# CS579 Project 1: Social Media Data Analysis

Priyam Dinesh Shah, William E Bodell III

Group - 20

## I. Project Details

**Title: Verifying the price hype for NFTs over the social network to understand the cause**

**Short Description**

We believe that many NFTs are blindly hyped by creators and marketers to make money. There are a number of major influencers on Twitter who ill-market the NFT based digital arts to their followers, often being paid to do so. We’re planning to correlate the price of a NFT use case ”ëthermon” and its tweets from the users to check the hype around it over a period of time during its launch and steep fall. During this analysis we are trying to uncover any trends which might suggest why this NFT projects have experienced organic growth and how prices got influenced by high-profile shilling.

**Price Chart of Ethermon (**[**https://coinmarketcap.com/currencies/ethermon/**](https://coinmarketcap.com/currencies/ethermon/)**)**



Based on the price chart we considered our analysis time period in two phases

1. Ethermon Project Launch Phase ( May’21)
2. Massive Price decline phase ( Jun’21)

### 1. Data Collection and Processing

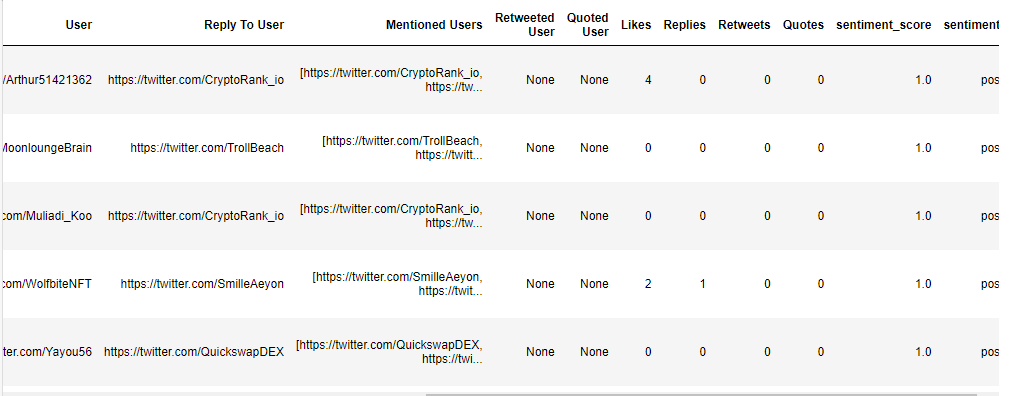
*We selected Twitter as a social media platform for our project. Initially we tried collecting data using the Twitter API, but it was not helpful for historical data collection as we required data of previous year. We later also tried applying for the academic research api access but had no luck in getting credentials for it. So finally, we used the snscrape library of python for scraping the historical data for specific time periods and a given query.*

*We scraped the following data using snscrape and stored them in a dataframe for both the time periods. The data collected is stored in dataframe and is used on the fly for all below calculations.*

**

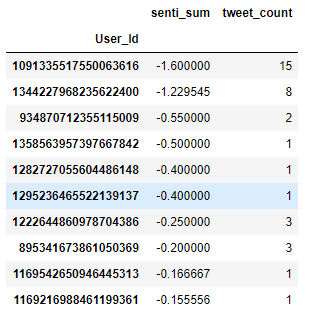
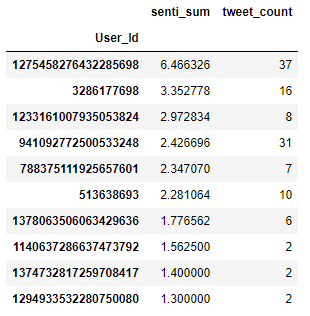
***Dataframe structure for tweet data we collected.***

*Since we required user reaction (red for positive reaction and blue for negative reaction in below results) for our graph to decide whether the sentiment of tweets for that time period is positive or negative. We used sentiment analysis on the stored data using textblob library and added an additional column called sentiment\_score.*

**

***Updated results of dataframe after sentiment analysis***

*We further filtered the data collected as it had similar users, so we decided to aggregate them and their sentiment scores. Finally we filtered our data to provide relevancy by adding a condition of keeping higher tweet count and sentimental score values only.*

**

***Higher Positive Sentiment Values Higher Negative Sentiment Values***

*For the weights on the edges of the graph we used the summation of likes count, replies count, retweets count, and quotes count. We also colored the edges according to sentiment, red for positive and blue for negative.*

**

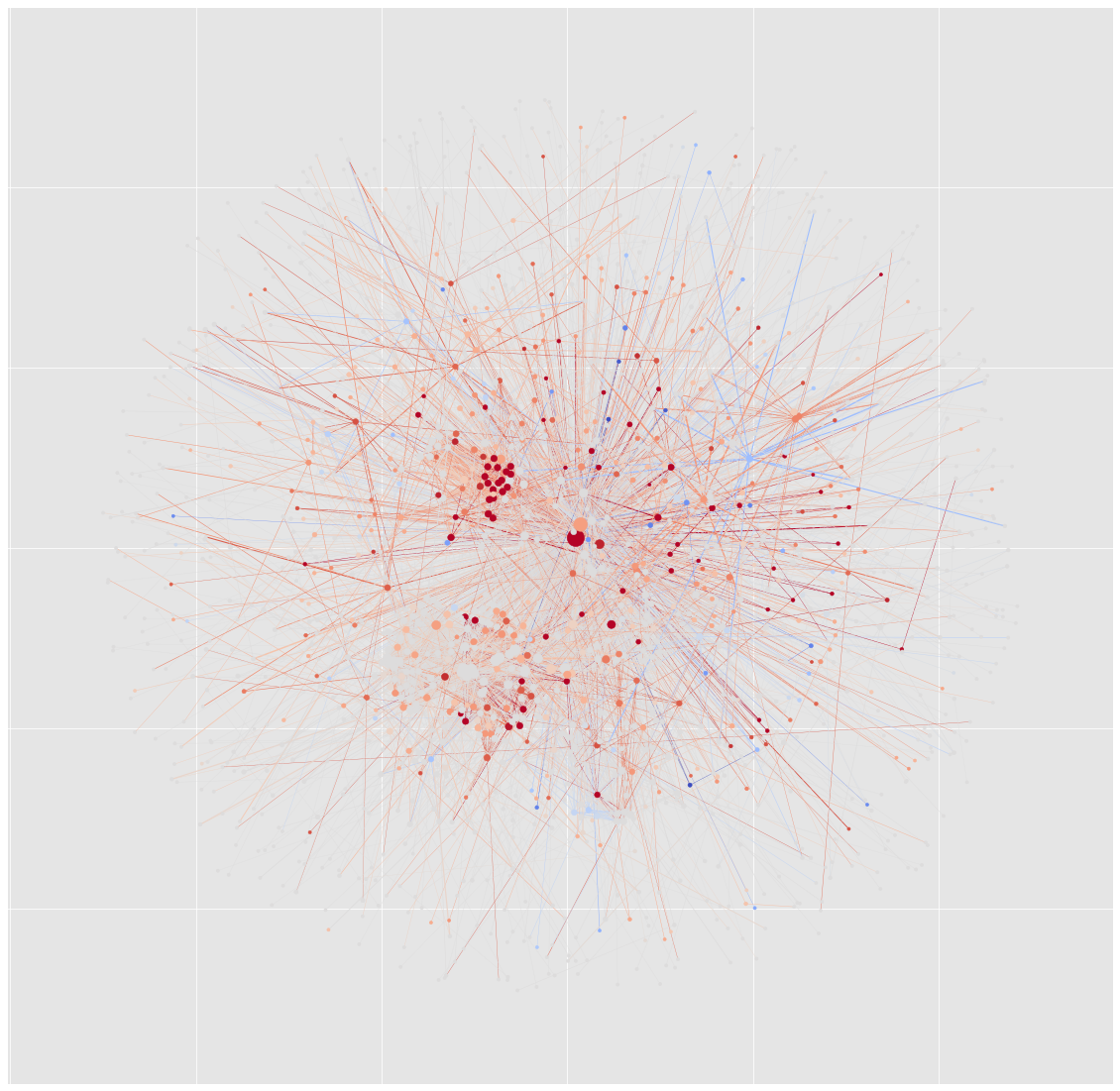
***Code Snippet for weights and edge color for interaction graph***

### 2. Data Visualization

*We have used the NetworkX tool for our graph visualization. We have plotted interaction graphs for two different time periods as mentioned above.*

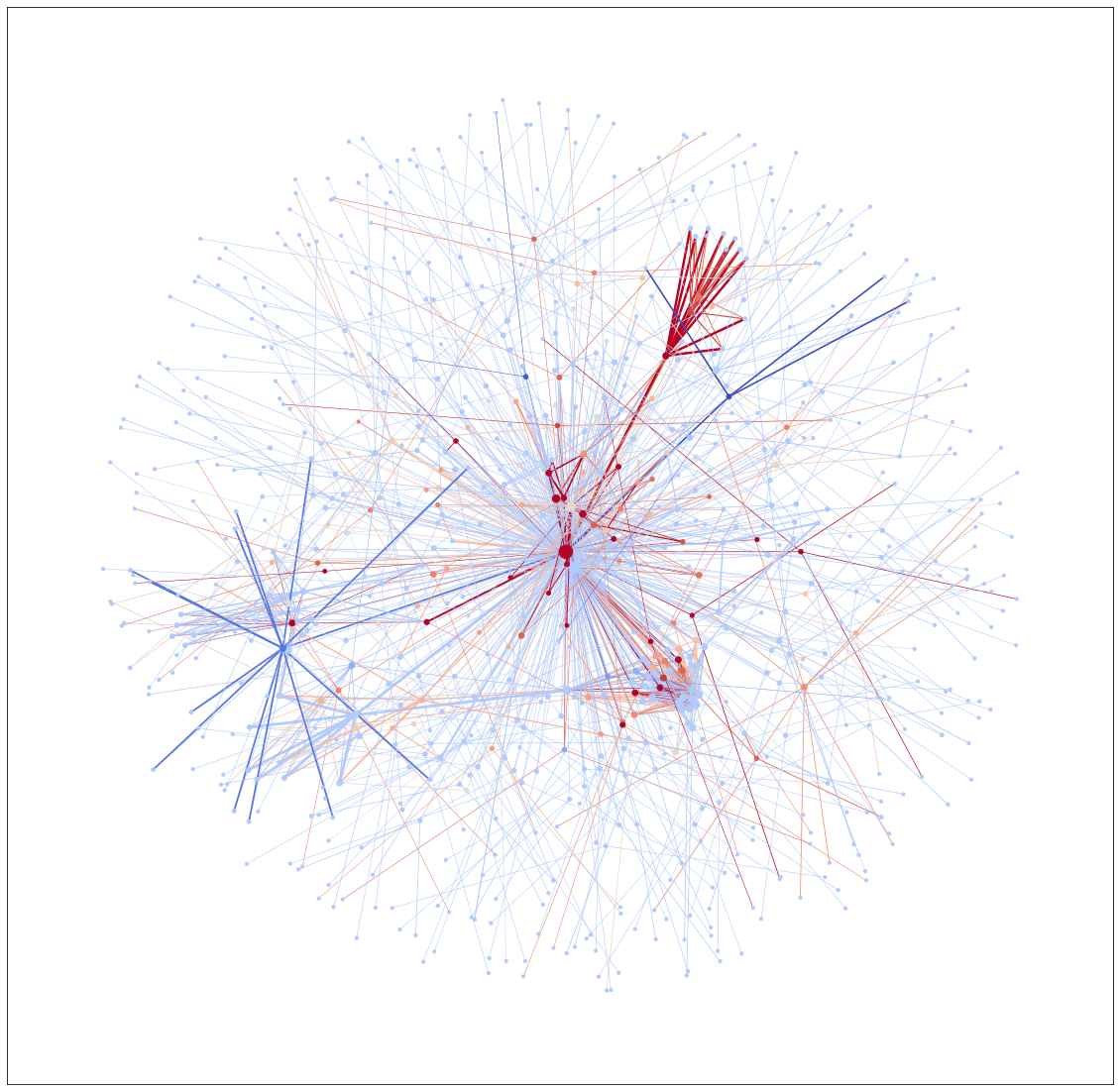
*The nodes are users, and edges interactions between them. Node size represents the number of tweets by the user, and its color the overall sentiment score, with red being positive and blue being negative. Edges are colored according to sentiment as well, and weighted according to the number of likes, replies, retweets and quotes the tweets received. If a user mentions another user in multiple tweets, we group the interactions into one edge.*

*Below are the graphs and their understandings.*

**

***Graph for May’21 phase when the price went from 0.38$ -> 0.139$ drop***

*As we can see from the graph for May’21 during the launch phase and till slight decrement in price had very high positive reaction for ethermon project among the users. The big red nodes which shows the influencers, creators and marketers promoting heavily to raise the price values.*

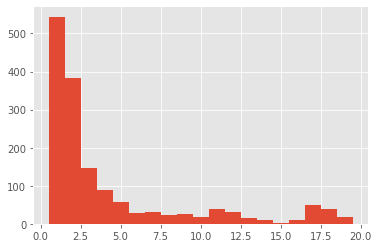
**

***Graph for Jun’21 phase when the price was completely down from 0.38$ -> 0.05$ drop***

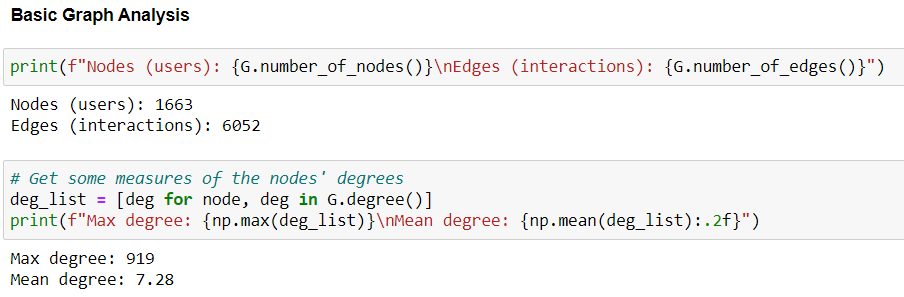
*As we can see during this time period the complete tweets were of negative sentiments, with few positive ones trying to still push the price up to recover their loss. But since the creators, influencers and money makers have already made their profits till mid-May now the project price went way down. Soon, after realizing the motives of the project, people started heavily commenting harsh about it.*

### 3. Network Measures Calculation

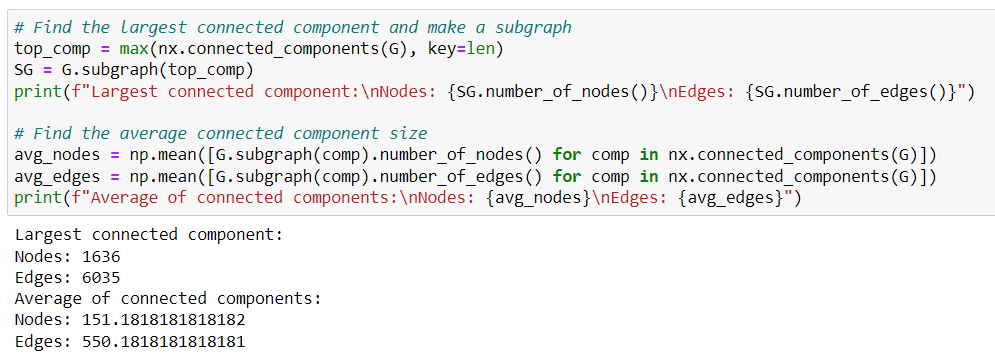
*For the graph of May’21 the network measures are as followed*

**

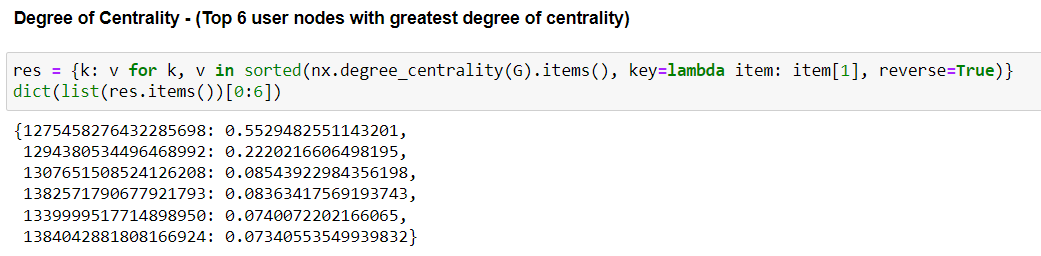
***Degree Distribution as a histogram of Graph for May’21***

******

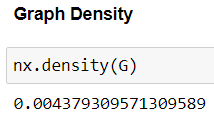
***Node and Edge Count, Max and Mean Degree of Graph for May’21***

******

***Connected Component Measure of Graph for May’21***

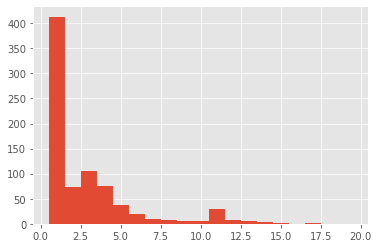
******

***Degree of Centrality for Top 6 user nodes with greatest values for May’21 time period***

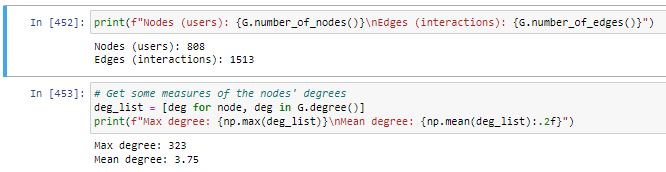
******

***Graph Density for May’21 time period***

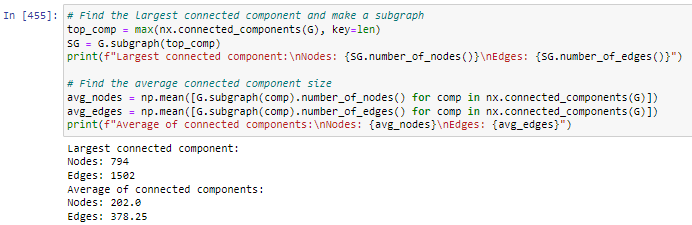
*For the graph of June’21 the network measures are as followed*

**

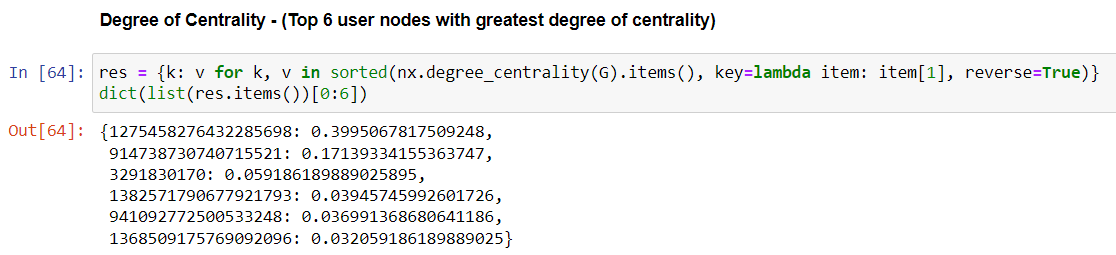
***Degree Distribution as a histogram of Graph for Jun’21***

**

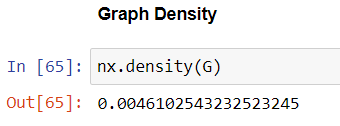
***Node and Edge Count, Max and Mean Degree of Graph for Jun’21***

**

***Connected Component Measure of Graph for Jun’21***

******

***Degree of Centrality for Top 6 user nodes with greatest values for Jun’21 time period***

******

***Graph Density for Jun’21 time period***

### 4. References

1. Python libraries used snscrape, np, re, textblob, pandas, matplotlib, networx
2. Tutorials referred
   1. <https://github.com/ugis22/analysing_twitter/blob/master/Jupyter%20Notebook%20files/Interaction%20Network.ipynb>
3. NFT use case link <https://coinmarketcap.com/currencies/ethermon/>
4. Master thread of similar nft use case https://twitter.com/zachxbt/status/1465464817813176324?t=8aEHr1wgWhf65K91e7DWrw&s=19