

# Twitter Sentiment Analysis for Cryptos

Eleftheria Trigeni ID 3308210043

Konstantinos Papatzimos ID 3308210032



# Introduction

Cryptocurrencies (or cryptos) consists a way of exchange that exists only in digital form. These days more and more people, even the young ages, invest on them because in the next years it is expected that their value will grow significantly.

There are many cryptos like Bitcoin, Ethereum, Binance coin (BNB), Dogecoin, Cardano and many more. One of their main characteristics is that they don't have a fixed price. Their price depends on the demand. So, practical, people determine their price and so their value. We decided to do twitter sentiment analysis for three Bitcoins and particularly for Bitcoin (which is the most famous crypto), BNB and Ethereum. Our aim is, using tweets that we extracted from twitter, to examine the sentiments of people about these coins. We wondered if people have the same sentiments about these three very common cryptos. Because of the fact that their price increases and decreases daily, we considered that may the sentiments of the people change according the cryptos' value.

## Methodology

We used Twitter API and in combination with Python we managed to collect 3707 tweets. Actually we collected more than these but there were a lot of duplicates that we abstracted in the sequel. The reason that we had duplicates was that the date was in a datetime format. So even we wrote a piece of code to avoid duplicates, because of the fact that every second the seconds in the date were changing, every new record was considered as unique. An other issue that caused duplicates is that every second the tweets get likes. So again, every record was considered as unique because it had a different number of likes even if it had the same text and was written by the same user.

So, we removed duplicates and carried on with the 3707 unique tweets. We did data preprocessing and we removed urls, emoticons and unwanted punctuations.

Our tweets dataset is consisted of the following attributes: text, username, user description, number of user's followers, likes, number of user's tweets, number of account's retweets, date that tweet was created, location, hashtags, sentiment and subjectivity. Subjectivity indicates if a tweet is a personal opinion so in this case the tweet is considered as subjective, or it refers in factual information, so the tweet is considered as objective.

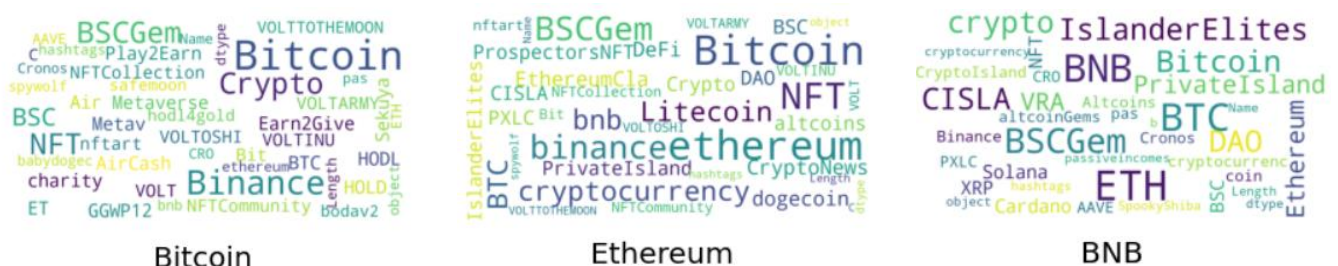
Using TextBlob we classified the tweets according to the sentiment as negative, positive or neutral and according to the subjectivity as objective, subjective or neutral. Also we used Matplotlib and Seaborn frameworks in order to create our plots.

In our MySQL database are stored 6838 tweets but we used only the unique 3707 tweets for the analysis. So all the below findings are based on these 3707 tweets. From them, 1762 have the hashtag Bitcoin, 1452 have the hashtag Ethereum and 1073 have the hashtag BNB. Bitcoin is the most popular cryptocurrency and that's why more tweets contain this hashtag as we will see in the sequence.

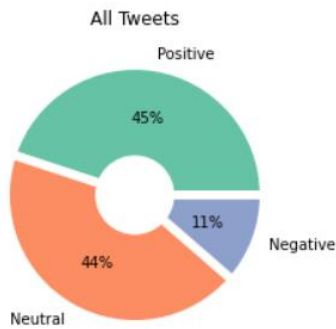
We collected these tweets in a period of 4 days (23/02/2022-26/02/22). In the next section we will present our findings.

## Results

To start with, we created some Wordclouds about the hashtags. Wordclouds contain the most common words of a text, so in our case, they contain the most common hashtags for every dataset. The purpose of this research was to examine what hashtags someone uses when writing a tweet about bitcoin for example, except of the hashtag #Bitcoin. According to our results, the hashtag «Bitcoin» is the most common. People use this hashtag when they refer to any type of crypto. It doesn't makes sense because many people identify cryptos with Bitcoin. So we just confirm this fact with our results. However, it seems that people do not write tweets only about a particular crypto. They write general comments that refer to many of cryptos, so they use several hashtags. That's why, for example, the hashtag #Ethereum appears in the tweets that refer to the Binance coin for example, or the reverse.

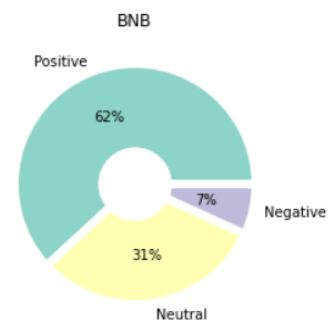
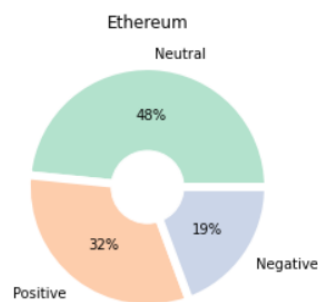
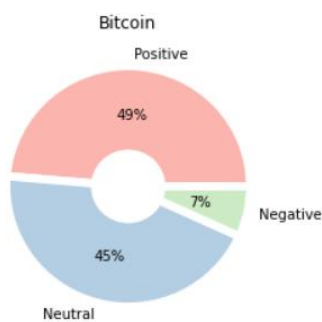


The next step is to examine people's sentiment for each crypto. Our goal is to determine if there is a significant difference between these three cryptos.

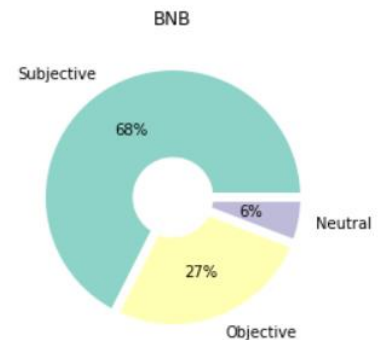
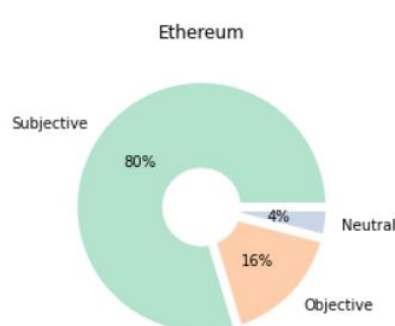
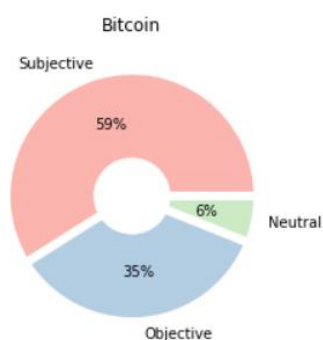


Firstly we take a first look at the sentiments of all the tweets and also at each hashtag individually. In total, as we see in the first plot nearly all of the sentiments are considered as positive or neutral. Only 11% have a negative content.

For Bitcoin, the percentage of positives is very close to the neutral percentage. The negatives are only 7%. Worthly is the fact that only 32% of sentiments for the Ethereum cryptocurrency are positive, while 19% are negative. The price of Ethereum probably decreased during our research, which led to the reduction of its value, which caused some negative reactions. On the other hand, BNB has the highest percentage of positive sentiments, so its value was probably high, resulting in increased profitability for those who have invested.



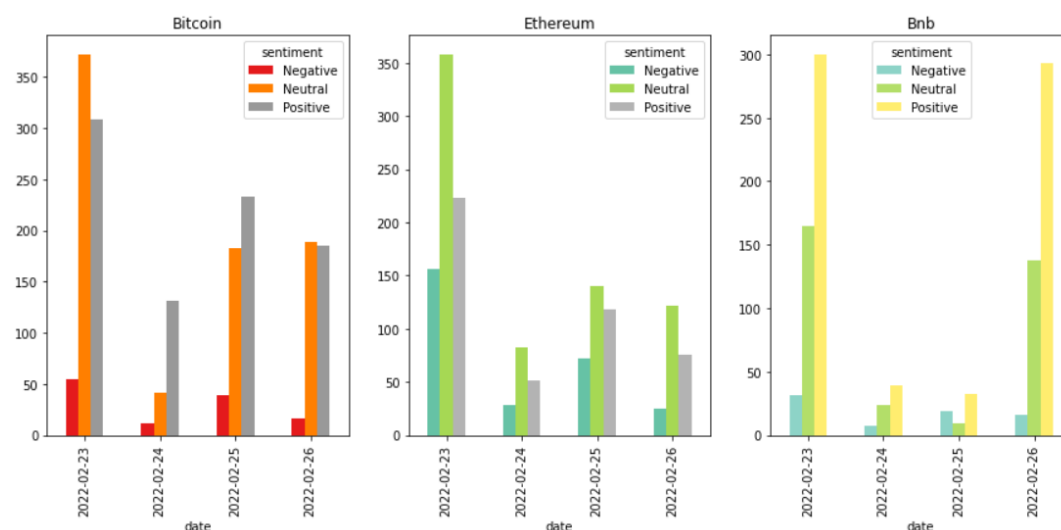
We created the same plots for the subjectivity of tweets for each crypto. Regarding Ethereum, a quite significant number of tweets are subjective. In combination with the fact that Ethereum has not many positive sentiments, may a drop of its price led people to express their personal opinion and complaints. BNB follows a similar pattern. More specifically, BNB has many positive sentiments, and as a result, many people have expressed their personal positive opinions. Consequently, the percentage of subjectivity increased.



As we mentioned above, we collected tweets from 23/02/2022 until 26/02/2022. We would like to see how the sentiments of people change by the time. Having a first look, it seems that irrespective of the date, the most tweets identified as neutral or positive.

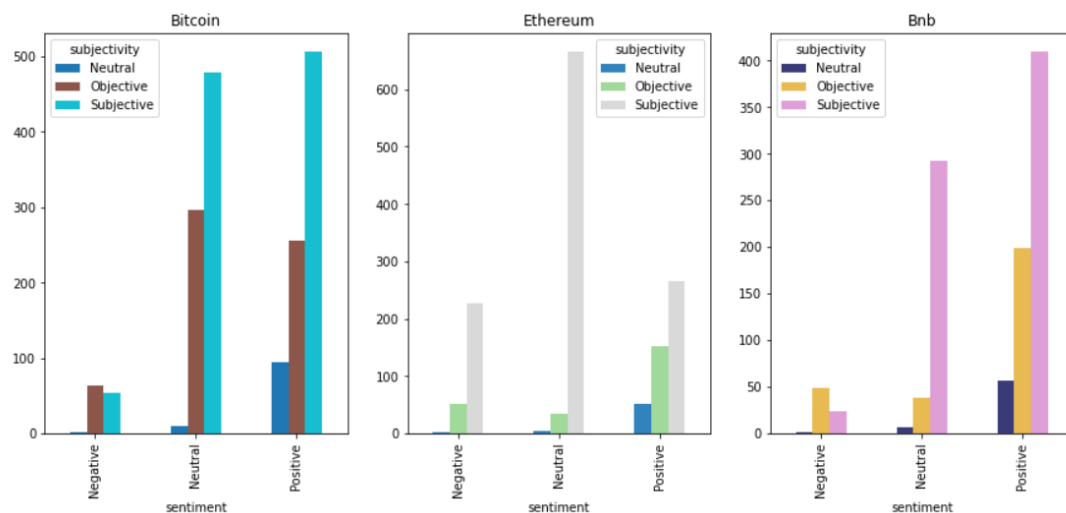


Let's see how the sentiments of people change by time for each of the cryptos individually. We begin with bitcoin. In general how matters the date, the most sentiments are neutral or positives. On 23th of February the most sentiments were neutral but on 24th and 25th of February the most sentiments were positive. On 26 of February it is not so clear. So we could say that may on 24 and 25 of February the value of Bitcoin increased. To talk about Ethereum, the most sentiments are neutral inspite of the date. In contrast with Bitcoin, as we mentioned above, there are and several negative sentiments and mainly on tweets that have been created on 23 and 25 of February. Finally, in this diagram it seems clearly the positives sentiments for the Bnb crypto. As we notice, the most tweets about Bnb that we collected were created on 23 and on 26 of February and the difference between positive and negative sentiments is significant.

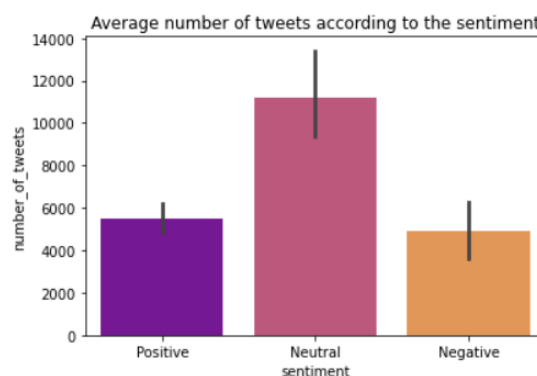


Next, we present our analysis that aims to examine if the subjectivity of a tweet affects the sentiment of it. We did this analysis for each crypto individually. We conclude that,

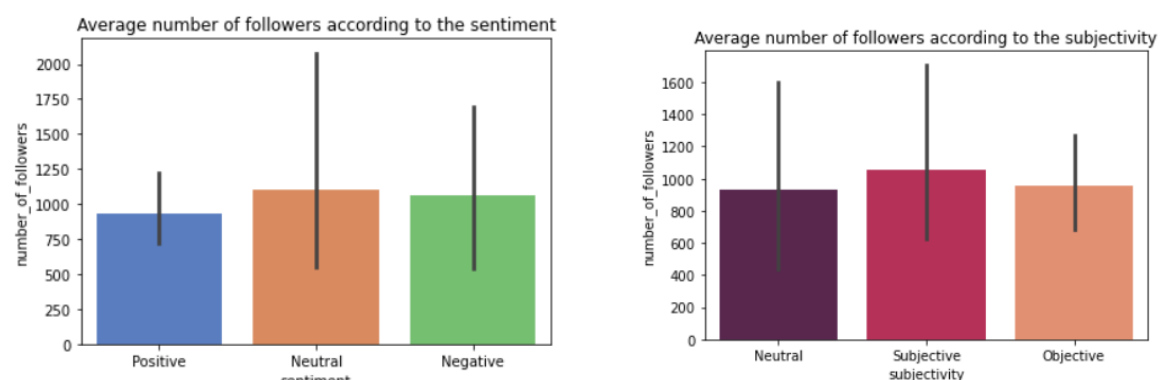
for our analysis, the sentiment does not affect whether the tweets are objective or subjective. For all the tweets and despite the sentiment, most tweets are considered subjective.



Another issue that we would like to research is whether the average number of tweets that a Twitter account has affects the sentiment of the tweets. In other words, if someone write constantly tweets, do these tweets have positive or negative content? The below diagram indicates that the tweets that have a neutral sentiment were created by users that write tweets very often. We can conclude that a user who is very active on Twitter tries to keep a balance and not to express his negative or positive sentiments.



In the sequel, we present our analysis to determine if the number of users' followers has an impact on their tweets' content. According to the below plots, we can say that this claim cannot be confirmed.



# Findings

We noticed that bitcoin is inextricably linked to the term cryptocurrency, and as a result, it appears as a hashtag in other cryptocurrencies such as Ethereum or BNB. Additionally, we discovered that tweets with a favorable or negative sentiment are often subjective. This is due to the fact that the price and value of cryptos are intimately connected. As a result, a higher price equates to a higher value, resulting in more returns for investors. When the value of a cryptocurrency falls, the opposite occurs. People communicate their personal perspectives on their earnings in a sequential manner, resulting in a high degree of subjectivity. Furthermore, Cryptocurrencies (Bitcoin, Ethereum, and BNB) have generally favorable and neutral sentiments. BNB, in particular, appears to be the most popular and profitable crypto for consumers during the timeframe of our analysis, with a largely positive sentiment. Bitcoin's popularity has remained very consistent, with only a small number of negative remarks and mostly favorable and neutral sentiments. Finally, Ethereum has the highest percentage of both negative and positive sentiments, and it is largely neutral.

# Discussion

We conducted a thorough sentiment analysis of tweets mentioning three popular cryptos. An idea for further research is to run our code on a regular basis and gather tweets to observe how people's attitudes evolve over time. It would be fascinating to track how the value of each cryptocurrency fluctuates day to day, then conduct this sentiment analysis to see if our findings match the actual changes. To clarify, it would be fascinating to know if a reduction in the price of a cryptocurrency creates negative feelings in individuals, and vice versa, if a rise in the price of a cryptocurrency causes positive feelings in people. Undoubtedly, this research could have huge interest if it was conducted for a greater variety of cryptos, as well as if we could extract results based on the locations.