EXPLORATORY DATA ANALYSIS OF SOLANA NFTs

Submitted By: GROUP 4

1. **Syed Bilal Hussain**
2. **Parnit Kaur**
3. **Shreeya Suryakant Bidkar**
4. **Dheekshith Kadavinahosahalli Malleshaiah**
5. **Damilola Olanrewaju**
6. **Anantha Krishnan Santhosh**

ABSTRACT:

NFTs have been gaining a lot of popularity and momentum over the past 5 years. This analysis focuses on the Solana NFT marketplace that first emerged during the latter part of the year 2017. The project intends to investigate on the various features like the market capital, sales, and their owners. Also examining the failed NFT projects and the numerous categories of the NFTs available. It can evidently be seen that ‘Collectibles’ and ‘Digital Art’ are the most saleable categories of NFTs compared to the rest that only contribute a fraction of the volume. The analysis also points out that more than half of the NFT projects that entered the marketplace failed to accelerate.

INTRODUCTION:

The term NFT has been trending heavily lately, and everyone wants to be a part of something that could be a highly profitable investment for the future. Anyone with a device connected to the internet, and the slightest enthusiasm towards the internet space have either come across or searched for this trending term, especially those who feel upset for missing out on the huge opportunity earlier, when Bitcoin’s price boomed.

Now the question arises, what exactly are NFTs?

NFTs' are non-fungible tokens embedded into a blockchain with unique smart contracts, therefore not interchangeable, while offering an exclusive digital certificate of ownership for the NFT [1]. These tokens are objects like digital art, collectible, audio, video, in-game purchases, etc. and are traded online with cryptocurrency. More importantly, an NFT blockchain enables us to establish the origin of the token, and distinguish between the original object and its duplicates, allowing to commodify and lodge various marketplaces for digital artists.

The minting and selling of NFTs began 5 years after the introduction of Bitcoin. The very first NFT was created by Kevin McCoy in May of 2014. McCoy created Quantum, which raked in $1.4 million at a Sotheby auction in November 2021 [3]. Interest in the NFTs heightened in the year 2021 when the market experienced record sales. Total trading volume increased by a staggering 704% between Q2 and Q3 of 2021. Further on the 12th of March 2021, "Beeple", a digital artist sold his 'The First 5000 Days' Metakovan, Founder of Metapurse, for an astounding $69,346,250". A digital art piece that sold for 69 million US dollars, provided hope and spread the news among the artist community that this is the future.

Why Solana? Why not Etheruem?

Originally NFTs were part of the Ethereum blockchain, but increasingly more blockchains have implemented their own versions and marketplaces of NFTs [4]. One such upcoming NFT marketplace based on the Solana blockchain has been analysed in this project. Although Ethereum is the pioneer of the blockchain market, it had its disadvantages, especially its speed. The Solana blockchain was created to solve the shortcomings of the Ethereum blockchain. Solana was launched during the peak of the 2017 crypto surge. The network was founded by Anatoly Yakovenko. The Solana

team includes some notable tech executives. For example, Greg Fitzgerald, Solana's CTO, worked for Qualcomm, Dropbox, and other well-known platforms [5]. ‘SOL’ is the token for the Solana network. The main objective of the Solana Blockchain is to deliver on scalability without sacrificing decentralization. Solana is one of the fastest blockchains in the crypto market. Its protocols have been benchmarked up to 29,171 transactions per second. Solana's network processes a block every 2.34 seconds in contrast to Ethereum which sits at around 15-transactions per second. Ethereum upholds one block approximately every 13 seconds, Solana being almost 6 times faster.

DATA SOURCE AND MODEL:

DATA COLLECTION:

The data used to conduct an exploratory analysis on the Solana blockchain NFT collections were obtained in the form of a CSV file from Kaggle.com, a subsidiary of Google, which is a crowd-sourced platform created to attract, nurture, train and challenge data scientists from all around the world to solve data science, machine learning and predictive analytics problems [6]. The author employed python using the requests, JSON, and pandas libraries to extract data of the top 600 NFTs present in the Solana Blockchain [7] from coimarketcap.com. The dataset consisted of 592 columns and 17 rows. The structure of the original CSV file containing the top 20 indices is displayed and explained below in figure 1.

1	Index	Name	Volume	Volume_US	Market_Cap	Market_Cap	Sales	Floor_Price	Floor_Price_	Average_Pri	Average_Pri	Owners	Assets	Owner_Ass	Category	Website	Logo
2	0	basis.marke	27256.63	4001818.4	708.14545	103969.92	366	39.5	5799.39	74.471667	10933.93	12	16	75	Collectibles	https://basi	https://
3	1	Skeleton Cr	11381.25	1670995.1	106.656	15659.234	2075	0.67	98.3694	5.4849398	805.29886	103	132	78.03		https://skel	https://
4	2	Flippies	6732.19	988420.14	1663.8632	244288.39	1504	0.8	117.456	4.4761902	657.19424	905	1638	55.25		https://ww	https://
5	3	Meerkat Mil	3734.76	548337.46	90.275	13254.176	395	3.79	556.4478	9.4550886	1388.1961	19	23	82.61		https://mee	https://
6	4	Stylish Stud	3191.15	468524.64	165	24225.3	729	4.3	631.326	4.3774348	642.69498	28	50	56		https://styli	https://
7	5	I'M AIKO	2904.7	426468.05	2530.8771	371583.38	1411	1.2	176.184	2.0586109	302.24525	599	1206	49.67	Art,Collecti	https://aiko	https://
8	6	Crypto Idolz	2315.74	339996.95	1168.042	171491.92	1418	0.6	88.092	1.633103	239.77218	244	529	46.12		https://cryp	https://
9	7	Megalodon	2116.61	310760.68	60.050375	8816.5961	5229	0.04	5.8728	0.4047829	59.430231	946	1309	72.27		https://t.co	https://
10	8	CorruptCatz	1983.92	291279.13	345.79	50768.888	1489	0.1	14.682	1.3323842	195.62064	636	916	69.43		https://corr	https://
11	9	Meta Waifu	1844.59	270822.7	989.49127	145277.11	1156	1.27	186.4614	1.5956661	234.2757	376	643	58.48	3D,PFP,Met	https://met	https://
12	10	Crypto Idolz	1756.29	257858.5	910.91546	133740.61	2906	0.24	35.2368	0.6043668	88.733138	413	847	48.76		https://cryp	https://
13	11	Doge Capita	1630.23	239350.37	5.62	825.1284	597	1.29	189.3978	2.7307035	400.92189	5	6	83.33		https://ww	https://
14	12	Hot Bunnies	1590.89	233574.47	527.85	77498.937	412	1	146.82	3.8613835	566.92832	243	391	62.15	Digital,Meta	https://hot	https://
15	13	BOSS BULLS	1236.33	181517.97	244.68667	35924.896	370	1.49	218.7618	3.3414324	490.58911	99	127	77.95	Art,Digital,P	https://ww	https://
16	14	Solana Dooc	1132.74	166308.89	327.88889	48140.647	1162	1.5	220.23	0.9748193	143.12297	123	175	70.29		solanadood	https://
17	15	Crypto Idolz	990.64	145445.76	962.15677	141263.86	1264	0.25	36.705	0.7837342	115.06785	388	1239	31.32		https://cryp	https://
18	16	Psychedelic	981.33	144078.87	0	0	60	15.5	2275.71	16.3555	2401.3145	12	13	92.31	Digital,Art,N	https://psyc	https://
19	17	169 Pixel Ga	970.46	142482.94	84.8475	12457.31	1195	0.1	14.682	0.8121004	119.23258	285	419	68.02		169pixel.co	https://
20	18	Solchicks	858.32	126018.54	0	0	126	9	1321.38	6.8120635	1000.1472	14	15	93.33		https://solc	https://

Figure 1. structure of the CSV file

The dataset contains the following information on the NFT collections:

1. Index: The index of the file.
2. Name: The names of the NFT collections.
3. Volume: The total volume of sales from the NFT collection in Solana (SOL).
4. Volume_USD: The total volume of sales from the NFT collection in United States Dollar (USD).
5. Market_Cap: The market capitalization, total value of the collection's items in circulation in Solana (SOL).

6. Market_Cap_USD: The market capitalization USD, total value of the collection's items in circulation in United States Dollar (USD).
7. Sales: The number of sales from the NFT collection.
8. Floor_Price: The lowest price at which any of the objects in the NFT collection were sold, in Solana (SOL).
9. Floor_Price_USD: The lowest price at which any of the objects in the NFT collection were sold, in United States Dollar (USD).
10. Average_Price: The average price of an NFT in the collection in Solana (SOL).
11. Average_Price_USD: The average price of an NFT in the collection in United States Dollar (USD).
12. Owners: The total number of owners of NFT's in the collection.
13. Assets: The number of objects in the collection.
14. Owner_Asset_Ratio: The ownership percentage of all the items in the collection.
15. Category: The category of the NFT collection.
16. Website: The associated websites of the NFT collection.
17. Logo: The associated images of the NFT collection.

METHOD:

DATA PRE-PROCESSING:

To prepare the obtained data for analysis, the data had to undergo some pre-processing. It was done completely using Mathematica, the data file was imported into the notebook using the ‘Import’ function and its dimensions were first checked. A function was introduced to import the images of the logos of different NFT collections into the dataset. Initially, the dataset was explored for duplicates, and then checked for missing data in the columns, and the total number of missing values in each column was obtained. The column containing the links to the collection websites and categories were removed as they were irrelevant to our analysis and a lot of data was missing, respectively. The rows containing the missing data were then deleted. Additional data on the categories of each NFT collection were manually imported into the dataset from solsea.com (Solana NFT marketplace).

DATA ANALYSIS:

After the data pre-processing, the next step is to analyse the data. The initial step was to discover the correlation between the entities in the columns, and therefore a correlation analysis was performed to understand further the association between the different column entities, that is the features of various NFTs. This was achieved by utilising the correlation function inbuilt in Mathematica applied to the dataset.

The next goal was to find the top-performing NFT collections in the dataset, which were discovered by employing the ‘SortBy’ function available in Mathematica which enabled the sorting of the objects concerning their market cap selected as their

performance attribute. Further analysis was performed on the floor prices and floor sales of individual NFT collections by plotting bar charts, done using the 'BarChart' function. Bar charts are frequently used in the analysis as they allow the reader to identify specific patterns and trends efficiently when compared to a table containing data.

A Bar chart and whisker chart was employed to discover the trends of the top-selling NFT collections from the dataset. The whisker chart was displayed using the 'BoxWhiskerChart' function available in Mathematica and is a convenient method to display key values evaluated from the sales column including the average, median, and upper and lower sales prices of the top 20 selling NFTs. Further analysis was done on the number of assets that belonged to each NFT collection, and the number of owners of each object type using the 'ListPlot' function, 2 graphs were plotted to find the relation between the number of owners, the number of assets that belong to each NFT collection, and the sales.

RESULTS:

(i) Correlation Analysis

Correlation analysis was employed to statistically evaluate and study the relationship between the different columns. A correlation between two columns singifies that in the occurrence of a systematic switch in one variable, it also affects the other, either positively or negatively, i.e., the variables alter together. A correlation coefficient is an idea to give value to a relationship. Correlation coefficient values range from -1 to 1. In case it becomes "0", it means there is no relationship between the variables at all. A correlation of -1 is known to be a perfect negative (increasing one decreses the other) and 1 is known to be a perfect positive (increasing one linearly increases the other).

	Volume USD	Market Cap USD	Sales	Floor Price USD	Avg Price USD	Owner	Assets	Owner Asset Ratio
Volume USD	1.0	0.356271	0.238439	0.0696146	0.605738	0.144694	0.126606	0.0667043
Market Cap USD	0.356271	1.0	0.400295	−0.00954017	0.110256	0.414119	0.535009	−0.0943579
Sales	0.238439	0.400295	1.0	−0.0346645	−0.0347812	0.810149	0.751806	0.0285603
Floor Price USD	0.0696146	−0.00954017	−0.0346645	1.0	0.207361	−0.0446786	−0.0525707	0.092825
Avg Price USD	0.605738	0.110256	−0.0347812	0.207361	1.0	−0.0691902	−0.0845218	0.0864633
Owner	0.144694	0.414119	0.810149	−0.0446786	−0.0691902	1.0	0.908035	0.045991
Assets	0.126606	0.535009	0.751806	−0.0525707	−0.0845218	0.908035	1.0	−0.14619
Owner Asset Ratio	0.0667043	−0.0943579	0.0285603	0.092825	0.0864633	0.045991	−0.14619	1.0

Figure 2. Correlation table

The correlation analysis was performed between Volume, Market Cap, Sales, Floor Price, Average Price, Owners, Assets and Owner Asset Ratio.

Following had a strong positive correlation:

1. Volume and Average Price → 0.60
2. Market Cap and Assets → 0.53
3. Sales and Owner → 0.81
4. Sales and Assets → 0.75
5. Owner and Assets → 0.91

From the above, we can state that the five variable-pairs have a strong positive correlation (> 0.5). This simply means that an increase in one of them will lead to an increase in the other. In this scenario, the highest correlation dependency would be between the columns ‘Owner’ and ‘Assets’ at 0.91, followed by ‘Sales’ and ‘Owner’ at 0.81. The least correlation is between ‘Market Cap’ and ‘Asset’ at 0.53 which is still significantly positive.

(ii) Top 20 NFT Collections based on Market Cap

We then proceeded to analyse the 20 NFT Collections with the highest market cap.

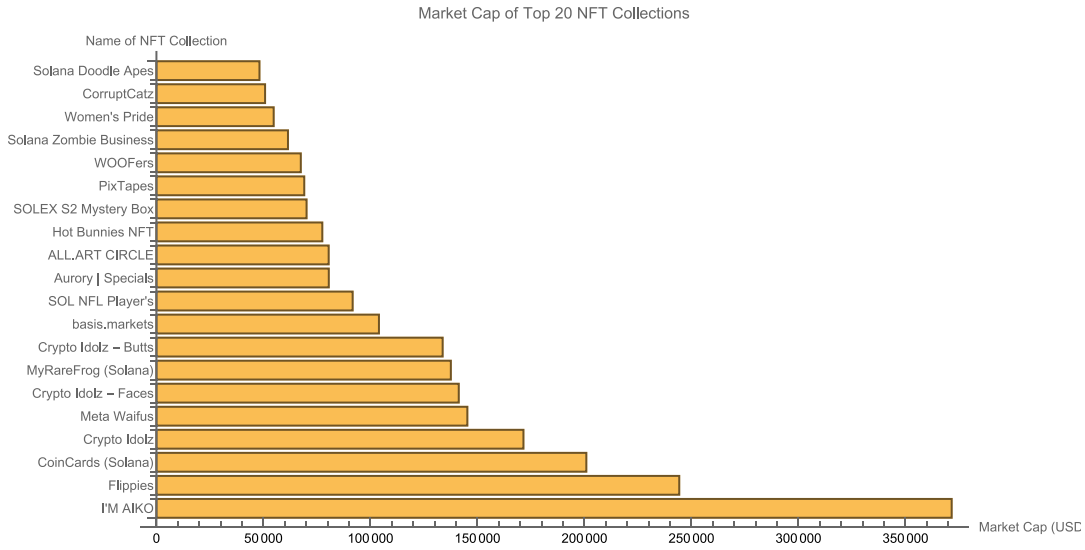


figure 3

A bar chart was plotted (figure 3) to evaluate the top 20 NFT collections with respect to the market cap.

Upon calculating, it was discovered that about the 3/5th of the Market Cap was held by just 20 NFT collections out of almost 600 total NFT collections.

Among the top 5 NFT collections with the highest market cap, *'I'M AIKO', 'Crypto Idolz'* and *'Meta Waifus'* fall in the NSFW category. Collectors are spending around $200-$300 to buy them and their Market Cap is around $150k-$370k, which is mind-boggling.

We then proceeded to analyse their average prices (figure4) and the following was discovered:

The average price for NFT Collection 'basis.markets' was astonishingly higher than any other NFT Collection, at 10933.9 $. This was followed by 'Flippies', and 'Hot Bunnies'.

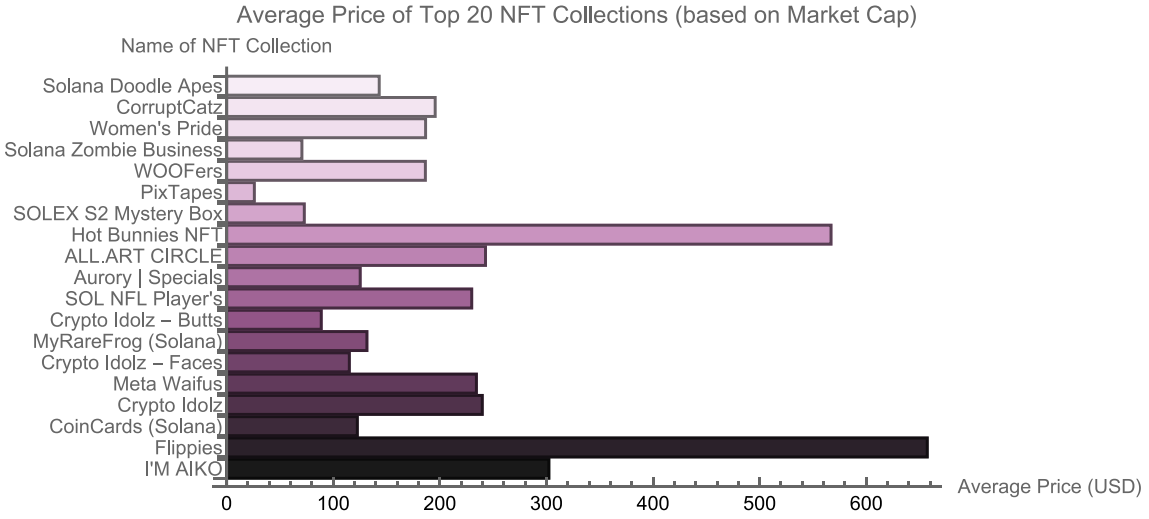


Figure 4

The bar graph was plotted using the 'Average Price' of the collections instead of the 'Floor Price', as the Floor price is decided by the individual who owns an NFT and it depends on a lot of factors like market demand, buyer's emotions, investment potential, creator's popularity and asset's uniqueness, to name a few. Due to this, it varies a lot. So we believe 'Average Price' would give a better monetary value understanding of the NFT collection.

Excluding basis.markets, the costliest is Flippies [$657.2], cheapest is PixTapes [$26] and the average is around $207.

(iii) Top 20 NFT Collections based on Sales

Next, we analysed the top 20 NFT Collections with the highest sales.

It was observed that 42.17% of the overall sales were made by just these 20 NFT collections.

A whisker chart (figure 5) of the ‘Average Price’ of these 20 NFT collections was plotted to get a deeper understanding of its distribution.

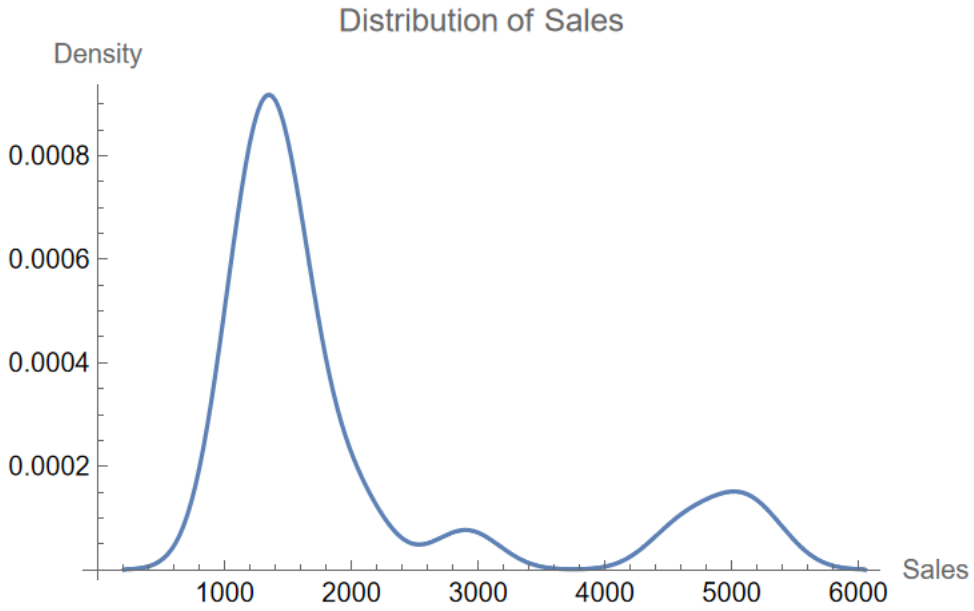
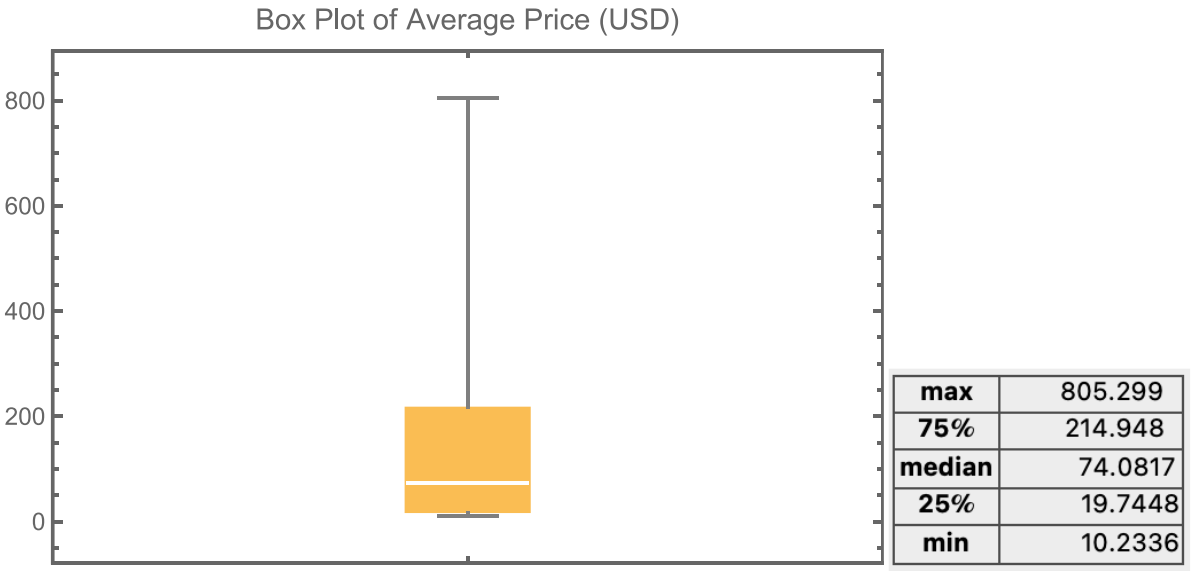


Figure 5



max	805.299
75%	214.948
median	74.0817
25%	19.7448
min	10.2336

Figure 6

Here it was observed that NFT collectors are mostly spending between $10 and $805, and on average $150 on NFTs in the top 20s.

The sales of 10 of the NFTs from the top 20 are found to be between $10 to $74, while the other half falls under the range of $74 to $800.

The majority (75%) of the sales were from around $10 to $215. Whereas, the remaining 25% of the top 20 NFT collections had sales ranging between $215 and $800.

(iv) Top 20 NFT Collections based on Number of Owners

After analysing Top 20 NFTs based on their Market Capital and Sales, analysis on top 20 NFT collection based on the total numbers of owners (figure 7) were performed and the conclusions were drawn from it. Firstly, with the help of the histogram, the distribution of owners among the top 20 NFTs were examined and then, using 'Total' function and basic statistics, the percentage of the total number of owners under top 20s were found, which was about 12%.

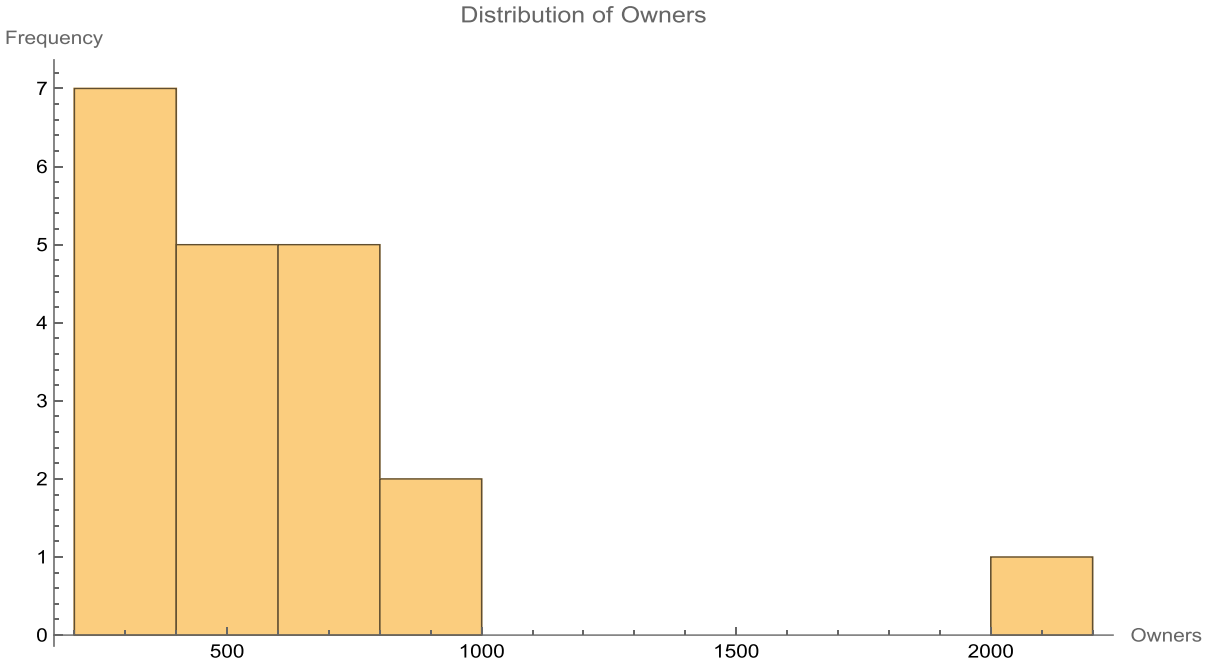


Figure 7

So, only around 12% of the total owners, that is 128,267 have bought the top 20 NFTs. After analysing the data further, it was noted that the 'Solameleon' NFT has the highest number of owners, which constitutes approximately 2% of the total owners.

Further, 'Solameleon' NFT was observed to have the second highest number of sales, according to our previous findings.So together with the above two unique factors of 'Solameleon', gives it the title of being 'The Most Popular NFT'.

Additionally, as previously examined, the number of the owners seemed to be highly correlated with the two features, that are 'Assets' and 'Sales'.

The correlation of the Owners with Assets was found to be '0.908035', and among the Owners and Sales was found to be '0.810149'. Further to support our findings, two scatter plots were drawn for each of the correlations observed, to understand it better.

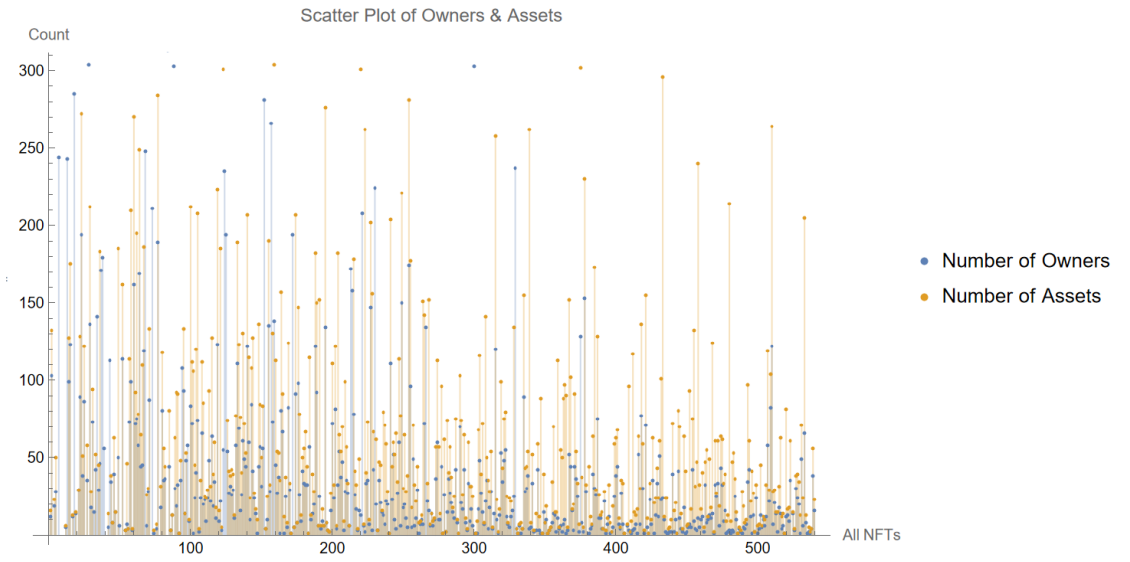


Figure 8

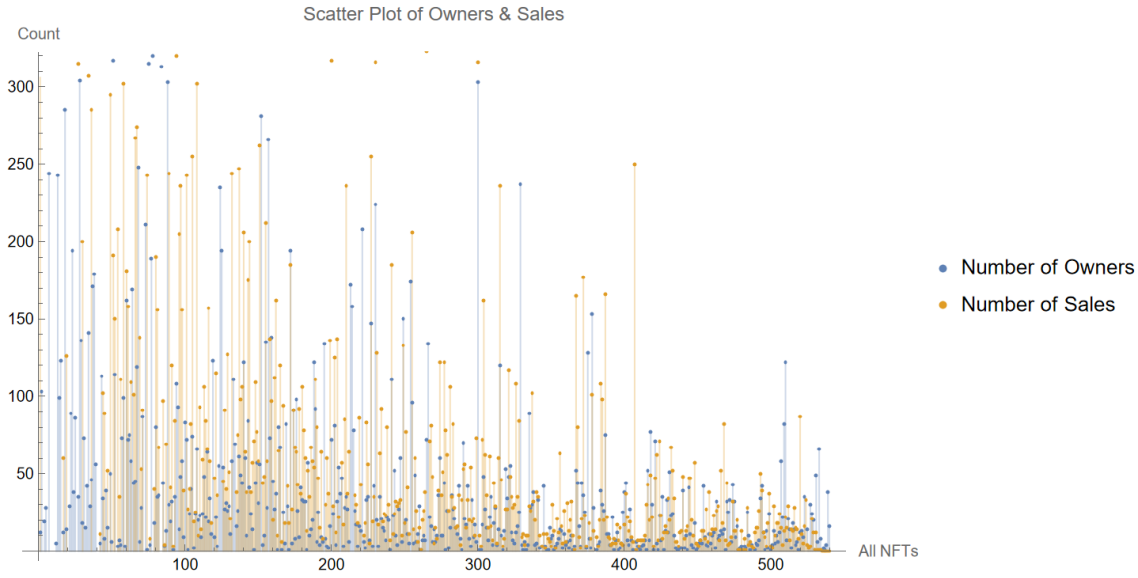


Figure 9

(v) Categories

After analysing owners, a quick analysis on the categories of the top 100 NFT Collections were performed. Since the data already given was not good enough (more than half of it was missing), we individually searched the top 100 NFT Collections (from our dataset) on solsea.io (NFT market that sells Solana based NFTs) and divided them into the following categories:

- Collectible

- Digital Art

- PFP (Profile Picture)
- 3D
- Music
- P2E (Play 2 Earn)
- Metaverse

We then created an excel file using these and used it to analyse the categories.

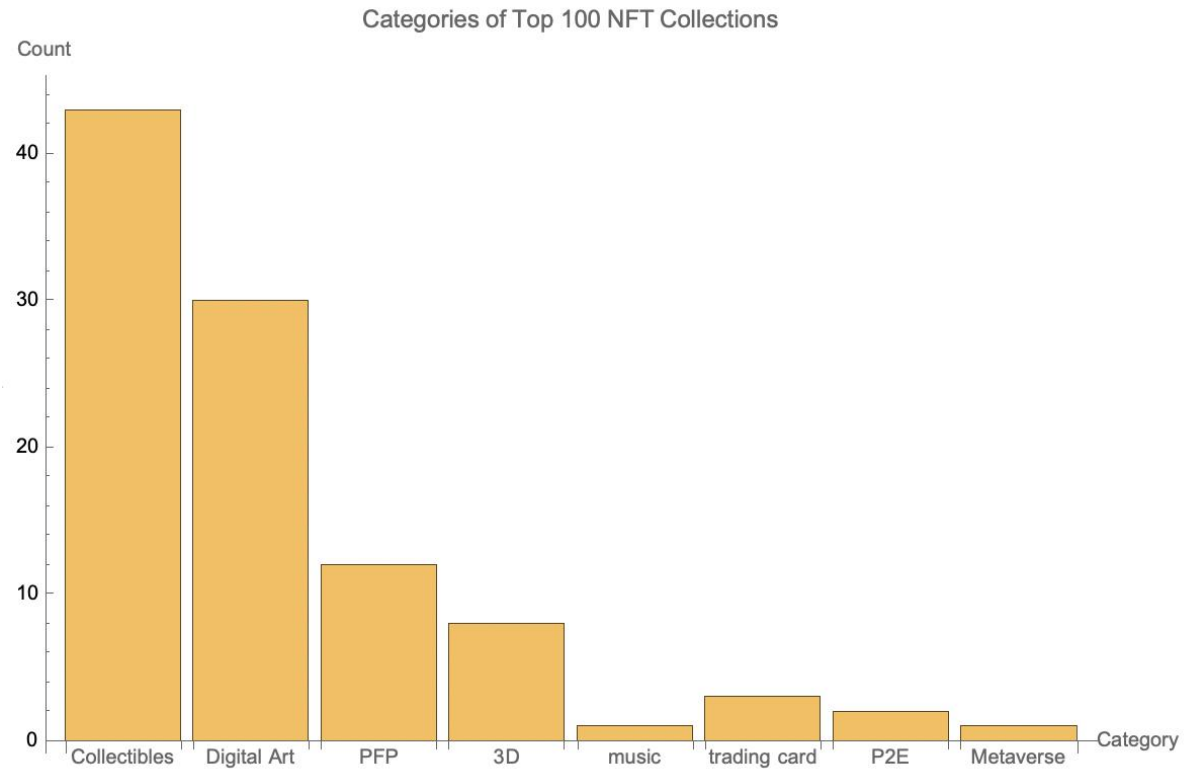


Figure 10

Turns out majority of them were either ‘Collectibles’ or ‘Digital Art’. The rest of them are possibly new entries into the NFT world, and might see a surge in the upcoming years. Music and Metaverse objects appear to be the least favoured ones.

When it comes to the world of NFTs, the answer to what could be an NFT isn’t defined. An NFT could be anything – a collectible item, an artwork, PFP (profile picture), 3D artwork, metaverse, tickets, music, P2E (play to earn), gaming, big sports moments, virtual fashion assets, memes, and even domain names.

Out of these major categories, 8 were observed to be in the top 100 best performing NFTs based on the market capital.

(vi) Failed Projects

For every NFT there is a Baby Baller, one of the thousands of NFT projects that have faltered not long after celebrated debuts [8]. On average, analytics have shown that one in three NFT collections have essentially expired, with little or no trading activity. Another third is trading below the amount it cost issuers to mint the tokens. As industry participants say the decline of NFT is more a sign of frenzied demand cooling rather than a bubble bursting.[9]

We first selected the NFT Collections with 0 ‘Market Cap’ and then sorted it by ‘Sales’ to get the worst performing of the bunch, in the top.

We noticed 9 NFT Collections with 0 Market Cap have no Sales, but they did have owners. Possibly, the owners are the creators themselves.

We calculated the mean of Sales to identify the average number of sales reported of these failed NFTs. It was 44.

We then divided them into two categories:

- Complete Failures: 0 Market Cap and 0 Sales
- Failure: 0 Market Cap but with few Sales

The NFTs which have completely failed are 9, and their percentage is 1.52%.

The collections which have failed after making a few sales are 311 and the percentage is 52.53%.

This shows that over half of the NFT projects had failed!!

The value of NFTs can also plummet if creators don't get a marketing plan together or if the collection is supplanted by newer and shinier projects. Many NFTs that are simply digital images are particularly vulnerable to price drops.

(vii) Exclusive NFTs?

The owner Asset Ratio is the ownership percentage of all items in the NFT collection. A high owner asset ratio indicates that the NFT Collections assets' ownership is diversified (more owners). A low owner asset ratio means that the assets were owned by a very small number of people.

This led us to think: Why do some NFT Collections have a very small number of owners? Are they exclusive NFTs, i.e., collected by a small group of people with very niche interests? Or are they just failed projects? Upon checking the dataset, we noticed many had 0 Market Cap. These are the ones that failed. We certainly don't need these to know more about the possibly 'exclusive' NFT collections. So, we removed them and chose the top 30 rows. We then visualised it with ‘List Line Plot’,

showing the floor price (USD) of these NFT Collections having a low owner asset ratio and non-zero market cap.

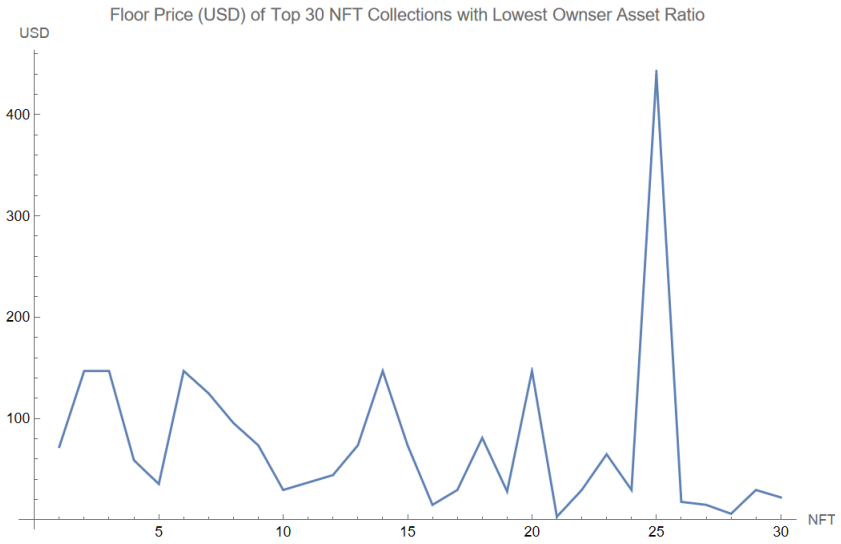


Figure 11

From the plot (figure 8), it appears the floor price ranges between $50 and $150 for most of them. This increased our confidence in our hypothesis that these must be exclusive NFTs that some people buy, even though they aren't popular or among the top-performing. They do not buy it for profit, they buy it because they have a genuine interest in these NFTs. Heart of a collector, we can say. Upon further analysis though, things change. Floor prices can vary a lot, so the average price gives a good sense of the NFT Collection's price. So, we then plotted the average price too.

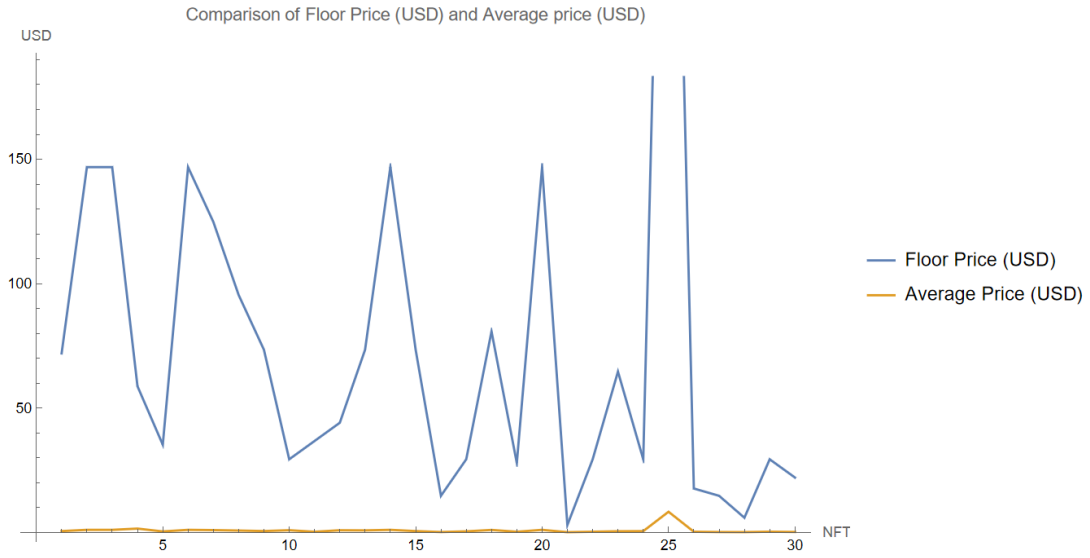


Figure 12

Almost all appear to be less than $1. This gives a clear answer and disproves our hypothesis. Collectors did not buy these NFT Collections due to personal interest. They must have been tracking these collections, observed the price drop from $50-$150 to less than a dollar, and grabbed the opportunity to just buy it. Probably thought,

a dollar is not much, so it is not a big deal if it drowns. But if the price goes back up, they could sell and make a profit. We're willing to bet the majority of the people who bought these, did not investigate the NFT collection. They simply saw the price drop, dropped in to utilize the opportunity, and bought it just in case it the prices increase. It was in the end just a gamble on accumulating NFTs.

CONCLUSIONS:

There are a number of factors that controls the value of NFTs, such as market demand, buyer emotion, investment potential, creator's popularity, and asset's uniqueness. Due to these factors, it is proving more and more difficult to predict the demand and prices of NFT objects on a day-to-day basis. It can be observed that most sales transactions that take place on the Solana blockchain, the Solana NFT marketplace by the investors are between $10 and $215.Of all the categories of NFTs considered, it can be clearly expressed that ‘Collectibles’ and ‘Digital Art’ (both constituting more than 75 percent of the total) are the most appreciated forms of NFT objects in the market when it comes to both market capitalisation as well as sales. On evaluating the trends, we also observed that, out of total NFT collection involved in the analysis 52.53 percent failed. When it comes to any blockchain that hosts a NFT marketplace like Solana and Ethereum the market trends seem to be similar, but we are yet to discover a relationship between the token prices of the blockchain and the prices of the NFTs due to restrictions of data scarcity and research into the NFT market. Whether NFTs are right investments for the future, and whether the value of NFTs at present is because of the immutability of the blockchain still remains a dilemma.

BIBLIOGRAPHY:

1. 1.Team, N. Non-fungible tokens quarterly report Q1 2021. (2021). https://nonfungible.com/subscribe/nf-report-q1-2021. (Accessed 4 May 2021). (**Nonfungible Corporation**).
2. Evans, T. M. Cryptokitties, cryptography, and copyright. *AIPLA QJ* **47**, 219–247 (2019).
3. Lounge, T. W. Choosing the right blockchain for your NFT. (2020). https://medium.com/phantasticphantasma/choosing-the right-blockchain-for-your-nf-d1df2bebae91. (Accessed 4 May 2021)
4. https://fortunly.com/statistics/nft-statistics/
5. https://www.securities.io/solana-sol-vs-ethereum-eth-everything-you-need-to-know
6. https://www.kaggle.com/getting-started/44916
7. https://www.kaggle.com/datasets/nenamalikah/nft-collections-by-sales-volume.

8. 8. https://nftexplained.info/guide-to-understanding-nft-metrics-what-is-floor-price/
9. non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations
10. Mapping the NFT revolution: market trends, trade networks, and visual features Matthieu Nadini1,2, Laura Alessandretti 3, Flavio DiGiacinto1,4, Mauro Martino5, Luca MariaAiello6 & Andrea Baronchelli