IT ACADEMY

CryptoPunks Market Analysis

Marta L. Roldan
Pau Gómez
Pol Sans

Abstract

The CryptoPunks market has recently attracted a lot of attention due to the massive amount of money it can move. Taking a closer look at the data reveals a very uneven market, concentrated in a few owners and where the clear winners are the initial investors. Many have deposited millions of dollars in this market, although its sustainability is in doubt: profits depend on a constant increase in prices. Regarding the internal causes that determine prices, no determining characteristics have been found to suggest a highly valued Punk, which indicates a random market without clear rules. Only some types (alien and zombie) and highly rare Punks see extreme purchase values, but a causal link cannot be determined. The market is also very sensitive to mediatic purchases, so price stability and liquidity are not guaranteed.

Index

1 - What are CryptoPunks ?

CryptoPunks are a Non-Fungible Token (NFT) collection made up of 10,000 digital illustrations, which can be individually bought and sold purely through digital interactions. In recent months they have experienced a significant increase in prices, reaching millions of dollars per transaction.

They are part of a larger asset class: NFTs, which are based on blockchain. Thanks to this technology, the owner of a crypto asset is easily and unambiguously identifiable by all users of the network, which adds a scarcity factor to digital files which are otherwise infinite and free to reproduce. This scarcity is what gives them their exchange value: each CryptoPunk can only have a single owner at a time, known to the whole world. The property certificates are registered in the blockchain, which is in practice impossible to hack since the code in question is stored by thousands of computers around the world. Specifically, this collection is stored on the Ethereum blockchain, which means that all purchase and sale transactions are denominated in this cryptocurrency.

The price of CryptoPunks and of the NFT and crypto space in general has exploded with particular virulence during the year 2021, meaning the interest in this type of goods is part of a larger trend.

What differentiates CryptoPunks from other NFTs? They were basically the first to make digital art part of traditional collectibles. But really, what makes people buy unconventional pixel-based digital art for the price of a mansion? Another factor that increases its perceived value is the collector's item status given by having been a pioneering digital item in the crypto space.

1.1 History of the CryptoPunks

CryptoPunks are an excellent example of the current boom in the NFT market. The path from niche forums to older auction houses has been incredibly fast. The reputable art auction house Christie's recognized the trend in advance, launching one successful NFT sale after another. Today, physical and virtual works of art in NFT format sell for millions of dollars.

This project was started by Matt Hall and John Watkinson, founders of New York-based software company Larva Labs, in mid-2017 when they created 10,000 images of people in 24x24 pixels that they distributed to members of the crypto community free of charge. Half a year later, the cost soared to several thousand dollars, a value that today has skyrocketed. CryptoPunks are some of the first NFTs, giving them extra added value due to their scarcity and significance, which differentiates them from later concepts such as CryptoKittes.

This collection of NFTs has had several peaks in popularity and price recently, sometimes caused by the influence of purchases made by relevant personalities, such as American athletes, the multinational Visa or businessman Gary Vaynerchuk.

2 - Market Characteristics

2.1 Punk Traits

The CryptoPunks are 10,000 in total, each one with a number assigned between #0 and #9999. As they are unique numbers and it is impossible to create more units, the offer of these NFTs is limited. Each CryptoPunk consists of an 8-bit illustration of a “punk” with their individual traits that differentiate them from the rest.

Each CryptoPunk has a skin and a type, determining its general category. There are only 5 types and 7 skins. In addition, they optionally have a maximum of 9 accessories out of a total of 86. A minority of these accessories are hidden accessories, that is, not visible but registered as present.

2.2 Transaction Characteristics

CryptoPunks can change ownership through different types of transactions. Except Transfer and Claim, the rest of the operations have an assigned value in terms of the Ether cryptocurrency. Thanks to blockchain technology, all transactions are publicly accessible and are recorded on the Ethereum blockchain.

The first operation of each CryptoPunk is Claim, when its first owner was assigned one of the 10,000 for free, mostly in 2017. Apart from Claim, the other action that can be done without an explicit cost is Transfer, which is not priced in Ethereum; that is, the owners are able to freely give away or transfer CryptoPunks.

The rest of the monetary transactions can be Offer or Bid, which can end with Sold and result in a transfer of ownership if they come to fruition. An Offer is a public sale price that the owner publishes and that any user can accept at any time. When the buyer provides such an amount, the CryptoPunk is automatically transferred. In contrast, a Bid represents an auction in which the seller only indicates the initial value of the auction. Offers and Bids can be withdrawn by the seller at any time.

3 - Project Structure

3.1 Goal

The goal of this project is to find the reasons behind price differences among the different CryptoPunks, in order to obtain a guide for future investment. With that objective in mind, the latest sale price of each Punk will be analyzed in an attempt to study which variables best predict that quantity.

3.2 Method

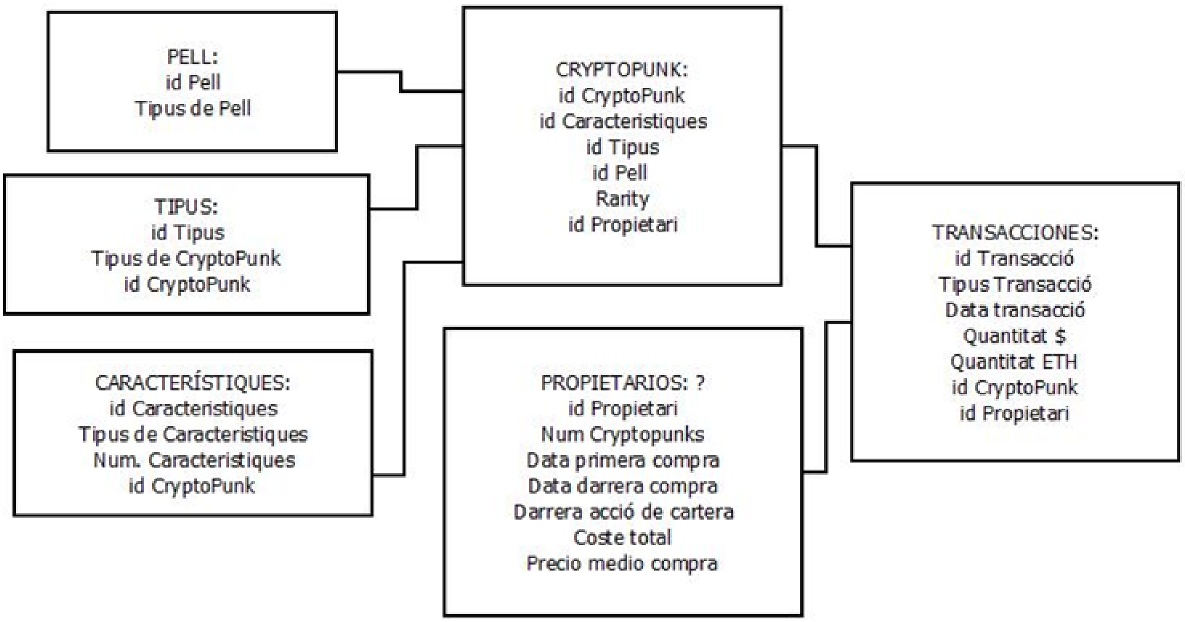
To obtain data, web scraping of the LarvaLabs and DefyPunk pages has been carried out using the Python Selenium library, since those websites provide reliable and updated information on each CryptoPunk, its transactions and owners.

Subsequently, a descriptive analysis of the data obtained will be carried out in order to obtain relevant information on the functioning of the market.

Next, unsupervised Machine Learning methods will be used in order to obtain possible patterns hidden from the human eye and of interest to achieve the objective. With these results in mind, supervised methods will be performed in order to clarify the relationship, if any, between CryptoPunks, their prices, and their characteristics.

3.3 Database Design

To facilitate management and processing, a relational database model has been created with three main tables: CryptoPunks, Owners, and Transactions. The CrytpoPunks table comprises the 10,000 units with their respective characteristics plus rarity index, owner and type. The Owners table mainly shows the financial records of each user and the CryptoPunks owned, and the Transactions table reports on all relevant data regarding property exchange operations. A unique key will be used to identify a CryptoPunk so that information can be easily transferred among tables.



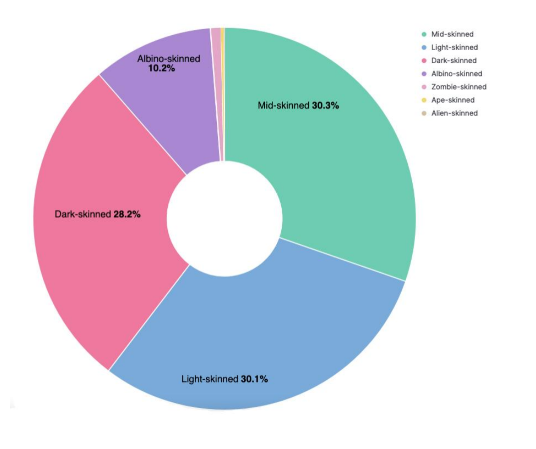
4 - Descriptive Analysis

4.1 CryptoPunks Analysis

4.1.1 Types and skins

Skins

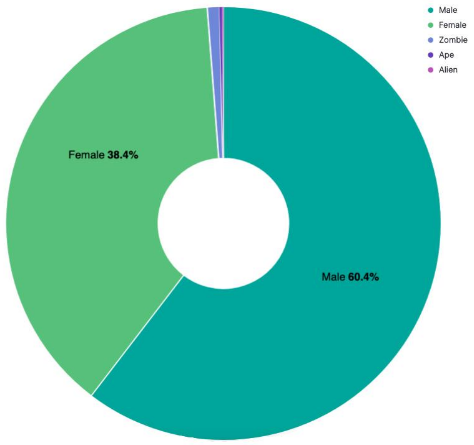
Figure 1 shows a distribution of CryptoPunks by skin and average value. It can be appreciated that the types of skins Ape, Alien and Zombie are those with a much higher average price yet represent a very low percentage with respect to the total of CryptoPunks.



Skin	Count	Percentage	Average price
Albino-skinned	1018	10.2%	112,862 USD
Alien-skinned	9	0.1%	2,274,354 USD
Ape-skinned	24	0.2%	724,404 USD
Dark-skinned	2824	28.2%	111,569 USD
Light-skinned	3006	30.1%	112.286 USD
Mid-skinned	3031	30.3%	102.283 USD
Zombie-skinned	88	0.9%	738.477 USD

Figure 1: distribution of CryptoPunks by skin and average price

Types



Type	Count	Percent	price average
Alien	**9**	**0.1%**	**2,274,354 USD**
Ape	**24**	**0.2%**	**724.404 USD**
Female	3840	38.4%	110,278 USD
Male	6039	60.4%	108,352 USD
Zombie	**88**	**0.9%**	**738,477 USD**

Figure 2: CryptoPunks distribution in terms of type and average price

Figure 2 shows that the Alien, Ape and Zombie Types have a higher average price than the Male and Female types. That is to say, in percentage, 1.2% of the total of CryptoPunks (marked in red: Ape, Alien and Zombie) distort the total price. This is an important point when approaching the present study since a small minority of cases can significantly distort potential results.

In Figure 3, where the relationship between rarity and price is studied, it can be seen that the most extreme values can be found in the highest rarity indices, suggesting a potential indicator of value.

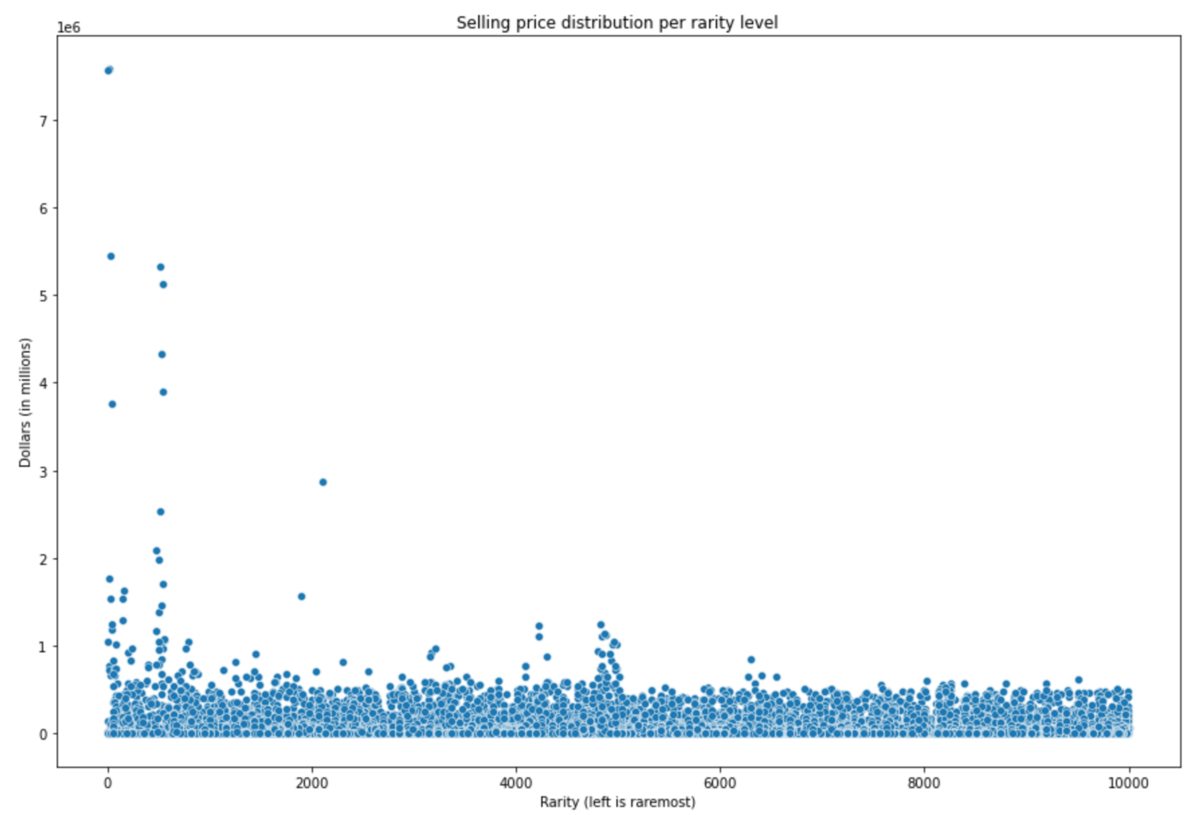


Figure 3: sale price according to rarity index (rarest to the left)

4.1.2 Characteristics

Of the 86 accessories that each CryptoPunk can have, this is the percentage distribution of the number of characteristics per item. Therefore, the largest group of Punks is the one with 3, 2 or 4 characteristics. The rarest are the extreme values, that is, many or few attributes (9, 8, 0…), which can be more likely to see higher prices.

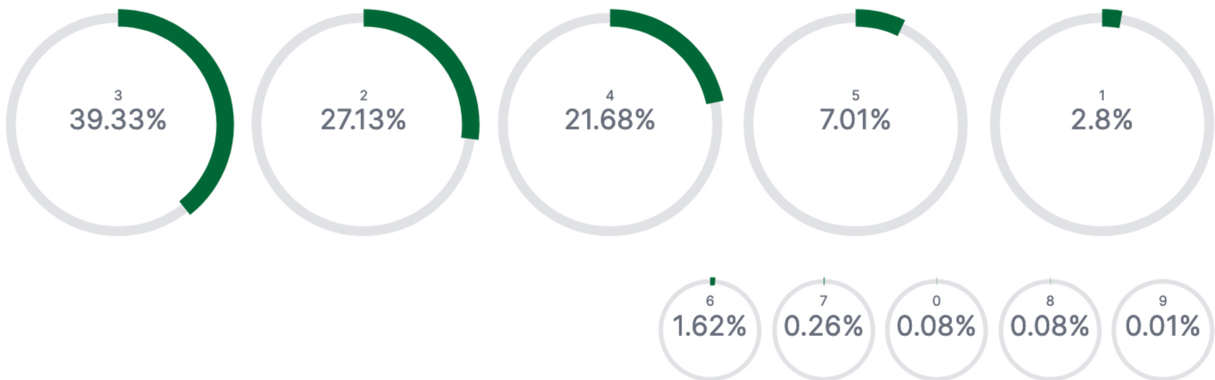


Figure 4: Distribution of CryptoPunks according to number of characteristics

Figure 5 is a price analysis according to the number of characteristics, which reveals higher prices for Punks with 0, 1 and 7 attributes. However, these differences are not as extreme as those appreciated in terms of the Types and Skins.

Regarding the distribution of the characteristics that each Cryptopunk has, it stands out that most elements are usually not too frequent. Those that appear on more occasions are the attributes Earring, Mohawk or Cap. Especially rare are hidden attributes, the most common of which is Hidden Earring.

Number of characteristics	Average price in USD	Distribution of 'shown' and 'hidden' characteristics
0	170,203	
1	202,431	
2	121,560	
3	117,458	
4	107,980	
5	108,806	
6	121,908	
7	233,820	
8	144,026	
9	758	

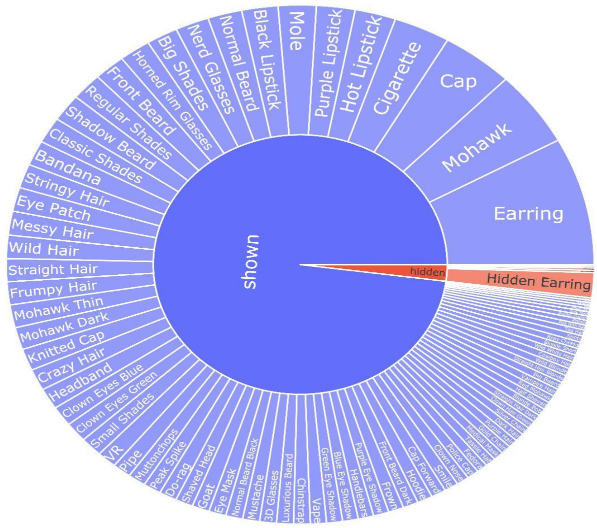
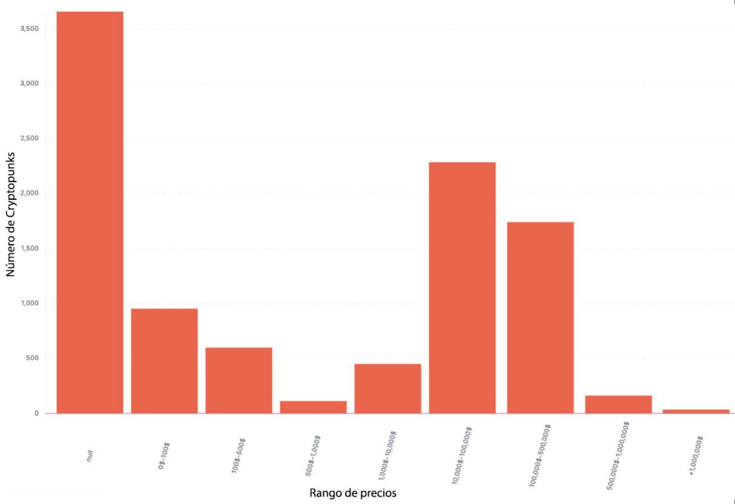


Figure 5: average price in terms of number of Traits and Traits distribution

4.1.3 Price

According to Figure 6, there is a high percentage of Cryptopunks that has never been sold in the open market: specifically, more than a third. In addition, a significant number of Punks sees values below $500. In total, these two groups make up more than half of the Punks. In addition to this bias in terms of unsold or cheap Punks, the price distribution also reflects a very marked skew to the right which shows those units that have been sold at a very high price.



Price range	Total	%
worthless	3656	36.57
0 $ -100 $	953	9.53
100 $ -500 $	599	5.99
500 $ -1,000 $	113	1.13
1,000 $ -10,000 $	451	4.51
10,000 $ -100,000 $	2285	22.86
100,000 $ -500,000 $	1741	17.42
500,000 $ -1,000,000 $	163	1.63
+ 1,000,000 $	36	0.36

Figure 6: CryptoPunks distribution according to price range

Top 100 most expensive CryptoPunks Analysis

Figure 7 analyses which are the characteristics of the CryptoPunks that have had a higher purchase price. It can be seen that the most expensive CryptoPunks have one thing in common: their rarity rating is low. This might point to a trend or correlation in which the higher the rarity of the CryptoPunk, the more expensive the Punk will be.

CryptoPunk	Rarity	Skin	Type	Traits	ETH Amount	Amount $	Date
3100	15	Alien-Skinned	Alien	Headband	4,200	7,580,000	03/11/2021
7804	12	Alien-Skinned	Alien	Cap Forward, Pipe, Small Shades	4,200	7,570,000	11/03/2021
5213	33	Ape-Skinned	Ape	Gold Chain, Knitted Cap	2,250	5,450,000	30/07/2021
7252	512	Zombie-Skinned	Zombie	Chinstrap, Crazy Hair, Earring	1,600	5,330,000	24/08/2021
2338	537	Zombie-Skinned	Zombie	Mohawk Thin	1,500	4,320,000	06/08/2021
2140	50	Ape-Skinned	Ape	Knitted Cap, Small Shades	1,600	3,760,000	30/07/2021
8888	2107	Dark-skinned	Female	Eye Mask, Red Mohawk	888	2,870,000	08/28/2021
7252	512	Zombie-Skinned	Zombie	Chinstrap, Crazy Hair, Earring	1,000	2,530,000	04/08/2021
3831	483	Zombie-Skinned	Zombie	Big Shades, Medical Mask, Vampire Hair	850	2,080,000	07/30/2021
6649	503	Zombie-Skinned	Zombie	Crazy Hair, Beard Dark Front	810	1,980,000	07/31/2021

Figure 7: Top 10 most expensive CryptoPunks

Figures 8 and 9 show which traits are the most common among the 100 most expensive CryptoPunks and those for the Punks yet to be sold. From the outset, some traits do repeat themselves, such as Earring, which is one of the most common traits among the zero value CryptoPunks, and at the same time it is also the most common trait among the 100 CryptoPunks with the highest selling price. Therefore, a clear correlation between traits and selling price does not seem apparent.

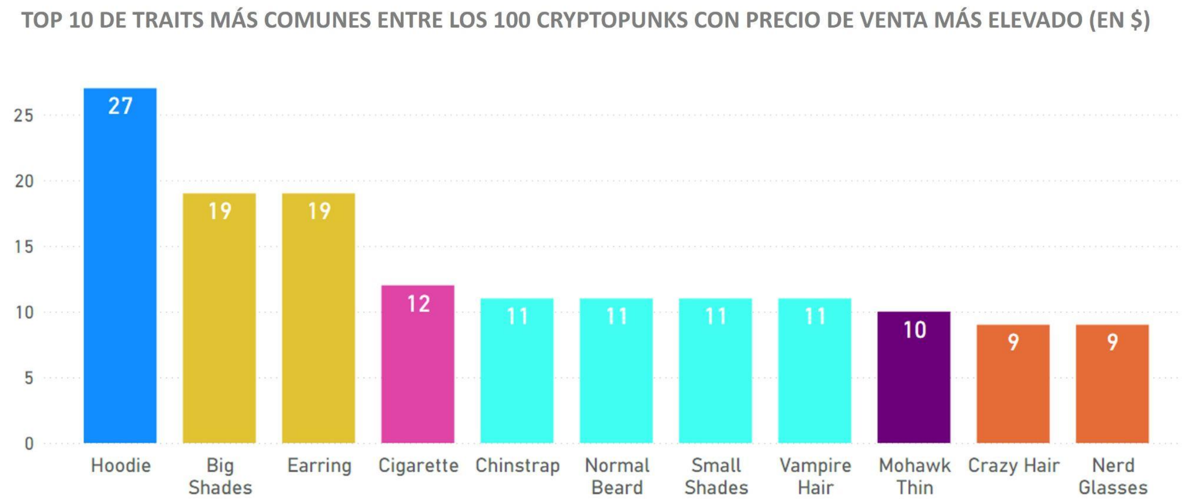


Figure 8: most common characteristics among the most expensive CryptoPunks

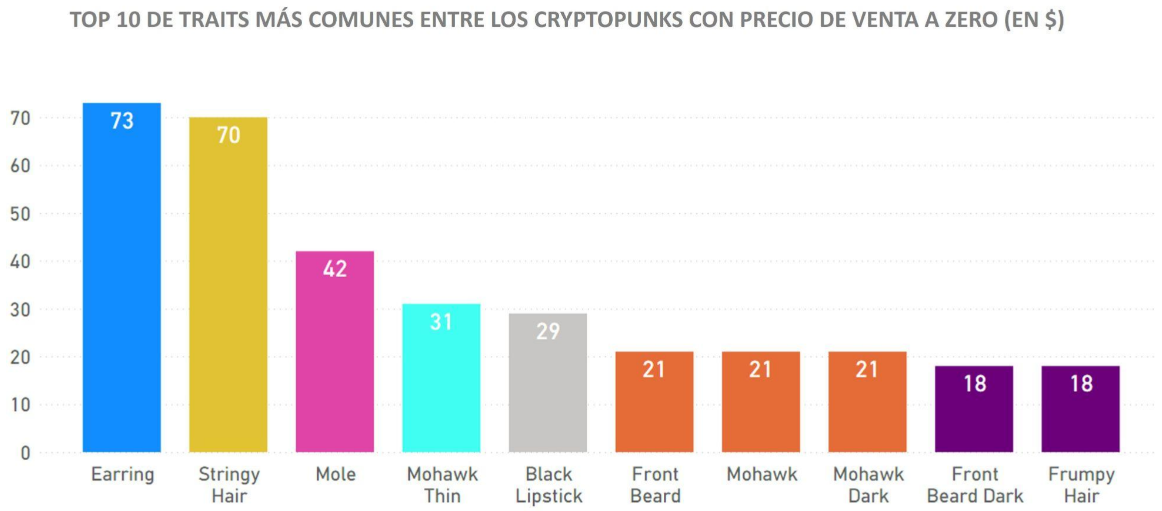


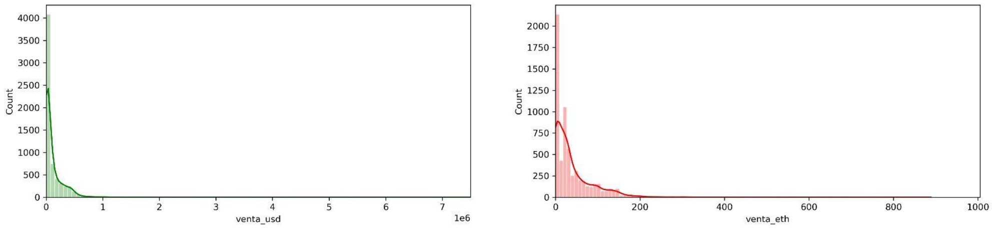
Figure 9: most common characteristics among the CryptoPunks sold with a value of 0.

4.2 Transactions Analysis

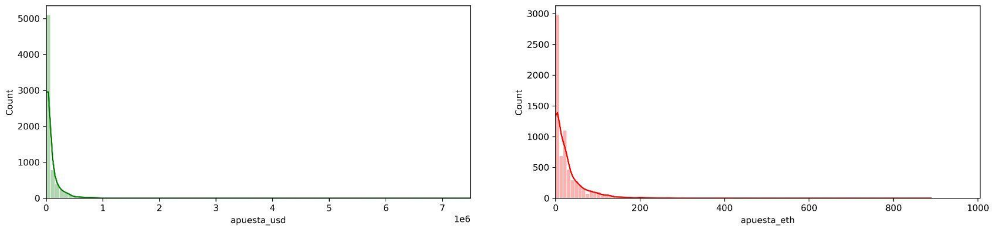
Number of transactions according to monetary mass

The vast majority of transactions have values below a million dollars, as seen in Figures 10 to 12. All three types of transactions are concentrated in relatively small values, which shows a market concentrated in prices around the thousands.

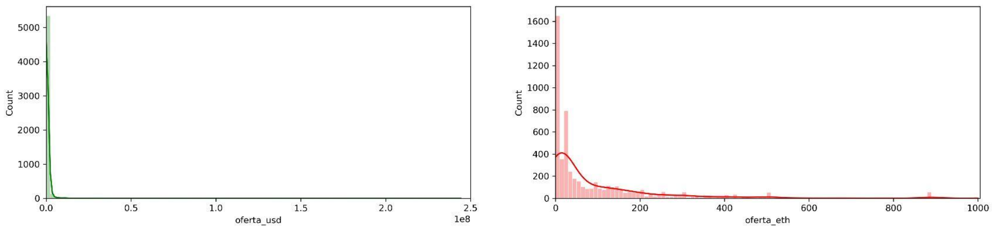
Sales (in USD and ETH)



Bids (in USD and ETH)



Bids (in USD and ETH)



Figures 10, 11 and 12: distribution of sales, bids and offers, respectively, according to frequency and amounts of dollars and ether

According to type of transaction

The evolution of the types of transactions over time reveals a relatively stable sales market since October 2020, with frequency peaks and periods of inactivity. However, these values are almost always above those seen before that same date. Regarding the number of offers, a large increase can be seen in 2021, although it is not accompanied by a proportional increase in sales.

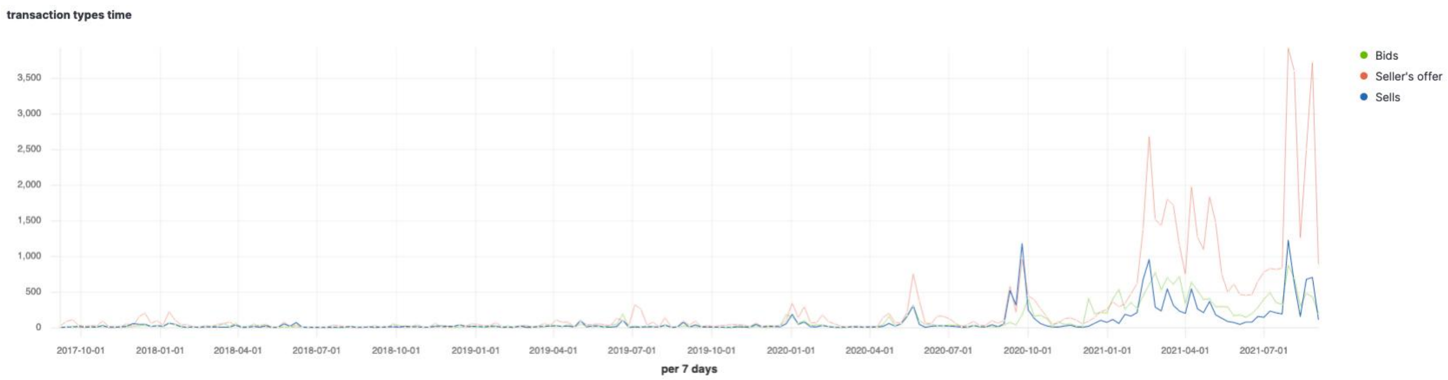


Figure 13: evolution of transactions types over time

Price

The price distribution graph gives very relevant information about the market. In it, it is observed that the maximum sales values observe very sudden and extreme peaks compared to the bulk of the market. More than 75% of values remain relatively close to each other, and below a million dollars. However, the bulk of offers has experienced a significant increase in recent months.

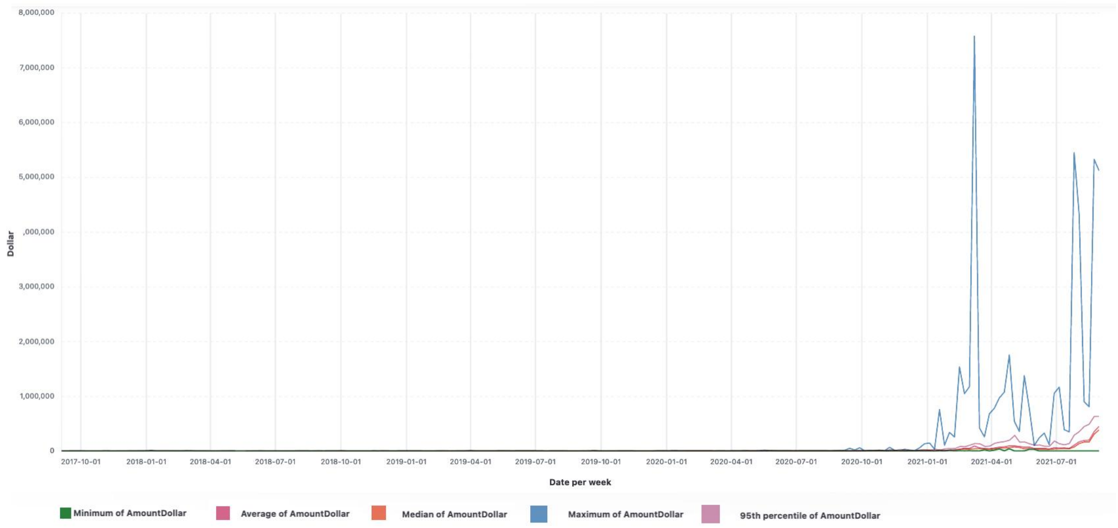


Figure 14: evolution sales according to price in dollars over time

Since March 2021 the price of CryptoPunks has reached maximums never seen before. However, as shown below, these peaks might not be completely random. The same days after the events listed below happened, a great movement in the NFT market was generated:

- **March 11, 2021.** Digital artist Beeple sells the first crypto work of art at the auction house Christie's. Beeple's work reached a value of $69 million. 2 months later and within the framework of the 21st Century Evening Sale, a batch of 9 CryptoPunks was sold in that same auction house for almost 17 million dollars.

- **July 30, 2021.** 104 CryptoPunks were released on the open market in block. The tide was started by businessman Gary Vaynerchuk, who bought a CryptoPunk for 1,600 Ethereums ($3.7 million). That same day, an anonymous investor named 0x650 spent 2,700 Ethereums ($7 million) on 80 CryptoPunks.

- **August 18, 2021.** VISA enters the CryptoPunk market on August 18, generating a large movement of CryptoPunks in the following days.

Punks by number of times sold

The following chart examines the number of times a same CryptoPunk has been sold in the market. Most of them have only been sold once or twice, and more than 75% of Punks have very little circulation (4 or fewer transactions). This stagnation is even more pronounced if we consider that 36% of Punks are excluded from this graph as they have never been sold (see Figure 5).

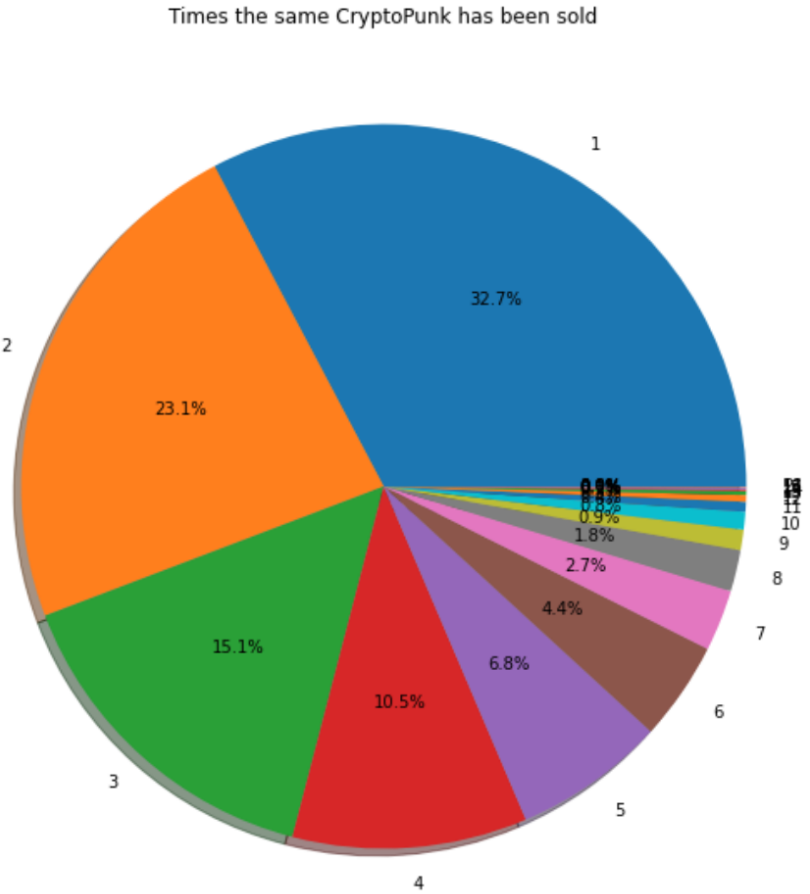
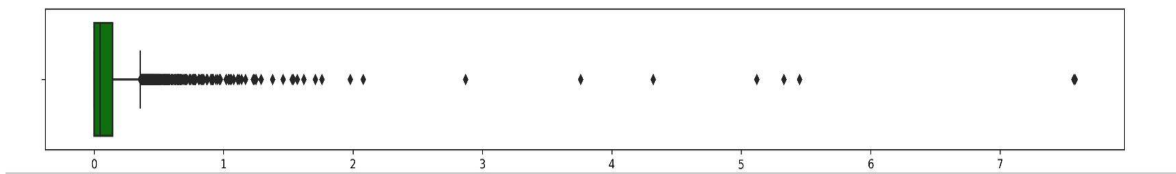


Figure 15: Number of Times Sold over Total Punks Ever Sold

When examining the prices paid the last time each Punk was purchased, the vast majority of prices are below 500,000. Values larger than that are very rare. Therefore, most of the most notorious cases of sale represent minority events within the CryptoPunks market.



Median (in USD)	Standard deviation (in USD)	25% (in USD)	75% (in USD)
4,361	255,718	611.75	143,362

Figure 16: distribution of the last price paid for a Punk (in millions of dollars)

4.3 Owners Analysis

A basic information table information about the current owners of CryptoPunks is quite revealing. First of all, they are very unevenly distributed. More than 75% of the owners have two or fewer Punks, meaning a small minority have a great control over the market: it could indicate a potential presence of *whales*, or excessively influential players.

	CryptoPunks owned	Dollars Spent	Dollars earned	Current Profit (in dollars)
Average	3.272668	359,520.4	340,591.8	-28,928.56
Standard Deviation	14.425881	1,246,380	1,835,839	1,539,829
Minimum	1	0	0	-14,400,000
25%	1	0	0	-173.402
Median	1	77.873	0	-25.626
75%	2	305086.5	45.341	0
Maximum	430	29,150,000	43,950,000	42,160,000

Figure 17: table of CryptoPunks owners and their financial results

One of the characteristics of the CryptoPunks market and emerging NFT markets in general is the disadvantage between the first owners and those who arrived later. This is one of those cases: the first owners got a Punk for free, while the newer buyers have had to spend thousands or millions. In the next figure the current profit (or losses) compared to the date the first Punk was obtained can be seen.

First, in terms of frequency, it can be seen that a large number of owners started in the current year, 2021, which indicates a clear interest in this market. Second, it can be seen that the first owners have certainly made the most money: since they did not have to invest a dollar, subsequent sales have inflated the profit figures. Finally, the conclusions are more mixed in terms of the results of new owners: the balance seems to tilt towards more profits than losses, but losses may also be caused by portfolios that have not yet started to sell their assets.

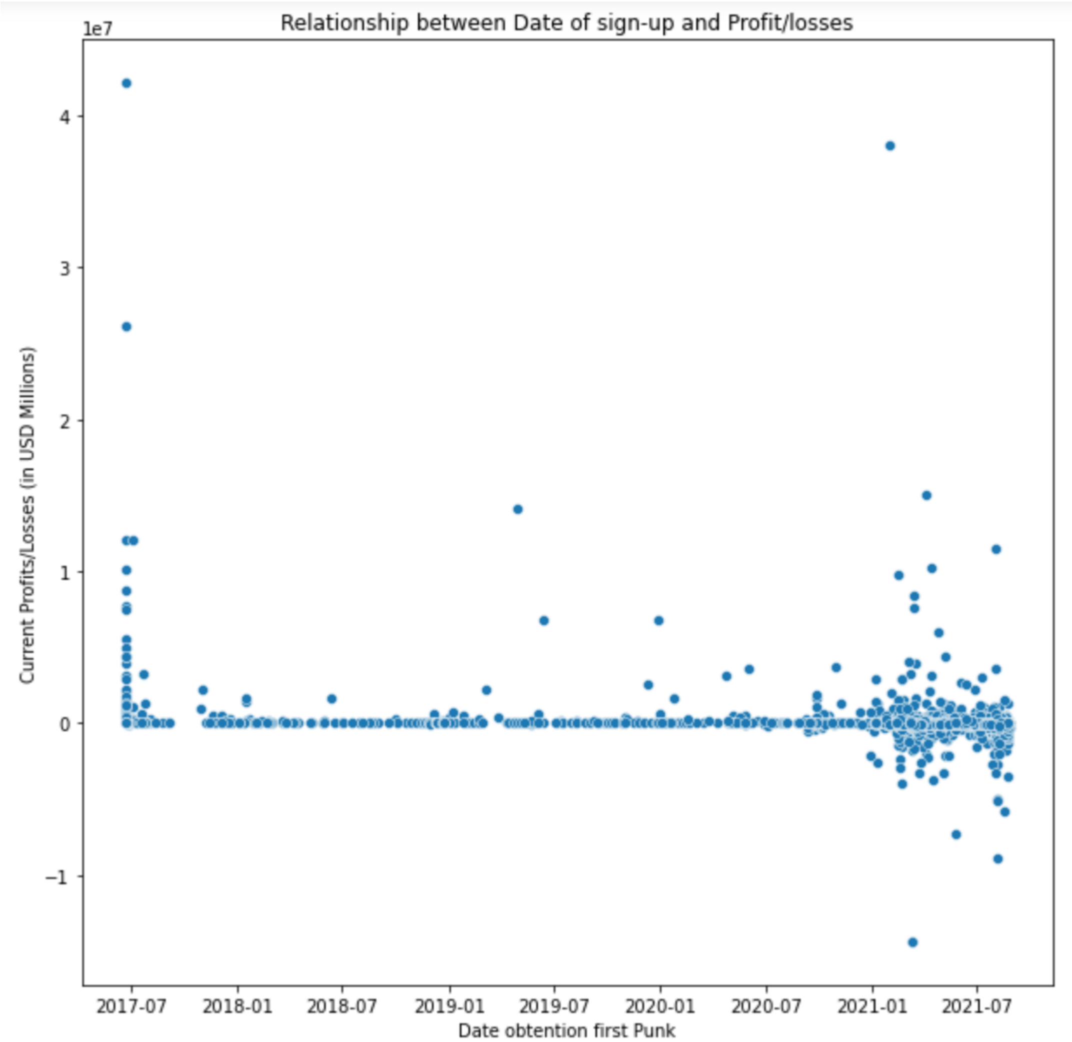


Figure 18: profits / losses (in millions of dollars) by owner according to registration date

5 - Predictive analysis

5.1 Considerations

The reference prices used in predictive analytics are expressed in dollars only. Prices in Ethereum are not shown given its high fluctuation and the fact that prices in dollars are more understandable to the general public. Delving into the analysis of the price fluctuation between Ethereum and dollar / euro is not the object of this study.

Only the transaction type 'Sold' is considered to feed the algorithm. The rest of the transactions, such as offers or auctions, are considered as the result of market volatility or speculation, instead of actual offers (see figures 10, 11, 12 and 13), hence not relevant for this study.

To prevent extreme cases from affecting predictive analytics, certain outliers are removed from the final database. Outliers are those transactions of more than one million dollars, in order to avoid extreme deviations in the predictive algorithm, and at the same time avoid excessive data deletion. These outliers represent less than 1% of the CryptoPunks (see Figure 6).

In the same way, CryptoPunks that have never been sold are removed from the dataframe due to the impossibility of determining the characteristics affecting their price. This group represents slightly more than 30% of the initial data (see Figure 6). These exclusions have been carried out in order to discern the patterns that determine the value of CryptoPunks in the general population.

It has been considered to nurture the algorithm with the Return Rate, since it provides the profitability actually obtained by an owner: having a high profit does not mean having a high profitability, since the purchase price could also be high. Since the second to last sale of 2087 CryptoPunks has a value of 0 (it was claimed for free), the calculation of the Return Rate generates 2087 infinite values, adding more arguments to ignore such cases.

It was tried to execute the algorithm with a binary column for each one of the traits (95 columns in total, including hidden traits). The silhouette score dropped to 0.19, so it was decided to omit these variables, which falls in line with the descriptive analysis shown in Figures 8 and 9 suggesting no relationship.

5.2 Unsupervised analysis

Unsupervised Machine Learning algorithms can be used to discern patterns unidentifiable to the human observer, thus discovering possible determinants of price or profitability that are impossible to detect with a simple descriptive analysis. After testing different unsupervised methods, a k-means method of 7 clusters revealed the maximum information, with a Silhouette Score of 0.63.

Brief description of the 7 clusters by number of Cryptopunks and basic characteristics

Cluster 0: Number of Cryptopunks: 653, all female and Dark-skinned
Cluster 1: Number of Cryptopunks: 387, all male and Albino-skinned
Cluster 2: Number of Cryptopunks: 1254, all male and Light-skinned
Cluster 3: Number of Cryptopunks: 694, all female and Mid-skinned
Cluster 4: Number of Cryptopunks: 1155, all male and Dark-skinned
Cluster 5: Number of Cryptopunks: 1243, all male and Mid-skinned
Cluster 6: Number of Cryptopunks: 922, alien, ape, female (only Albino-skinned) and zombie

Cluster	tipus_Alien	tipus_Ape	tipus_Female	tipus_Male	tipus_Zombie	skin_Albino-skinned	skin_Dark-skinned	skin_Light-skinned	skin_Mid-skinned
0	0	0	653	0	0	0	653	0	0
1	0	0	0	387	0	387	0	0	0
2	0	0	0	1254	0	0	0	1254	0
3	0	0	694	0	0	0	0	0	694
4	0	0	0	1155	0	0	1155	0	0
5	0	0	0	1243	0	0	0	0	1243
6	5	13	869	0	35	215	0	654	0

Figure 19. Clusters obtained after applying a k-means clustering algorithm

As seen in figure 19, 6 of the clusters obtained comprise only Male and Female Types, and one cluster (num. 6) groups the Alien, Ape and Zombie Types, as well as some Female. This falls in line with the descriptive analysis carried out at the beginning of the present study (Figures 1 and 2). Cluster number 6 has a mean and median only slightly higher than the rest of the clusters, so the elimination of outliers (and not types) was effective.

Profitability by clusters

CLUSTER	Average	Maximum	Minum	Median
0	63,091.86	691,939	-59,635	17,716
1	69,257.01	629,866	-38,250	17,666
2	65,720.57	901,932	-90,836	15,071
3	62,756.87	973,803	-116,947	18,440
4	62,820.54	913,863	-155,524	16,026
5	59,117.56	970,009	-138,938	17,282
6	71,857.32	954,650	-155,577	21,040

Figure 20. Profitability of the last sale (in USD) ordered by clusters.

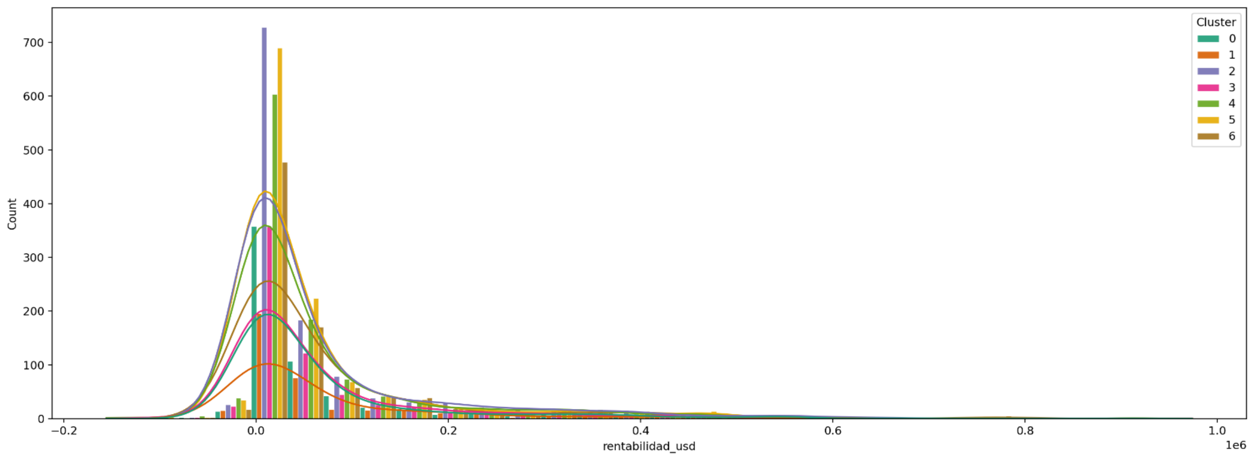


Figure 21. Histogram of profitability by cluster

Figures 20 and 21 show the profitability data (from the last sale) of the different clusters. Although the data seems similar overall, some difference can be considered that will be developed in the conclusions section.

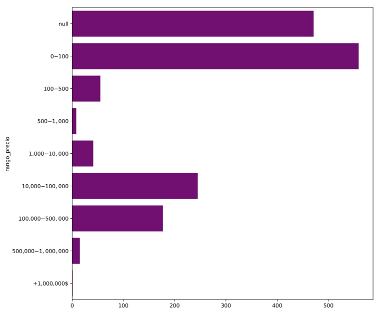


Figure 22a

Cluster 0
1572 CPs in total
30% (471 CPs) have never been sold and 1 exceeded one million in the last purchase.

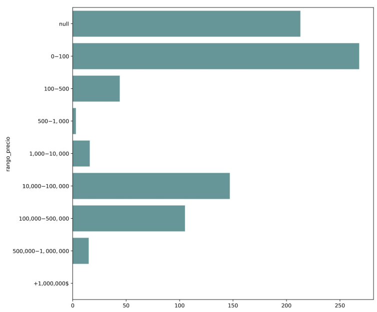


Figure 22b

Cluster 1
811 CPs in total
26% (213 CPs) have never been sold and none exceeded one million in the last purchase.

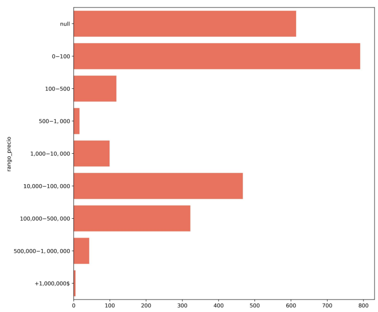


Figure 22c

Cluster 2
2475 CPs in total
25% (614 CPs) have never been sold and 5 exceeded one million in the last purchase.

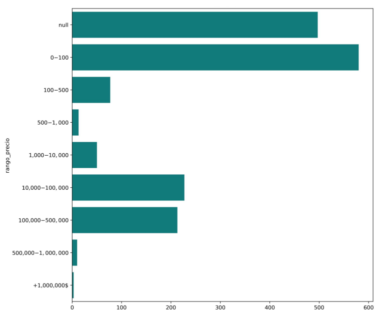


Figure 22d

Cluster 3
1670 CPs in total
30% (497 CPs) have never been sold and 3 exceeded one million in the last purchase.

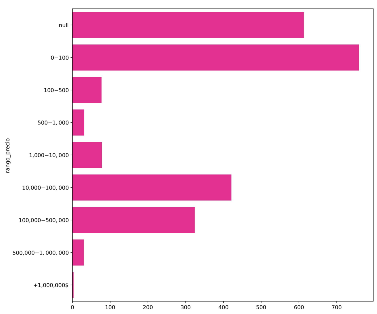


Figure 22e

Cluster 4
2336 CPs in total
26% (613 CPs) have never been sold and 3 exceeded one million in the last purchase.

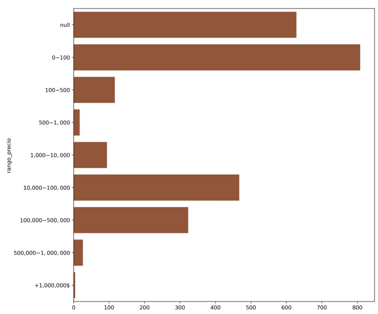


Figure 22f

Cluster 5
2483 CPs in total
25% (628 CPs) have never been sold and 4 exceeded one million in the last purchase.

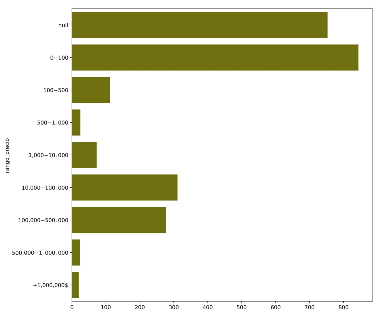


Figure 22g

Cluster 6
2439 CPs in total
30% (753 CPs) have never been sold and 20 exceeded one million in the last purchase.

Figures 22a to 22g show the distribution of the latest purchase prices of each cluster over the total of 10,000 CryptoPunks. There are no significant differences to rule out the hypothesis that there are no differences between CryptoPunks but there are nuances that can influence decision-making, especially in clusters 1 and 6. Cluster 1 groups a very low number of CryptoPunks (811, compared to the average of 2,162 of the other 6 clusters), it is the least dispersed, the only one with no sales above one million and the one with the less remarkable losses: 6% over those which have had more than one sale, with a maximum loss of 38,250 USD (39% loss). It also has a characteristic type of skin (Albino-skinned), which although not exclusive, is relatively rare. Cluster 6 stands out for grouping not only the rarest types but also for having achieved the highest number of sales over a million dollars and having the highest loss: 155,577 USD (a 48% loss).

5.3 Causal analysis

A one-way ANOVA analysis was performed with the clusters resulting from the unsupervised analysis. The goal of this method is to identify whether there are significant differences between the value distributions of the different clusters; more specifically, in relation to the last sale price and profitability. A p-value lower than 0.05 would reject the null hypothesis, which consists of no difference between the clusters.

Since each cluster contains different numbers of cases, the difference should have been imputed as if they were null values. This is due to the need for ANOVA to have the same number of samples in each variable. Depending on how they are imputed, a lower or higher value of statistical validity is obtained, which can be a limitation for analysis. The first relationship studied is between the different clusters and the last sale price.

Replacing NaNs with:	F-score	p-score
Column mean	0.007010181844102753	2.9533474945171285
Column median	17.608762400176733	2.223983671508818e-20

Figure 23. ANOVA test results; target variable: last sale price

An ANOVA test was also carried out on the relationship between the different clusters and the profitability of the last sale. The results are as follows:

Replacing NaNs with:	F-score	p-score
Column mean	0.023433893112453804	2.4380344747663147
Column median	17.66123493384489	1.9146669966000232e-20

Figure 24. ANOVA test results; objective variable: profitability

These data indicate a statistically significant difference between the cluster means. Even so, due to the nature of the ANOVA analysis, it is not possible to determine which clusters differ from the rest. Due to this and the bias caused by the addition of artificial values (mean or median), a *post hoc* test is performed to further study the relationship between variables. The Tukey HSD test compares the difference between all the groups and allows the analysis of a database with unequal groups, thus eliminating the bias caused by the imbalance of the number of components of the clusters.

Figures 25 and 26 show the results of the Tukey test in relation to the sale price and profitability. The null hypothesis of no relationship cannot be rejected in any of the cases, indicating the non-existence of determining price patterns. This analysis also discredits the previous run of ANOVA, probably affected by the imputation of null values. In conclusion, there is no cluster that is significantly different from the rest.

According to profitability						
Group 1	**Average**	**Group 2**	**Average**	**Difference (absolute value)**	**p-value**	**Rejects H0?**
0	63,091.87	1	69,257.01	6,165.14	0.9	No
0	63,091.87	2	65,720.58	2,628.71	0.9	No
0	63,091.87	3	62,756.88	334.99	0.9	No
0	63,091.87	4	62,820.55	271.32	0.9	No
0	63,091.87	5	59,117.56	3,974.31	0.9	No
0	63,091.87	6	71,857.32	8,765.45	0.7483	No
1	69,257.01	2	65,720.58	3,536.43	0.9	No
1	69,257.01	3	62,756.88	6,500.13	0.9	No
1	69,257.01	4	62,820.55	6,436.46	0.9	No
1	69,257.01	5	59,117.56	10,139.45	0.7271	No
1	69,257.01	6	71,857.32	2,600.31	0.9	No
2	65,720.58	3	62,756.88	2,963.7	0.9	No
2	65,720.58	4	62,820.55	2,900.03	0.9	No
2	65,720.58	5	59,117.56	6,603.02	0.7739	No
2	65,720.58	6	71,857.32	6,136.74	0.891	No
3	62,756.88	4	62,820.55	63.67	0.9	No
3	62,756.88	5	59,117.56	3,639.32	0.9	No
3	62,756.88	6	71,857.32	9,100.44	0.6974	No
4	62,820.55	5	59,117.56	3,702.99	0.9	No
4	62,820.55	6	71,857.32	9,036.77	0.5825	No
5	59,117.56	6	71,857.32	12,739.76	0.1623	No

Figure 25. Tukey HSD test results for profitability

According to sale price						
Group 1	**Average**	**Group 2**	**Average**	**Difference (absolute value)**	**p-Value**	**Reject H0?**
0	105,748.95	1	111,541,61	5,792.66	0.9	No
0	105,748.95	2	107,278.66	1,529.71	0.9	No
0	105,748.95	3	99,348.22	6,400.73	0.9	No
0	105,748.95	4	109,366.31	3,617.36	0.9	No
0	105,748.95	5	97,401.79	8,347.16	0.7399	No
0	105,748.95	6	114,067.78	8,318.83	0.9	No
1	111,541,61	2	107,278.66	4,262.95	0.9	No
1	111,541,61	3	99,348.22	12,193.39	0.9	No
1	111,541,61	4	109,366.31	2,175.3	0.9	No
1	111,541,61	5	97,401.79	14,139.82	0.6954	No
1	111,541,61	6	114,067.78	2,526.17	0.9	No
2	107,278.66	3	99,348.22	7,930.44	0.9	No
2	107,278.66	4	109,366.31	2,087.65	0.9	No
2	107,278.66	5	97,401.79	9,876.87	0.7114	No
2	107,278.66	6	114,067.78	6,789.12	0.9	No
3	99,348.22	4	109,366.31	10,018.09	0.6466	No
3	99,348.22	5	97,401.79	1,946.43	0.9	No
3	99,348.22	6	114,067.78	14,719.56	0.5692	No
4	109,366.31	5	97,401.79	11,964.52	0.1941	No
4	109,366.31	6	114,067.78	4,701.47	0.9	No
5	97,401.79	6	114,067.78	16,665.99	0.1559	No

Figure 26. Results of the Tukey HSD test for sale price

6 - Conclusions

The CryptoPunk market is a highly attractive market for investment, especially due to the price evolution observed since March 2021 (see figure 14). Extracting lessons to predict which CryptoPunks represent a better investment is the objective of this report, as well as to determine the risk to which investors are exposed when betting on this market.

After analyzing the traits or descriptive characteristics of the CryptoPunks (95 in total counting the hidden traits) it is concluded that they have not been decisive until now to define the price of CryptoPunks and, therefore, having one or the other trait does not affect their value. It is also important to note that there are 36% of CryptoPunks that have not yet had any type of transaction and, therefore, data is lacking to be able to predict for what price they could be sold in a hypothetical future. In the descriptive analysis (Figures 8 and 9), it is already observed that several of the descriptive characteristics are repeated between the CryptoPunks with the highest selling price and the CryptoPunks that have had a zero sale price. Hence, it is concluded that the descriptive characteristics of the CryptoPunks are not an important variable for their performance.

Another interesting variable is the rarity ranking. Although initially the descriptive analysis points to the possibility that said index could influence the sale price of a CryptoPunk (figure 3), once the data of extreme cases is cleaned, a clear and direct relationship is not observed. Although some of the more expensive CryptoPunks are considered rarer, this study does not find this a determining index to define the future price of a given CryptoPunk.

Furthermore, a descriptive analysis of the CryptoPunk Owners has been carried out. There is a great inequality between the amount that some owners had to spend to acquire a Punk while the more pioneering users received many CryptoPunks for free. Many owners have either obtained all of their Punks when they were free or have not yet sold them (see figure 17), which may indicate many current inactive users. A really high and unequal amount in terms of invested capital is observed, since the vast majority of transactions are of low value. Still, a generalised price increase can be seen even beyond extreme transactions. In the same way, the Punks in circulation also move unevenly: less than 75% have changed hands more than 3 times. Finally, many owners seem to have lost money so far, and in large amounts, which could indicate a large speculative bubble or otherwise a future investment.

A considerable price disparity is observed: from CryptoPunks that have been given away, to some with a purchase price of 7.5 million dollars. To improve analysis, price dispersion has been reduced (Figures 1 and 2) with the treatment of outliers. This treatment is considered effective, since the clusters resulting from the predictive analysis are grouped based on the price of the last sale and the type of skin. However, such clusters do not allow a price prediction of CryptoPunks that are not yet for sale. It is considered that in order to carry out a reliable price prediction study, the market is not mature enough: it would be convenient if all CryptoPunks had had at least some transaction and / or have a longer sales history throughout time.

As mentioned in the Considerations section of this study, the profitability rate of CryptoPunks that had been sold was analyzed. Yes, it can be stated that until September 8, 2021 (data collection date), 92.72% of sales have obtained a profit and, of these, 71.76% have obtained a profitability rate higher than 50%. Thus, it is stated that up to the aforementioned date, the investments made in the CryptoPunks market have generated a high rate of return for its investors. However, the unsupervised algorithm is not able to establish a clear relationship and allow us to make future predictions between this rate of return and the traits, skin and / or types of CryptoPunks. Through the Tukey HSD test (figures 25 and 26), where the relationship between sale price and profitability is analyzed, the non-existence of determining price patterns is evident. The high volatility of the NFTs markets also makes it difficult to establish a relationship between past rates of return and possible future profitability tables. And the word 'possible' is used since it is not known to this day if in the future NFTs products will be considered a safe haven value (such as gold) or are only the result of speculation due to their non-fungible nature and, therefore, may crumble in value in the future.

The types and skins of the CryptoPunks are the variables that the unsupervised algorithm has considered for the grouping of the data in clusters. In the profitability analysis of the resulting clusters, we observe that the data are quite homogeneous (figure 20), which is confirmed in the result of the Tukey HSD test for the profitability of the clusters (figure 25): no significant differences are observed between them.

For all that was discussed earlier in this section, we conclude that this study has not found sufficient evidence to detect credible price patterns for CryptoPunks. Predictive analytics has not revealed any specific group with greater economic potential, once extreme cases have been ignored. Although the clustering algorithm reveals possible behaviors differentiated by type or skin, subsequent analysis and descriptive measures of the clusters rule out this possibility.

From the analysis it appears that CryptoPunks are a highly speculative market: values are rapidly and dramatically changing and explanatory patterns appear to be non-existent. In the same way, it is a very unequal market in all aspects: from the monetary amounts at stake to the Punks in possession, the risk of economic losses is very unevenly distributed, to the disadvantage of new players. It is a high risk market, as a decline in interest or investment could seriously weigh down investors' profits, which in turn are further dependent on the endless growth of the market.

In addition, the shadow that the NFT market is a big bubble has been a constant fear since its inception. For this reason, the entry of large economic corporations such as VISA gives credibility to the market and a greater perception as a safe value in the face of the great intrinsic volatility of NFTs. Even so, the liquidity when trying to recover the investment is not guaranteed at all and depends on the future growth of the market.

Regarding this last point, it should be noted that the value of this digital art is also largely influenced by the actions of public personalities such as celebrities (such as Ashton Kutcher), athletes (Josh Hart, NBA) or businessmen (for example, Shalom Meckenzie, the largest shareholder in DraftKings). The sale of some CryptoPunks at one of the world's leading auction houses, Christie's, reinforces this argument. Most mediatic interventions tend to see a significant increase in the pace of the market, constituting further proof of the volatility of the environment.

7 - References and sources:

Bybit Learn (2021, June 9). What Exactly Are CryptoPunks and Why Should You Care? *Bybit Learn.* Recovered from: https://learn.bybit.com/crypto/what-are-cryptopunks/

DefyPunk (2021). Recovered from: www.defypunk.com

Gonzalez, J. (2021, July 29). 10 celebrities who talk about bitcoin, buy it or love it (2021). *Criptonoticias*. Retrieved from: https://www.criptonoticias.com/comunidad/entrentación/famosos-hablan-compran-aman-bitcoin/

Havelli, D. (2021, August 2). NFT Madness Accelerates on Ethereum After User Purchases 88 Punks for $ 5,000,000. *BSC News.* Retrieved from: https://www.bsc.news/post/nft-madness-accelerates-on-ethereum-after-user-purchases-88-punks-for-5-000-000

Havelli, D. (2021, 24 August). Crypto Punks Continue Mainstream Interest Stemming From Visa and Odell Beckham Jr. Purchases. *BSC News.* Recovered from: https://www.bsc.news/post/crypto-punks-continue-mainstream-interest-stemming-from-visa-and-odell-beckham-jr-purchases

JD Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007

Larva Labs (2021). CryptoPunks. Recovered from: https://www.larvalabs.com/cryptopunks

Luzan, A. (2021, May 8). Art Reimagined: NFTs are changing the market for collectibles. *COINTELEGRAPH in Spanish.* Recovered from: https://es.cointelegraph.com/news/art-reimagined-nfts-are-changing-the-collectibles-market

McSweeney, M. (2021, August 29). NFT Wrap: The $2 billion month. *The Block.* Recovered from: https://bit.ly/3yNR4SE

NumPy (Version 1.21) [Computer Software]

pandas (Version 1.3.3) [Computer Software]. (2008)

Python (Version 3.7.9) [Computer Software]

Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011

seaborn (Version 0.11.2) [Computer Software]. (2012)

Selenium (Version 1.3.3) [Computer Software]