NFT Price and Volume

Matthew M Hoang

Ohio Dominican University MBA Program

# What is an NFT?

An NFT or Non-Fungible Token are digital assets that hold value and are usually connected to a blockchain technology. Blockchain technology is usually associated with cryptocurrencies such as Bitcoin and Ethereum. Many NFT’s are built off of the Ethereum blockchain as it allows more advanced technologies to be added to the code than the Bitcoin environment. NFT’s can range from art to digital land you can own and build on. The vastness of NFT’s is what is most confusing. For this study we will focus on digital avatar known as CryptoPunks. CryptoPunks according to their creator larvalabs are “24x24 pixel art images, generated algorithmically. Most are punky-looking guys and girls, but there are a few rarer types mixed in: Apes, Zombies and even the odd Alien.” (Larvalabs)

CryptoPunks required the artist to only create parts of a face and let a computer program put those parts together in different combinations to create the full set of 10,000 characters. CryptoPunks use the Ethereum blockchain to monitor and enforce the transaction of these avatars between people. CryptoPunks have become one of the highest value NFT’s on the market as the average bid for one is nearly 3 ether or $9000. The lowest sale for a CryptoPunk currently is 61 ether or $171800. The largest sale being in February of 2022 sold for around 8000 ether or $23.7 million. These prices have skyrocketed since their creation and continue to grow.

The Ethereum blockchain is a decentralized platform that allows people to connect and verify code called smart contracts. Smart contracts are the backbone of Ethereum transactions and NFTs. Together with the decentralization aspect, it allows total visibility of all transactions. The benefit of Ethereum over Bitcoin is that you can build applications off the Ethereum blockchain. Bitcoin does not allow this feature. The ability to build within this environment is how NFTs came about.

# History of CryptoPunks

CryptoPunks are one of the earliest NFTs on Ethereum and helped shape the standard for building new NFTs. Created by Larva Labs by John Watkinson and Matt Hall, they launched the art project in June 2017. This raised the questions of “What does ownership mean in the digital age? Will people have any interest in paying for the equivalent of a digital certificate of authenticity?” (Levith, 2021) While no two CryptoPunks are exactly alike, they share the core elements that make up each individual so many can be found with the same hair style but might have a piercing. Initially Larva Labs offered 9000 CryptoPunks for free and kept the remaining 1000 for themselves. They continue to auction off the remaining CryptoPunks at renown auction houses such as Sotheby’s or Christie’s. The creation of CryptoPunks has created a wave of enthusiasm for NFT’s and digital art. Many new projects continue to build the standardization of NFT’s and continue to flesh out the smart contracts that allow new artist to create more value in their work.

## CryptoPunks Value connected to popularity.

NFTs have most of its value driven by the popularity of that art. Art has an interesting dynamic in that it cannot create more supply, so its value is heavily dependent on demand. The demand side being as impactful as described, I wanted to test popularity of an NFT and how that might affect the price. To do this I think my preferable variable would have been web scrapping tweets and google search results and modeling those against price. That would be a bit too much work and decided to use volume of trades as the next best. Volume could be depicted differently as CryptoPunks are much more expensive than most commodities and would be affected by its price.

The data set used is from Kaggle. The data set is called CryptoPunks and contains prices, listings, bids, and sales of all CryptoPunk trades. This data was saved as a .jsonl file that was imported through the gui functions of R.

The libraries used are the typical libraries in class with the addition of the rjson library, Splitstackshape, and sqldf. Rjson allows us to manipulate json data files through R but that was not necessary as R seemed to read the json file as a data frame. The Splitstackshape library allowed me to separate strings of characters like the text to column function in Microsoft excel. And sqldf allowed me to manipulate the data frame using SQL terminologies for ease of filtering columns.

Data preparation took the longest as I was using a large and complex data set. I started with splitting column 1 which depicted the type of transaction to separate the types for easy filtering. Next, I used sqldf to use a where SQL statement to create a data frame with only the sold transactions. I then separated dates and the numeric value of ether into separate columns. This allowed me to manipulate the dates in the json file. The json file dates were imported as character file types and needed to be converted to date filetypes. I played around with the best method to go about this and decided on created two date columns. One would be the actual date of the transaction and the other would be the month that transaction took place. By separating these columns out, I could group by the dates and by the months to get the average price of ether CryptoPunks were going for as well as tally up the total volume of trade per month. I then merged these data frames together for my final data set.

**Regression and lagged model.** The regression of my model consisted of price equals volume. This generated a result of volume having a coefficient of .017 and a T value of 31.4 and a P value between 0 and .001. This means our result is statistically significant and volume could be impacting the price of a CryptoPunk by .017 ether or nearly $50 per sale. Next, I wanted to test for homoskedasticity in our data by plotting the squared residuals. Figure 1 shows the data is flat until 2021 when the NFT became much more mainstream. These number of observations though in our data is nearly 19000 and would be safe to say a line of best fit would still be flat even with the outliers in 2021.

Next, intuitively I felt price would be dependent on the previous price of a CryptoPunk and wanted to check for stationary data. Our book *Principles of Econometrics* by R. Carter Hill, William E. Grffiths, and Guay C. Lim explain stationary data as “variables have means and variances that do not change over time and autocorrelations that depend only on how far apart the observations are in time, not a particular point in time.” (p.427) To test for this we can create a correlogram to easily see. Figure 2 shows the correlogram with a rho of .234 is insignificant in two spots. A dicky-fuller test also shows low p values and shows our data to be stationary.

To follow up with my intuition that price is dependent on previous price I plugged our model into a first difference equation. Principles of Econometrics explains first difference stationary as “variable is one that is stationary after differencing once... variables that are only stationary once and not cointegrated, then a suitable regression involving only stationary variables I one that relates changes in y to changes in x, with relevant lags included.” (p. 586) Our equation looks like Price is equal to -1 plus price t-1 plus volume plus residual t-1. This equation gives us low coefficients for both price and volume while having high p values. These results are not proving to be statistically significant. I also tried to duplicate this effort on the average price of ether over a month. My equation looked more like Average price equals -1 plus average price t-1 plus volume t-1 plus residual t-1. This gave a much better result. The coefficient for average price lagged was -.49 with a p value of .0001. The coefficient of Volume was .02 with a p value of .007. The residual squared graph figure 3 shows more heteroskedasticity than our original model.

Our models need more tuning to account for a myriad of different factors that are influencing price. These basic models helps us understand the difference a first difference model and a normal regression. We also introduced lagged variables to help account for some of the other factors playing into price. In the end, prices of NFT’s may be even more volatile than stocks and just as unpredictable. These art pieces don’t have quarterly metrics that they must hand out and performance reviews to watch. I think to improve the model going forward would be web scrapping hits that an NFT is making as well as watching the number of participants in the discord server associated with that specific NFT.

References

Hill, R. C., Griffiths, W. E., & Lim, G. C. (2018). *Principles of econometrics*. Wiley.

LarvaLabs. (n.d.). *Cryptopunks*. Retrieved May 1, 2022, from https://www.larvalabs.com/cryptopunks

Levith, W. (2021). *Cryptopunks: A short history*. Retrieved May 1, 2022, from https://www.withotis.com/mag/cryptopunks-short-history

Figures

Chart

Description automatically generated

Figure 1: Squared Residuals of first regression

Chart

Description automatically generated

Figure 2: Correlogram of first regression

Chart, scatter chart

Description automatically generated

Figure 3: Residual Squared of averaged lagged regression

R Appendix

1. Library set up
2. Data preparation
3. Initial Regression
4. Squared Residuals
5. Correlogram and stationary test
6. Lagged Model
7. Averaged Lagged Model