**Determining the Relationship between NFT Performance Measures and Collector Sentiment**

**By**

Aleena Naeem, Brian Rochford, Raj Sidhu, Saif Sarwar

Submitted to Prof. [Yilu Zhou](mailto:yzhou62@fordham.edu)

**Executive Summary**

Non-fungible tokens or NFTs have gained immense popularity in the last year. The topic is extremely polarizing and unique which has caused it to gain much traction. The reason why this topic is growing in popularity is because of the fact that it is a brand-new technology that aims to give certain industries such as the art industry a new platform to own, produce and distribute its products. All of this is done on the blockchain which has many benefits of its own, such as anonymity, wealth transfer/generation, and security, to name a few. NFTs must be purchased currently using Ethereum (Ether) which is the second largest cryptocurrency at the moment in terms of price and market cap and once purchased the NFTs are stored on the Ethereum blockchain. This is important to know as the price of Ethereum also affects the volume of NFT sales.

We aim to look at Opensea, the largest NFT marketplace used to buy and sell NFTs, as well as twitter, a large social networking service that has helped to increase NFT popularity, to determine if customer sentiment affects NFT prices. Our goal is to find insights regarding the correlation between online customer sentiment on the project price, volume traded and number of sales via running regression on the three factors against user sentiments.

We used a combination of methods to gather this data and then analyze it. We started with utilizing web analytics methods in order to collect the data from Opensea and twitter for an 8-day period during the month of November 2021. We chose the top 31 collections and crawled Opensea using their API in order to gather the NFT data for the previous 30 days when collected. There was a total of 24 attributes collected for each collection, but the key attributes chosen were Thirty Day Price Change, Thirty Day Volume, and Thirty-day Sale. We scraped tweets from twitter using Snscraper. Once the NFT data was scraped for each collection we pooled the data into excel to organize and clean it and create our NFT Masterfile.

In order to do the sentiment analysis, we used the Vader Sentiment Analyzer which iterated through the NFT collection data frame of tweets to get positive, negative, neutral and the compound polarity score to each tweet for each collection. We chose the time period of 26-34 days prior because of our hypothesized lag in reaction to sentiment that would reflect in the overall price of a collection and other performance measures. Following this we calculated the sentiment ratio for each project tweets respectively using the count of positive and negative tweets for each project. The sentiment ratio was determined by dividing the positive tweets by the negative tweets.

After we collected all the data and obtained the sentiment ratio along with the compound polarity score, we ran 4 different regression models in order to see if we could find a significant relationship between customer sentiment and NFT collection price, volume, and sales. After running the first regression, we removed outliers and were left with a total of 27 collections. The regression was run in R studio to obtain more information about the model’s accuracy and significance. With the first regression we had a low accuracy of 0.1637% and did not provide us with any support for our hypothesis. The second model also showed an insignificant negative relationship between the sentiment of tweets for a specific NFT collection and thirty-day price change of that collection. The third model we ran showed us that the sentiment ratio has an effect on the thirty-day trading volume for a collection. This model was accurate and showed us conclusive evidence that twitter sentiment is tied directly to NFT collection price and volume traded. The fourth model we ran also yielded significant results with the dependent variable being thirty-day sales and the independent variable as the average compound polarity score. Again, we see that twitter sentiment is showing a negative effect on a different performance measure for NFT collections.

Conclusively speaking, our methodology and analysis resulted in evidence that twitter sentiment impacts NFT collection volume and sales. Moving forward, this analysis can be expanded to gain better insights and accuracy by building a stronger dictionary using nomenclature commonly found in space and taking a larger volume of time instead of 30 days as we did for our preliminary analysis. Additionally, other online resources can also be utilized for further insights such as discord and Instagram.

**Problem Statement**

NFT collections have caught huge attention since people are buying these pieces of art and selling it for higher prices, eventually making a huge profit. NFT collections are pieces of digital art that are bought and sold on market platforms using Ethereum as the currency. Ethereum is the second largest crypto currency in terms of price and market cap. With our research, our main objective was to determine if, and what, social media sentiment impacts the prices of NFT collections. A common belief would be to assume that positive sentiment would have a positive impact on a collections price, we look to prove or disprove this hypothesis.

We analyzed a number of performance indicators for an NFT collection such as the number of sales a collection saw over the past month as well as the volume traded over the past month. We aim to see if the sentiment of collectors of NFT collections over social media influence these performance measures. Again, a normal assumption would be that positive discussion about a collection would lead to a higher number of sales for a collection and an increase in trading volume. We look to prove or disprove this belief with statistical analysis. Our statistical analysis will consist of linear regression modeling using several different combinations of inputs in search of statistically significant results.

**Business Goal Analysis**

The NFT marketplace has seen constant growth recently and is predicted to grow exponentially over the next couple of years. Our goal has been to analyze the market in great depth to find insights that can be used by individual and institutional investors to participate in trading.

The application of web analytics was a crucial element in analyzing the industry for sentiment analysis on sales, volume and prices of NFTs. Web Analytics was our go to for our data collection, web crawling was used to crawl one of the major marketplaces for NFTs - Opensea and Twitter for NFT statistics and sentiments of investors respectively. After which, we moved on to analyzing the data via linear regressions.

By analyzing the Opensea marketplace data for NFT collections and studying the impact of investor and collection sentiment about a NFT collection, we aim to provide valuable insights for these retail investors. This report could also be aimed towards people who have not yet invested in a NFT because of uncertainty. Our goal is to provide them with some certainty in regard to the behaviors of an NFT collections performance based on what is being said about it on twitter.

This analysis report also aimed towards creators or aspiring creators of an NFT collection. We hope this analysis could provide transparent interpretations in regards to the performance indicators of a collection so they can make adjustments to their NFT collections marketing strategy if the performance measures of their collection don't align with the results of our models.

Lastly, our analysis can be utilized by active traders within the market to analyze patterns and retrieve insights for real time trading purposes.

**Dataset Description**

We looked at the top NFT collections for the month of November 2021 as that was the chosen period of time for our analysis. We chose the top 31 collections which were as follows:

BoredApe, Sandbox, Punks, MutantApe, Meebits, Punks Comic, Rumble Kong, Voxies, Cool Cats, X Copy, Lazy Lions, Bored Ape Kennel Club, DeadFellaz, Fluffworld, Treeverse, Pudgy Penguins, Parallel Alpha, Curio Cards Wrapper, VeeFriends, Zed Run Official, The Dodge Pound, Gen Dot Art, Forgotten Runes, World of Women, Bored Ape Chemistry, Cyber Kongz, Sup Ducks, Super Rare, Gutter Cat, Cyber Kongz\_VX, Emblem Vault.

Post project selection, we moved to the data collection process of our project. We targeted two platforms: Opensea for NFT statistics which were later used for regression analysis and Twitter for NFT tweets which were later used for Sentiment Analysis.

The tools used for scraping Opensea was the Opensea API and HTTP requests. We scraped the platform for NFT statistics of each collection. Our data contained a number of attributes from which we performed data cleaning and recleaning after which we ended up with the most relevant attributes for our analysis. The key attributes chosen were Thirty Day Price Change, Thirty Day Volume, and Thirty-day Sales. Thirty Day Price change is interpreted as the percentage change in the NFT collection average price from 30 days prior to the scrapping of the statistic compared to the average price on the day we collected

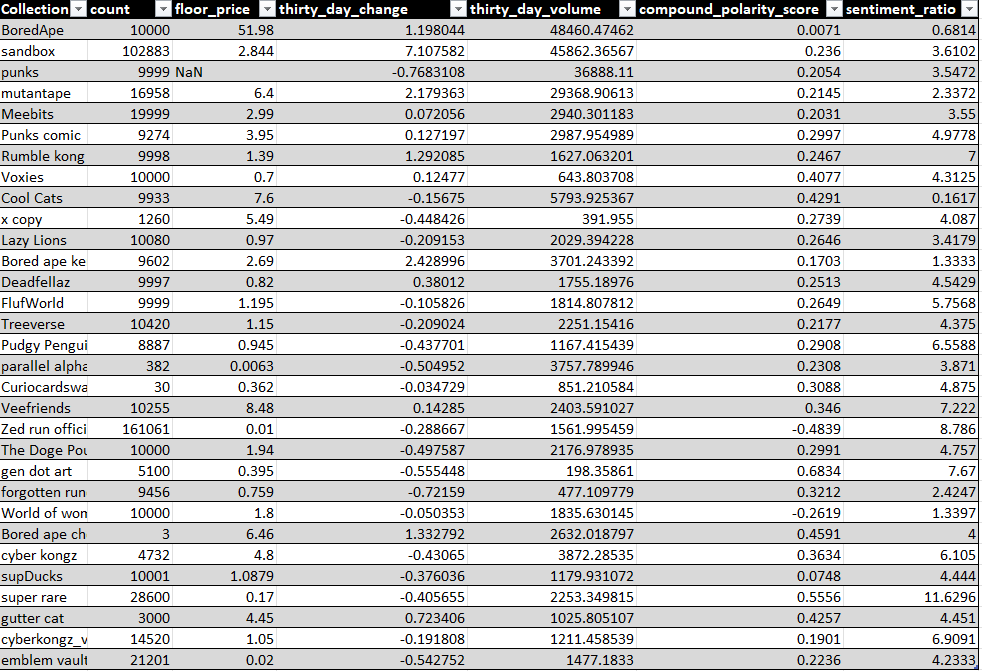
 The price is in terms of Ethereum not U.S dollars, so a percentage change in Ethereum is much more significant than dollars. One Ethereum is currently valued at $4,031. Figure 1 shows our data after we finished scraping Opensea:

Figure 1: NFT Statistics Collected from Opensea

We calculated statistics for our preliminary data, in Figure 2 and 3 you can see the Total Volume and Sales vs Volume by Collection respectively:

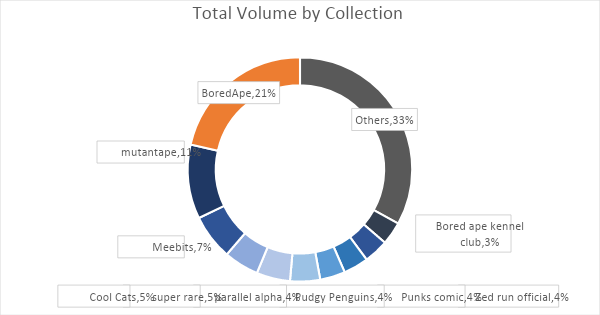


Figure 2: Total Volume by Collection

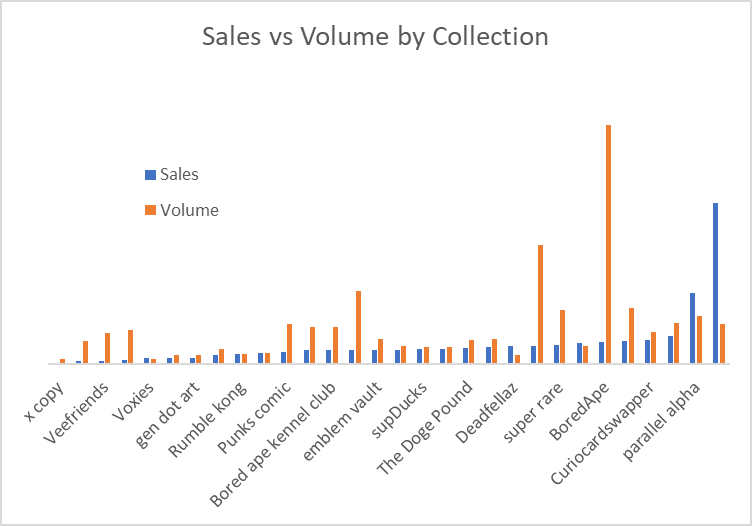


Figure 3: Sales vs Volume by Collection

We aimed to scrape tweets for an older time period to view the impact of sentiment on the price, volume and sales of the NFT projects. The biggest challenge we faced with Tweepy and Twint was the limitation where we could not pull tweets past a week. Hence, we went ahead with using Snscraper to scrape Twitter. These tweets were then used for sentiment analysis. We used the lexicon rule-based sentiment analysis library package titled Vader Sentiment Intensity Analyzer. This was used to assign a positive, negative or neutral score to each tweet. Which was then used to calculate a compound polarity score and the sentiment ratio.

The tweets were collected from November 4th, 2021 to November 12th, 2021. We chose these dates because the date in which we collected the NFT summary statistics was December 8th so 30 days prior is November 8th. Therefore, we collected tweets for 4 days prior and 4 days after this date to allow for some variation in tweets or changes in beliefs. This also will help mitigate retweets being calculated more than once into the overall average compound polarity score for an NFT collections list of tweets. We will go into more detail regarding the average compound polarity score and the sentiment ratio in the system design.

**System Design**

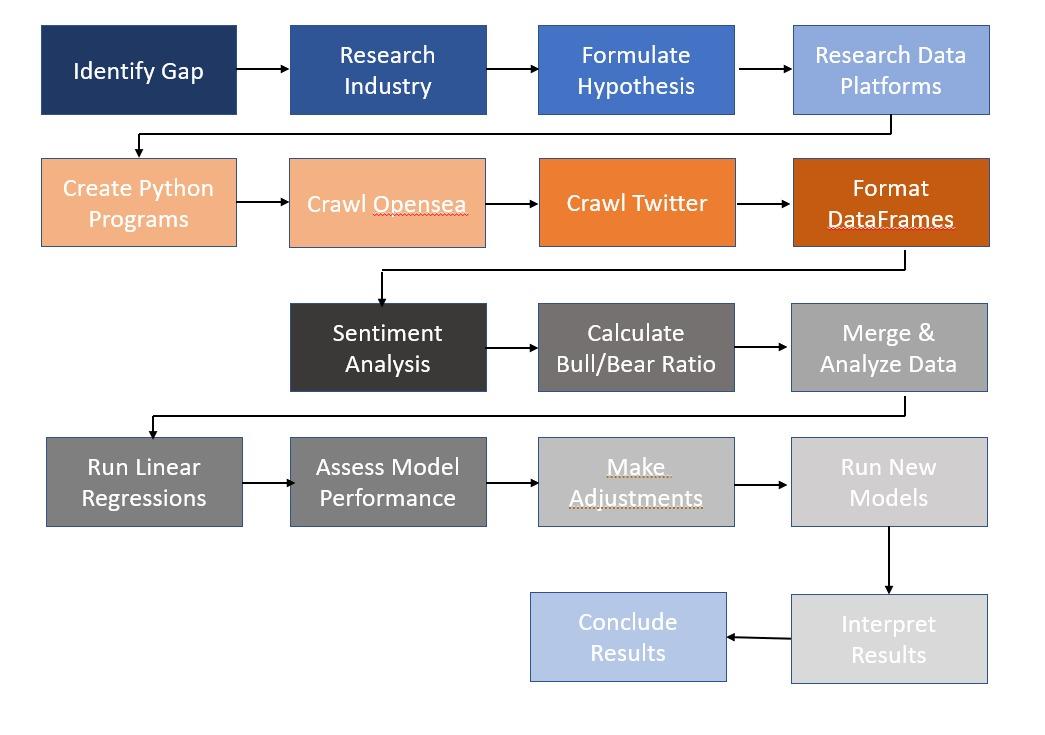
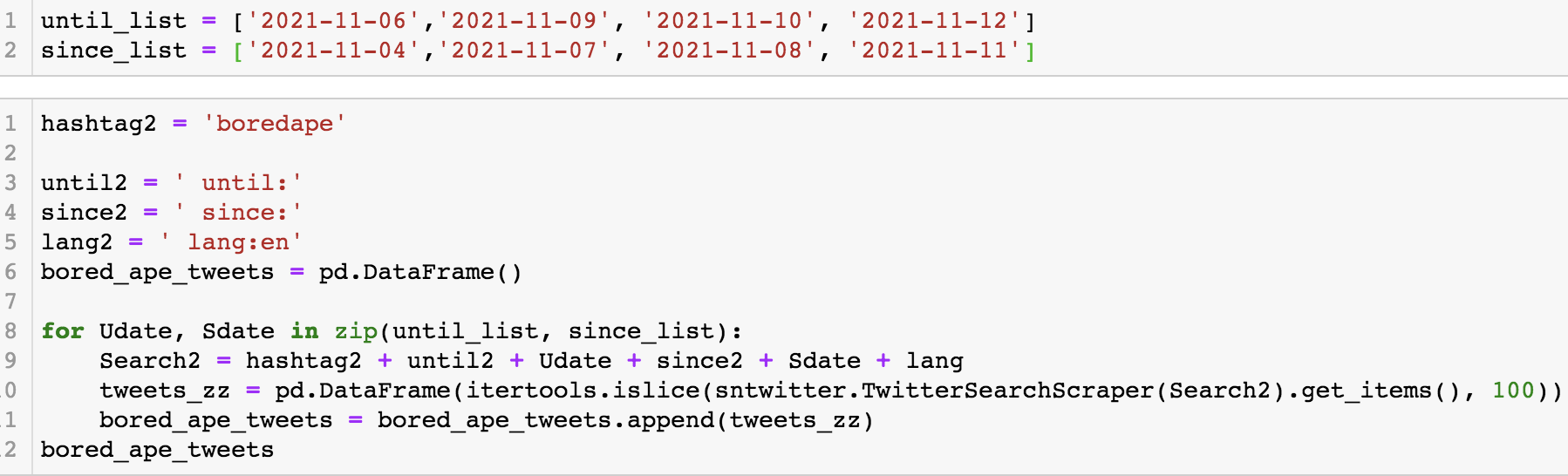
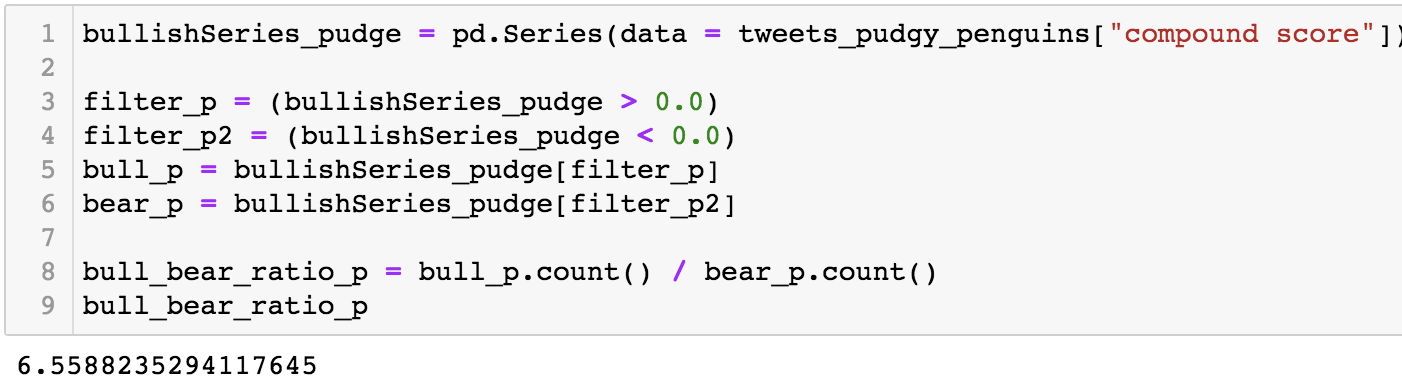


Figure 4: System Design Process

As mentioned above before we were able to build our models, we had to first decide what NFT projects should be included in the dataset. After considering the possibility of looking at some very popular and some unpopular ones we came to the conclusion that this kind of data collection and the subsequent results of our models may mislead NFT projects not at the top or not at the bottom. This is why we ultimately decided on choosing only the top 30 NFT projects. Therefore, NFT project startups and ones still just gaining traction can disregard the results of this analysis as it is only geared towards these high-volume collections.

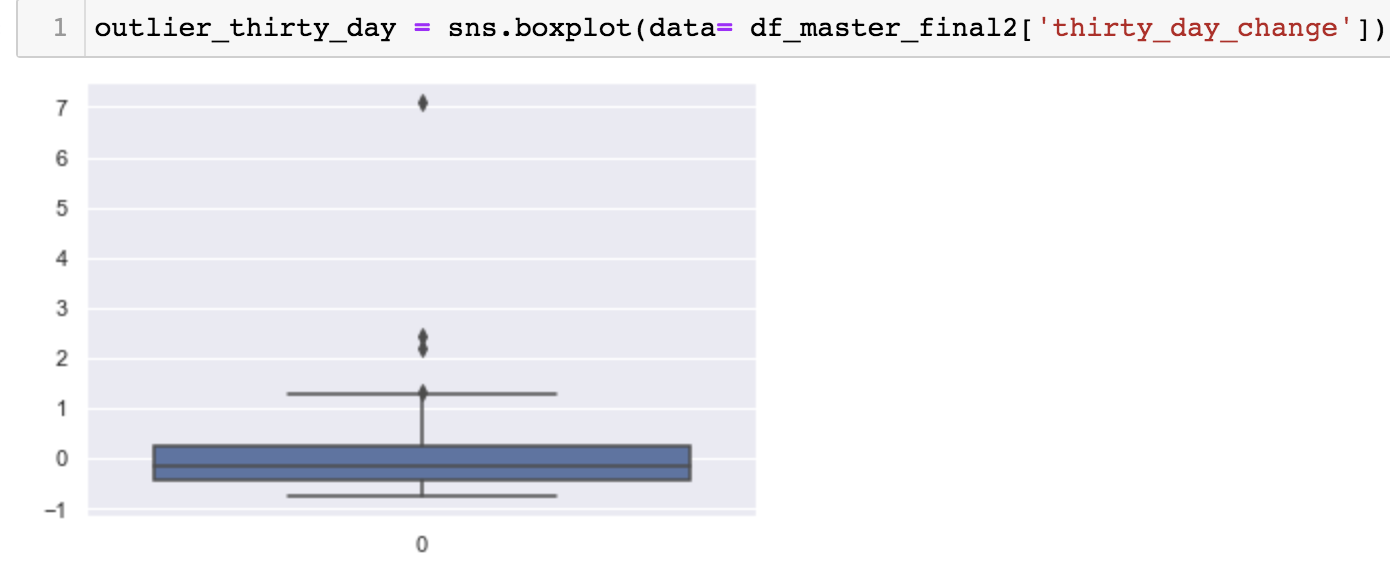
After we collected the summary statistics, we decided on the optimal time interval to collect tweets for that would give these NFT statistics enough time to see the effects of those tweets within those statistics. After consideration we choose 28-34 days ago. This not only was a longer duration of time to see the effects, but a longer period of time also cancels out some day-to-day noise we would see in our model had we chosen just the 7-day period.

We then decided on what sentiment scores we would use in our models output from the Vader Sentiment analyzer. First, we considered taking the average for the positive scores and negative scores of all tweets per collection and then finding the average of those two averages. Ultimately, we decided to just average out the compound polarity scores for all tweets per collection. Although this wasn't enough, we needed another measurement indicator with a different scale. After doing some research we came to the idea of calculating a sentiment ratio for all tweets of each collection. This ratio is often used when conducting sentiment analysis for stock price movements. Which is why it was given the name Bull-Bear ratio. To calculate this ratio, we counted the total number of positive tweets for a project (bullish ratio) and the total number of negative tweets for a project (bearish ratio). Where the positive tweets are the numerator, and the negative tweets are the denominator.

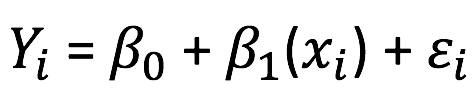
After calculating these two new attributes per collection we combined it with the summary statistics dataset. Finally, we were able to begin developing our statistical models. Through trial and error and tedious tinkering with several different models we were able to find some interesting results. The attributes to be used for our methodology was the 30-day price change (percentage terms), 30-day volume, 30-day sales, the average compound polarity score, and the sentiment ratio for each NFT collection.

**Initial Methodology Implementation**

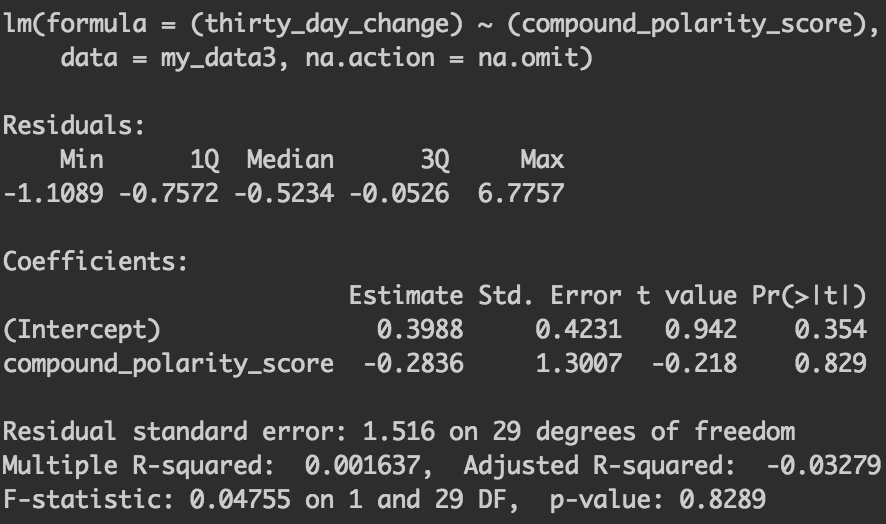
Our main objective in our research on NFT’s was to quantify the effects of overall sentiment for a collection. We aimed to determine if this sentiment had an impact on the prices for an NFT collection. We firmly believed that our model should be able to capture a significant positive or negative relationship with a NFT collection's price and the sentiment of that collection. The NFT market has been growing and becoming increasingly complex. Although, still nowhere near the size or complexity of the stock market. Besides the fluctuation in Ethereum prices, it is our belief that there are only a few other variables that could be affecting an NFT collections price. A couple of them being investor and collector sentiment. We also believe that the social media sentiment would not have an instantaneous effect on the price. It takes time for word to spread about an art piece from a collection being a good purchase. Unlike a stock price, where both public and private information is believed to be already factored into the price according to the efficient market hypothesis. We choose to look at tweets from 26-34 days prior to the collection of the summary statistics for the NFT collections. We chose this time period because of our hypothesized lag in reaction to sentiment that would reflect in the overall price of a collection.

 After calculating the average compound polarity score out of all the tweets for each collection and appending it to the NFT summary statistics the next stop was to build the linear regression model. Before doing so we wanted to find any outliers in our dataset that may have seen abnormal changes in price over the previous 30 days. This box and whiskers plot showed 4 outliers which we then removed. The NFT collections removed were Bored Ape Kennel Club, Sandbox, Mutant Ape, and Bored Ape Chemistry club. Three of these 4 as you can see by the names are sub collections of the Bored Ape Yacht Club. There are very few art pieces in these collections and as a result the prices are high and only get higher as the supply decreases.

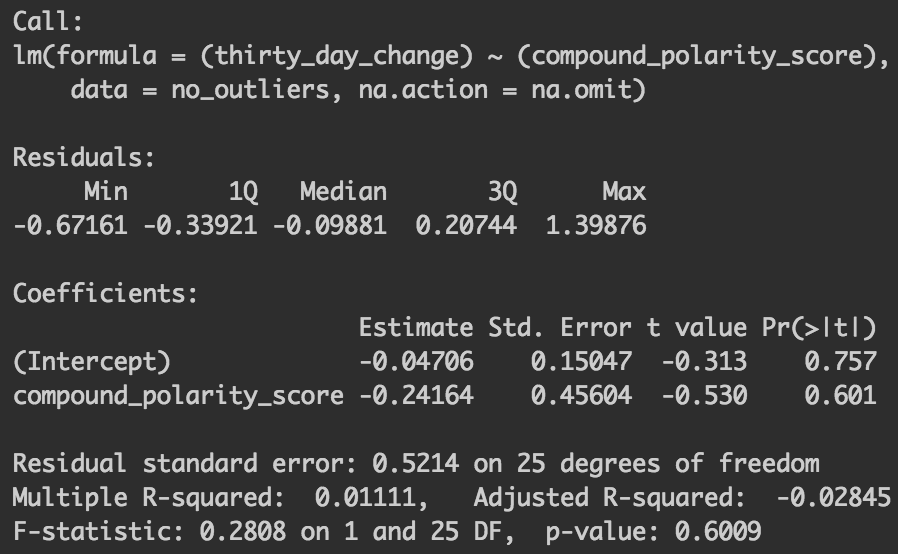
We now have a new dataset of 27 collections that all have seen relatively normal price changes in the last 30 days in comparison to one another.I have chosen to run these regressions in R studio to obtain more information about the model’s accuracy and significance.

Y = thirty-day percentage price change 

X = average compound polarity score



The first model shown to the right is with all the collections including the outliers. As you can see in the results the multiple R-squared is .001637. This is interpreted as .1637% of the thirty-day percentage price changes variance can be explained by the variance of the average compound polarity score. This accuracy is extremely low. Consequently, it does not provide us any evidence that would support our hypothesis of social sentiment having an effect on an NFT collections price.

Now using the dataset with the 4 price change outliers omitted we see similar results, but with a 578.68% increase in accuracy (the percentage change of the R-squared compared to previous model). Now this model shows that 1.11% of the thirty-day percentage price change variance can be explained by the variance of the average compound polarity score. In regard to statistically significant interpretations of the coefficient for the average compound polarity score, it has again failed to pass any t-tests to provide us with conclusive evidence about the effect twitter sentiment has on NFT collection prices. It failed the t-test for an 80% confidence interval by a large margin.

One sided t-test with 25 degrees of freedom

T-value = -.530, T-stat = .856

|-.530| < .856

Therefore, the null hypothesis that the results are significant at a 80% confidence interval is rejected.

The story is also the same when running the regression with thirty-day percentage change as the dependent variable and the sentiment ratio as the independent variable. The model yields insignificant results with or without NFT collection outliers.

**Refined Methodology Implementation**

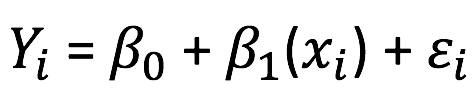
Price may be the most determinant factor of an NFT collection's success in the present day. Although, there are some other influencing factors when it comes to a collections long term growth and sustainability. Trading volume and number of sales are two other important measurement indicators for an NFT collections overall performance. An increase in either can mean an increase in popularity as more people list their digital assets and agree to sell when the right buyer bids enough to reach the seller's desired profit margin. A decrease in the number of sales or overall trading volume could signal that the NFT collection is losing momentum and that its digital bubble may soon pop. When the number of sales is decreasing it can either mean that people are listing their digital assets but are not getting the desired offers that they want. It could also mean that buyers aren't selling during a certain period of time because of the volatility the currency is traded with (Ethereum). A similar interpretation could be made for a decrease in trading volume. If an NFT collection is showing a decline in trading volume it could be caused by sellers not getting the offers they desire or that buyers aren't selling because of the volatility in Ethereum.

The sentiment from investors and collectors of an NFT collection could also have an impact on its trading volume or number of sales. We have run 2 more linear regressions in an attempt to quantify the effects twitter sentiment has on trading volume and number of sales for an NFT collection.

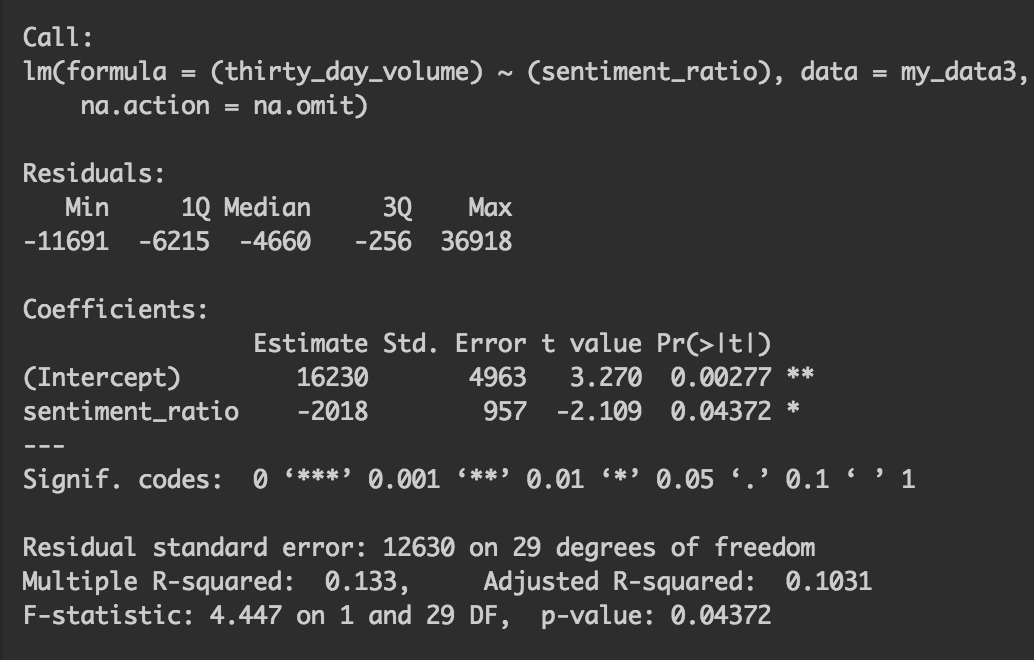
*Relationship between 30 Day Trading Volume and the Twitter Sentiment Ratio:*

Now for this first linear regression model's independent variable we chose to utilize our other measurement indicator of sentiment – the sentiment ratio (bull/bear ratio). This is a different measure as the average compound polarity score of all the tweets compound polarity scores for a given collection. To reiterate, for an NFT collection this sentiment ratio is calculated by counting the number of tweets that had a compound polarity score greater than 0. These are classified as bullish tweets and are the numerator. The bearish tweets counted have a compound polarity score of less than 0 and are the denominator of ratio. Any tweet for a given NFT collection that has a compound polarity score of 0 (summated as neutral) will not be counted. If the overall sentiment ratio for an NFT collection is greater than 1 than all the tweets for that NFT collection are classified as predominantly bullish, if it is less than 1 it is considered that all the tweets for that NFT collection are classified predominantly bearish. A larger positive number will obviously be classified as showing very bullish (or positive) sentiment for a particular NFT collection. Even though the average compound polarity score and the sentiment ratio are strongly correlated they still hold different weights and meaning in a linear regression model.

For this linear regression model, we have the trading volume over the last 30 days as the dependent variable (Y). We chose to use the sentiment ratio as the independent variable because of the larger scale and variance that it holds due to the difference in calculation compared to the average compound polarity score. This aligns with the wide variation of the dependent variable for this model (trading volume). Some collections may have many trades, but the volume is the summation of the Ethereum used in all of these trades for a given collection. Therefore, if a collection has a large number of sales but each digital asset is not sold for that much Ethereum then the trading volume will still not be that high. A collection could have fewer digital assets being sold but if the cost per asset is very high then the trading volume will still be relatively high. This similarity in calculation of the dependent variable in comparison to the calculation of sentiment ratio is why we chose this measurement indicator as the independent variable for this particular model.

Y = Trading Volume over the last 30 days for a given NFT collection

X = Sentiment ratio of tweets 26-34 days ago for a given NFT collection

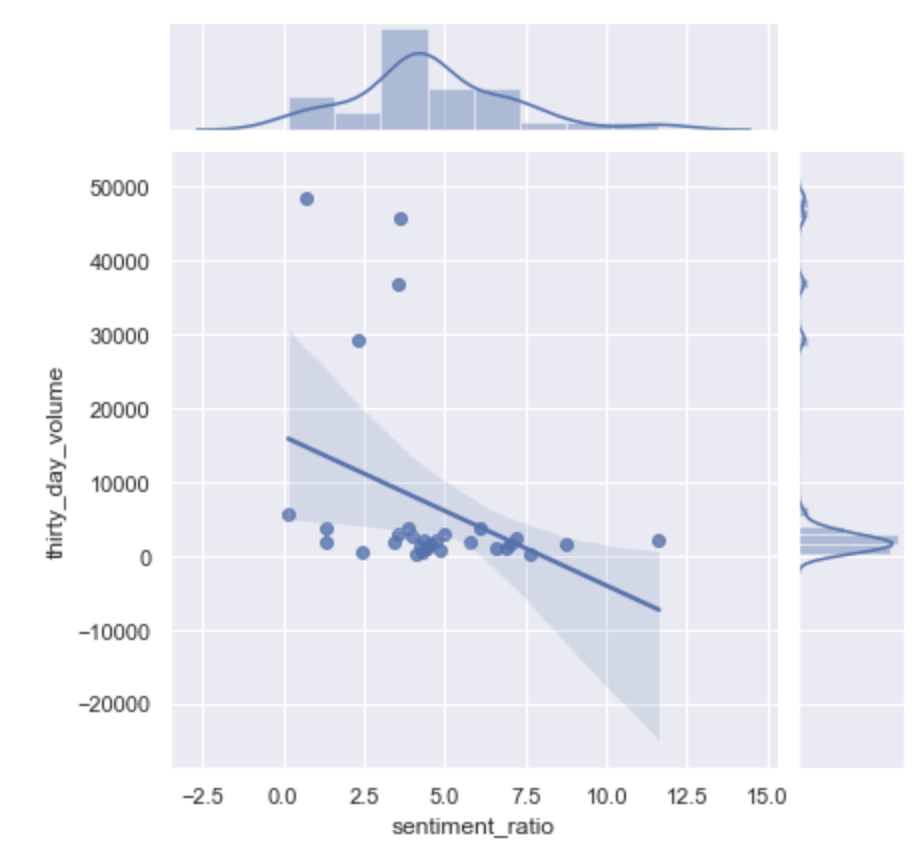
Interpretations of results for this model:

In this model shown to the right you can see a significant negative relationship with the sentiment ratio of tweets for an NFT collection and the thirty-day volume of that NFT collection. For example, when inputting the sentiment ratio for the Cyber Kongz collection as X to be multiplied by the sentiment ratio coefficient plus the intercept.

The model outputs an estimated value for the Cyber Kongz collection past 30-day volume:

16230 + -2018(6.105) = 3910.11 Ethereum

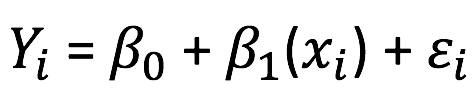
The 30-day volume traded can never be negative but theoretically it could be 0 if it was an irrelevant NFT collection. We looked at the Top 31 most popular NFT collections so the 30-day volume traded values of each NFT collection in our dataset shows 30-day volume traded values of 198.35 or higher. This model's estimation of the 30-day volume traded for the Cyber Kongz collection was very accurate. The actual value of the 30-day volume traded for Cyber Kongz was 3872.28 Ethereum. The significance code meanings are shown in the output for the model. You can see that the sentiment ratio having this negative relationship with the 30-day volume can be interpreted as significant for a 90% confidence interval.

 Here is the fitted regression line of this model on the right. You can see there are 4 noticeable outliers with extremely high trading volume compared to other top NFT collections. These have been the consistent top 4 NFT collections since the origination of the Opensea marketplace. Which is the Bored Ape Yacht Club, Sandbox, CryptoPunks, and Mutant Ape. We attempted to omit two of these collections in our initial methodology (Sandbox, Mutant Ape). Since this model for trading volume still produces significant results it is best to keep them. Omitting data unnecessarily can lead to questions of biases in regard to the accuracy of a given model and the artificial way in which it was obtained.

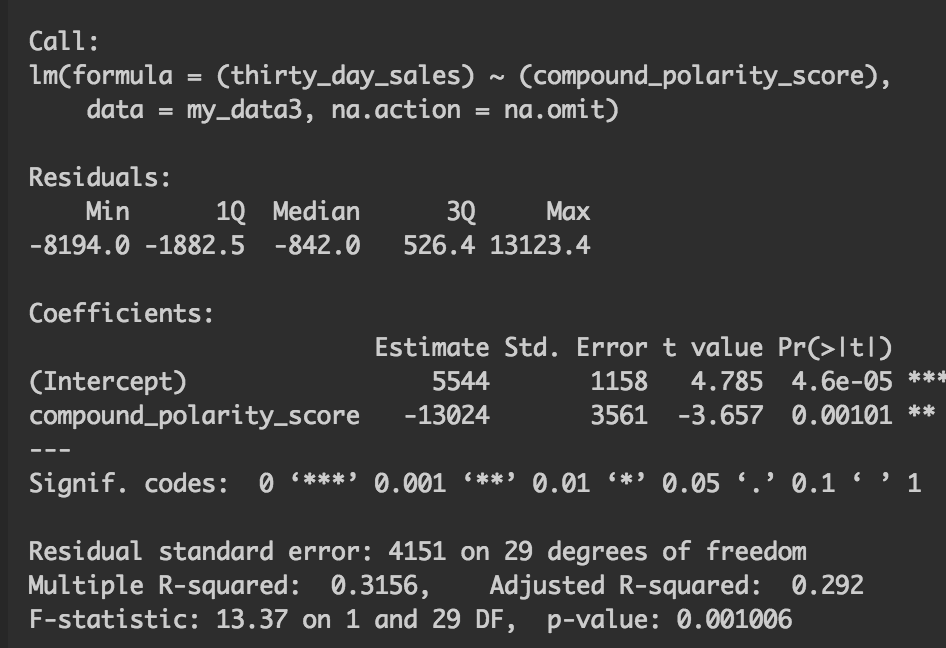
*Relationship between 30-day Number of Sales and the Tweets Average Compound Polarity Score:*

For the last linear regression model built, it’s aim is to determine the relationship between the number of sales an NFT collection saw over the past 30 days and the average compound polarity score of tweets about that collection from 26-34 days ago. The choice to use the average compound polarity score as the independent variable was again, based on logic. Volume traded was a summation of the cost of each collection and the sentiment ratio was a count summation of each score that was positive or negative. Now for this model, the number of sales is not in terms of total Ethereum but just a count of the number of sales a collection saw. The average compound polarity score is just one value that represents the entire frame of tweets for a collection. The variation of both of these variables is smaller which is why it makes sense to be used for the model.

This model is represented as:

Y = Number of Sales over the last 30 days 

X = Average compound Polarity Score



Interpretations of results for this model:

The outputted results show a similar trend as the 30-day trading volume regression on the sentiment ratio. This would make logical sense since the trading volume and number of sales are correlated. The sentiment ratio and average compound polarity score are also correlated. The output shows a significant negative relationship between the preceding average compound polarity score of tweets for an NFT collection and its subsequent number of sales over the course of the month. We are able to state this relationship is significant for a 95% confidence interval. As the two significance coded stars represent that 95% confidence interval.

For an example of the interpretation of this model we will use the average compound polarity score of the NFT collection: DeadFellaz. Its average compound polarity score for all tweets from 26-34 days was .2513.

Y = 5544 + (-13024)(.2513)

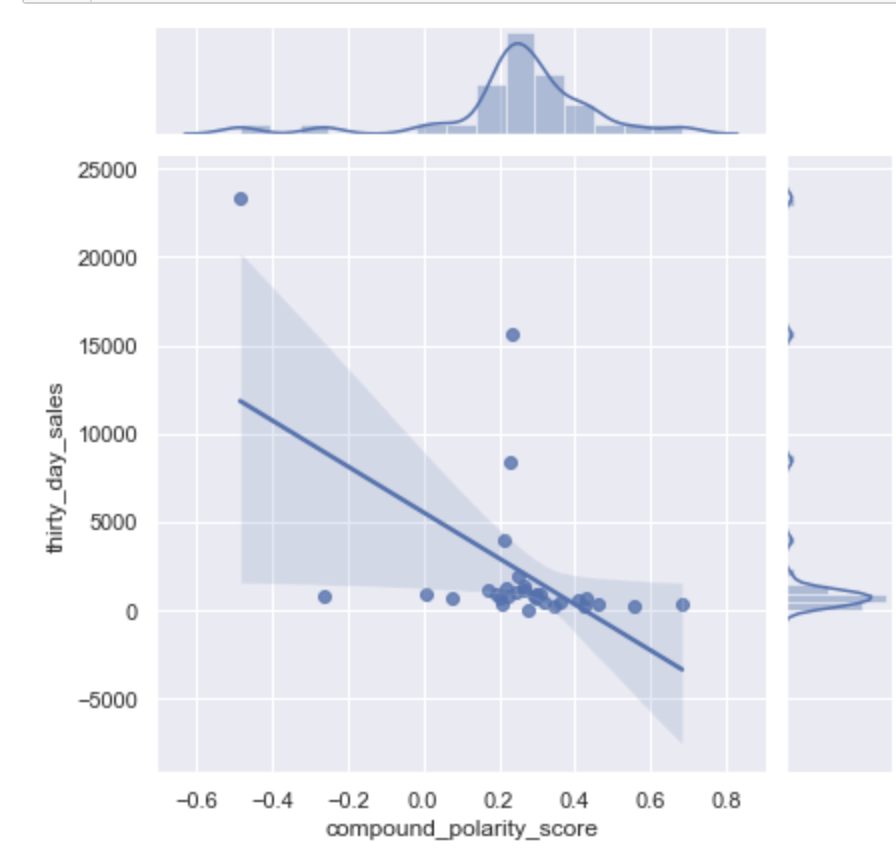
Y = 2271

Where:

Y = DeadFellaz 30-day number of sales,

X = DeadFellaz average compound polarity score

Therefore, the model’s calculation estimated DeadFellaz would sell 2,271 digital assets based on twitter sentiment for the collection. The actual number of sales the collection saw was 1,955. There is certainly some measurement error, but this model is still accurate enough to make conclusive insights. If the average compound polarity score was negative it would have increased the model’s estimation of the number of sales for the collection.



As you can see the fitted regression lines show a lot of variance for NFT collections that have a negative average compound polarity score. This variance decreases dramatically as the compound polarity score variable goes to .2 and starts to widen again as average compound polarity scores go towards .8. Although this variance is not nearly as large as the variance seen with the negative average compound polarity scores.

**Evaluation**

For our NFT sentiment analysis our findings were inconclusive. We found unexpected results, and this can be due to multiple factors. First and foremost, this is a brand-new field where unforeseen external factors can sway the prices of large NFT collections. When looking at stock prices, sentiment can be taken into consideration due to the historical data largely available to the public. With the NFT space being so small a single individual could potentially sway prices and cause a collection to boom or fail at a whim. With more information and a stronger dictionary, we might see better results. Our regression analysis did a wonderful job of showing the correlation between what we used and basically gave us a road map to better understand the market as it is now. If someone were to do a similar analysis, they could take our findings as a baseline and build on top of it**.**

**Conclusion and Future Direction**

The results of our regression analysis yielded unexpected results. Based on these results we can conclude that when tweet sentiment has a bullish ratio you can expect to see a decrease in the volume traded for that NFT collection. This could be due to investor beliefs that the NFT collection is overhyped and/or overpriced or this could be due to the fact that people start holding on to the collections that they have hoping to sell it at a greater price for a greater return.

The results of our regression model with 30 days sales as the dependent variable and the average compound polarity score also yielded the same unexpected results. Based on the results we can conclude that when the sentiment of tweets for an NFT collection is positive the number of sales will decrease over the next month. This could also be because NFT collectors are hesitant to sell their digital asset when the sentiment is high. In hopes that the price will continue to increase over the long run.

These results could be biased. NFTs are purchased with Ethereum. Based on the fact that Ethereum prices have been steadily dropping over the last month. This uncertainty in value for the currency NFTs are traded with may also be decreasing sales or decreasing the value of each sale for an NFT collection because people don’t want to sell it for a less dollar amount so they start holding hoping that the price would rise back, and they would then sell it in order to maximize their gains.

The low compound polarity score for Bored Ape calculated by the Vader sentiment analyzer could be due to this analyzer interpreting the word ‘Bored’ as a negative score. But even after excluding bored ape in the regression analysis the results are still the same.

What we could do in the future is take a look at a larger time interval such as 6 months or a year when that data is available and see if we can find a positive correlation. We could also look at creating a more in-depth dictionary that takes into consideration the vast terminology used by NFT investors to see if we can find a greater correlation. Going forward we could also potentially take into consideration the volatility of the cryptocurrency market as a whole and look into other online avenues where NFTs are discussed such as discord and reddit.

**References**

<http://ceur-ws.org/Vol-2989/long_paper10.pdf>

<https://medium.com/atoti/is-it-enough-to-focus-just-on-twitter-sentiment-9e1f65a867c2> (Bull/Bear ratio idea)

<https://www.finder.com/cryptocurrency/nft-collections>

<https://docs.opensea.io/reference/retrieving-collection-stats>

<https://towardsdatascience.com/data-science-on-blockchain-with-r-afaf09f7578c>

<https://www.nature.com/articles/s41598-021-00053-8>

<https://www.analyticsvidhya.com/blog/2021/01/sentiment-analysis-vader-or-textblob/>