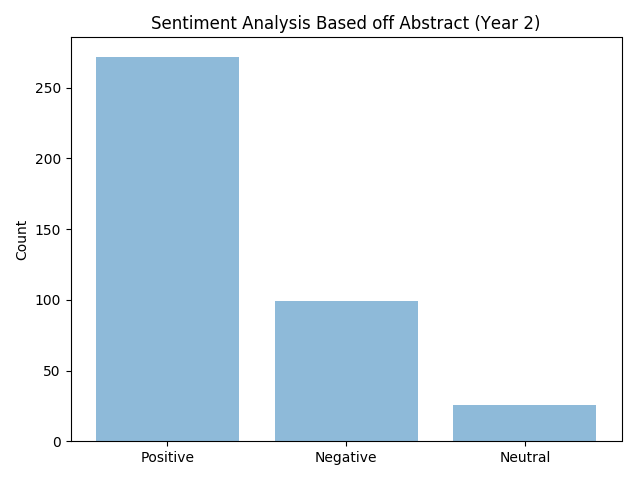
*Figure 1: Sentiment analysis graph performed on Year 1 data that includes context stop word bias*

This algorithm was then further refined to remove context stop words that were previously in our data that may have resulted in bias. The figure below shows the change in results. Though the number of positive abstracts are still very high compared to the negative and neutral it is not as prominent with a number of 103 positive abstracts, 35 negative abstracts and 12 neutral abstracts. This means that a number of abstracts that were previously identified as positive were actually falsely classified as positive and should have been either negative or neutral.



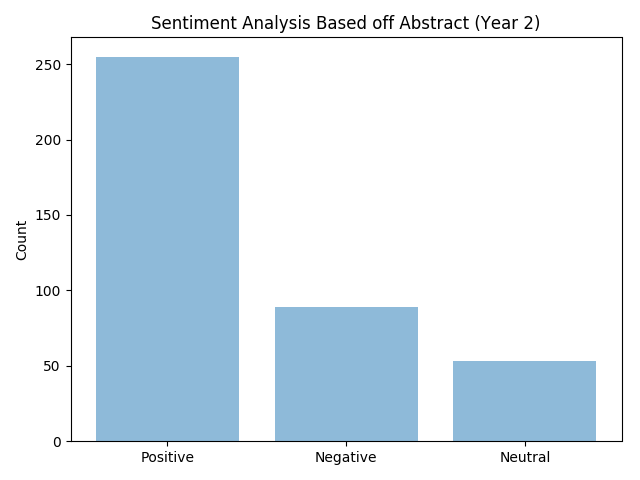
*Figure 2: Sentiment analysis graph performed on Year 2 data that includes context stop word bias*

Applying the same algorithm to the Year 2 data gave the following results. Out of the 398 abstracts in the CSV file, there were 272 positive abstracts, 99 negative abstracts and 26 neutral abstracts which has also been plotted below to better show the ratios of the 3 categories.



*Figure 3: Sentiment analysis graph performed on Year 1 data that excludes context stop word bias*

Similarly with the first year, context stop words were removed from our analysis to result in less bias. The bar graph below shows the end result while using the refined algorithm. While the amount of positive abstracts still dominated, the number of abstracts classified as neutral greatly increased as many of the articles initially classified as negative or positive turned to neutral instead. The final results ended as 255 positive abstracts, 89 negative abstracts and 53 neutral abstracts.



*Figure 4: Sentiment analysis graph performed on Year 1 data that excludes context stop word bias*

Looking at the four graphs gives some interesting findings. The number of positive abstracts clearly overpowers the number of negative and neutral abstracts in both years. By refining the algorithm to exclude context stopwords in the analysis, the amount of positive abstracts was found to decrease, as the abstracts were falsely classified initially. The number of abstracts that were classified as neutral was also the lowest in both years and that logically makes sense. In a research paper, it is very common for a writer to have a stance on whatever their research topic is on so it makes sense for the number of positive and negative abstracts to be high and not be neutral. Also a reason that many abstracts were classified as positive could be because researchers generally aim to try to prove their hypothesis and say that they were successful in their findings.