

Tarea Bases de Datos NOSQL

Alumno: Paloma Holgado Pretell

Introducción

NFT es una expresión que se escucha cada día más, incluso fuera del mundo crypto. Las siglas NFT significan Non-Fungible Token, estos activos son únicos que no se pueden modificar ni intercambiar entre sí, ya que no hay dos NFTs que sean equivalentes. Se puede pensar en estos activos como obras de arte.

En este trabajo voy a analizar las transacciones de los Cryptopunks en OpenSea, la plataforma más conocida como Marketplace para NFTs. Los Cryptopunks son un proyecto de LarvaLabs que consiste en 10.000 NFTs tokenizados en la blockchain de Ethereum. Los Cryptopunks son avatares de 24x24 pixeles basados en la estética punk de los 70. Cada Cryptopunk se genera aleatoriamente a partir de una lista de características, que incluyen desde las “pieles” que pueden ser personas, zombies, alienígenas o simios, hasta accesorios, como gorros, lentes, piercings y más.



Algunos datos curiosos sobre los Cryptopunks son que fue uno de los primeros proyectos en convertir NFTs en la blockchain de Ethereum y, cuando se lanzaron en 2017, 9.000 Cryptopunks estaban disponibles gratuitamente para el público mientras que LarvaLabs se quedó con los 1.000 restantes. A finales de 2020 fue cuando la demanda por estos coleccionables se disparó y finalmente explotó en 2021 con 2.500 millones de dólares en volumen de transacciones en los primeros seis meses.

Base de Datos

La base de datos que utilicé en este trabajo contiene las transacciones que se realizaron con los cryptopunks entre el 2017 y el 2021. La fuente de esta base de datos es kaggle.com. Contiene campos como el id de la transacción, el tipo de transacción, el emisor y el receptor de la transacción, la fecha y la hora, la fuente, el precio denominado en Ethereum, el número del punk, la dirección de la billetera virtual de la cual sale y a la cual llega la transacción, el tipo de punk, que es un array, y los accesorios que tienen los punks, que también es un array.

Para esta tarea, realicé el ejercicio tanto en MongoDB como en Python por lo tanto este documento tiene a continuación un apartado con las consultas hechas en NOSQLBooster, y luego las mismas consultas realizadas en Jupyter Notebooks con Python.

NoSQLBooster

Para iniciar el ejercicio importé el documento “cryptopunks.json” en NoSQLBooster. Cree un database llamado “tarea”, y adentro de esta, la colección “cryptopunks” que tiene los documentos del json importado. Haciendo un find de toda la colección de cryptopunks podemos ver que esta posee 167.492 documentos, es decir que tiene información de 167.492 transacciones. A su vez se puede ver a continuación que los campos de “type” y “accesories” contienen arrays.

```
db.cryptopunks.find()
```

| 3 | db.cryptopunks.find() |
|-------------------------------|--------------------------------------|
| 4 | |
| | cryptopunks 0.026 s 167,492 Docs |
| | |
| Key | Value |
| ↳ (1) 643edf4270bbafc5deecd53 | {13 fields} |
| ↳ _id | 643edf4270bbafc5deecd53 |
| ↳ txn_type | Bid Withdrawn |
| ↳ from | 0xe73a1d |
| ↳ to | |
| ↳ date | 2021-09-04 |
| ↳ timestamp | 2021-09-04T00:00:00 |
| ↳ source | larvalabs |
| ↳ eth | 321 |
| ↳ punk_id | 0 |
| ↳ from_wallet_address | null |
| ↳ to_wallet_address | null |
| ↳ type | Array[1] |
| ↳ 0 | Female |
| ↳ accessories | Array[3] |
| ↳ 0 | Green Eye Shadow |
| ↳ 1 | Earring |
| ↳ 2 | Blonde Bob |

Consultas

Como primera consulta quise verificar el número de cryptopunks que existen. La consulta que realicé saca primero una lista con los valores únicos de “punk_id” y luego los cuenta. Podemos ver en el resultado que efectivamente hay 10000 cryptopunks únicos en el mercado.

```
db.cryptopunks.distinct("punk_id").length
```

| | |
|---|---|
| 7 | db.cryptopunks.distinct("punk_id").length |
| 8 | |
| 9 | 0.187 s |
| 1 | 10000 |

Para iniciar también quise averiguar cuando se realizó la primera transacción que está en la base de datos, y la última. La primera podemos ver que realizó el 23 de junio de 2017 a las 00:00, y fue una venta del punk 0 por 0,98 Ether, que hoy en día serían aproximadamente 1.675 euros. Por otro lado, la última transacción en la colección se realizó el 7 de octubre de 2021 y fue una oferta por el punk número 9039 por 288.88 Ether.

```
db.cryptopunks.find({}, {"_id":0, "from_wallet_address":0, "to_wallet_address":0, "source":0}).sort({timestamp:1}).limit(1)
```

| | |
|--|---|
| 11 | db.cryptopunks.find({},{ "_id":0, "from_wallet_address": 0, "to_wallet_address": 0, "source":0}).sort({timestamp:1}).limit(1) |
| 12 | |
|  | |
| | |

```
db.cryptopunks.find({}, {"_id":0, "from_wallet_address":0, "to_wallet_address":0, "source":0}).sort({timestamp:-1}).limit(1)
```

| 13 | db.cryptopunks.find({}, {"_id": 0, "from_wallet_address": 0, "to_wallet_address": 0, "source": 0}).sort({timestamp:-1}).limit(1) | | | | | | |
|-------------|--|----|------------|---------------------|--------|---------|----------------|
| 14 | | | | | | | |
| cryptopunks | | | | | | | |
| | 0.163 s 1 Doc | | | | | | |
| txn_type | from | to | date | timestamp | eth | punk_id | type |
| Offered | 0x38b80a | | 2021-10-07 | 2021-10-07T18:28:15 | 288.88 | 9039 | [{"Array": 1}] |

Tipos

Para conocer los distintos tipos de cryptopunks que existen utilice la siguiente query y podemos ver que hay 5 tipos distintos de cryptopunks: Alien, Ape, Female, Male y Zombie.

```
db.cryptopunks.distinct("type")
```

```
13 db.cryptopunks.distinct("type")
14
0.137 s | 5 Docs
value
1 Alien
2 Ape
3 Female
4 Male
5 Zombie
```

A continuación, hice una consulta para determinar cuántos cryptopuks pertenecen a cada tipo. Para hacer esta consulta utilicé el método de agregación dado que, primero necesitaba realizar un “unwind” del campo “type” porque contiene arrays y necesitaba que queden como valores únicos. Después agrupé por los campos “punk_id” y “type”, dado que tengo cada punk_id y su tipo repetido por cada transacción que se realizó con ese NFT. Luego realicé otra agrupación por el “type” porque lo que quiero es contar cuantas veces se repite cada tipo, no necesito los punk_id para esto. Finalmente, ordené la salida por orden descendente en la cuenta de los tipos, de forma que quede primero el tipo que más se repite y último el que menos. Como podemos ver en el resultado, el tipo más común es “Male” y el más raro es “Alien”.

```
db.cryptopunks.aggregate( [ { $unwind: "$type" },
    { $group: { _id: { punk_id: "$punk_id", type: "$type" } } },
    { $group: { _id: "$_id.type", count: { $sum: 1 } } },
    { $sort: { count: -1 } } ] )
```

```
16 db.cryptopunks.aggregate( [ { $unwind: "$type" },
17 | | | | | { $group: { _id: { punk_id: "$punk_id", type: "$type" } } },
18 | | | | | { $group: { _id: "$_id.type", count: { $sum: 1 } } } },
19 | | | | | { $sort: { count: -1 } } ] )
```

| cryptopunks | | 0.346 s | 5 Docs | | | 100 | | p. 1 |
|-------------|--------------|---------|--------|--|--|-----|--|------|
| _id | count | | | | | | | |
| 1 Male | 6,039 (6.0K) | | | | | | | |
| 2 Female | 3,840 (3.8K) | | | | | | | |
| 3 Zombie | 88 | | | | | | | |
| 4 Ape | 24 | | | | | | | |
| 5 Alien | 9 | | | | | | | |

Accesorios

En la siguiente consulta vemos los accesorios que más se repiten y los que menos. Hay 87 accesorios distintos y como se puede ver el accesorio más común es el “Earring”, habiendo 2.459 punks que lo tienen. A este le sigue el “Cigarette” con 961 punks que lo tienen. Entre los menos comunes están el “Choker” con 48 punks y el “Beanie” con 44 punks que lo tienen. A su vez, hay 8 punks que no tienen ningún accesorio.

```
db.cryptopunks.aggregate( [ { $unwind : {path: "$accessories", preserveNullAndEmptyArrays: true} },
    { $group: { _id: { punk_id: "$punk_id", accessories: "$accessories" } } },
    { $group: { _id: "$_id.accessories", count: { $sum: 1 } } },
    { $sort: { count: -1 } } ] )
```

```
26 db.cryptopunks.aggregate( [ { $unwind : {path: "$accessories", preserveNullAndEmptyArrays: true} },
27 | | | | | { $group: { _id: { punk_id: "$punk_id", accessories: "$accessories" } } },
28 | | | | | { $group: { _id: "$_id.accessories", count: { $sum: 1 } } },
29 | | | | | { $sort: { count: -1 } } ] )
```

| cryptopunks | | 0.546 s | 88 Docs | | | 100 | | p. 1 |
|----------------------|--------------|---------|---------|--|--|-----|--|------|
| _id | count | | | | | | | |
| 1 Earring | 2,459 (2.5K) | | | | | | | |
| 2 Cigarette | 961 | | | | | | | |
| 3 Hot Lipstick | 696 | | | | | | | |
| 4 Purple Lipstick | 655 | | | | | | | |
| 5 Mole | 644 | | | | | | | |
| 6 Black Lipstick | 617 | | | | | | | |
| 7 Nerd Glasses | 572 | | | | | | | |
| 8 Big Shades | 535 | | | | | | | |
| 9 Horned Rim Glasses | 535 | | | | | | | |
| 10 Regular Shades | 527 | | | | | | | |
| 11 Shadow Beard | 526 | | | | | | | |

| | | |
|----|-----------------|-----|
| 78 | Top Hat | 115 |
| 79 | Pink With Hat | 95 |
| 80 | Pigtails | 94 |
| 81 | Welding Goggles | 86 |
| 82 | Buck Teeth | 78 |
| 83 | Orange Side | 68 |
| 84 | Tiara | 55 |
| 85 | Pilot Helmet | 54 |
| 86 | Choker | 48 |
| 87 | Beanie | 44 |
| 88 | null | 8 |

Para la siguiente consulta busqué la cantidad de punks que tienen cada cantidad de accesorios. Como podemos ver Hay 333 punks que tienen un solo accesorio, la mayoría de los punks tienen 3 accesorios y solo 1 tiene 7 accesorios.

```
db.cryptopunks.aggregate( [ { $unwind : {path: "$accessories"} },
    { $group: { _id: { punk_id: "$punk_id", accessories: "$accessories" } },
    { $group: { _id: "$_id.punk_id", count_accessories: { $sum: 1 } } },
    { $group: { _id: "$count_accessories", count: { $sum: 1 } } },
    { $sort: { _id: 1 } } ] )
```

```
34 db.cryptopunks.aggregate( [ { $unwind : {path: "$accessories"} },
35 | | | | | { $group: { _id: { punk_id: "$punk_id", accessories: "$accessories" } } },
36 | | | | | { $group: { _id: "$_id.punk_id", count_accessories: { $sum: 1 } } } ,
37 | | | | | { $group: { _id: "$count_accessories", count: { $sum: 1 } } } ,
38 | | | | { $sort: { _id: 1 } } ] )
```

| _id | count |
|-----|--------------|
| 1 | 333 |
| 2 | 3,560 (3.6K) |
| 3 | 4,501 (4.5K) |
| 4 | 1,420 (1.4K) |
| 5 | 166 |
| 6 | 11 |
| 7 | 1 |

Realicé una última consulta para saber cuál es el punk con 7 accesorios y es el número 8348.

```
db.cryptopunks.aggregate([{$project: { punk_id : 1, count: { $size:"$accessories" } }},
    { $sort: { count: -1 } },
    { $limit: 1 } ])
```

```
48 db.cryptopunks.aggregate([{$project: { punk_id : 1, count: { $size:"$accessories" } }},
49 | | | | | { $sort: { count: -1 } },
50 | | | | | { $limit: 1 } ])
```

| _id | punk_id | count |
|-----|--------------------------|-------|
| 1 | 643edf5070bbafcf5df0e282 | 7 |



Transacciones

A continuación, hice una consulta para saber los tipos de transacciones que contiene la colección. De aquí sale que hay 9 tipos distintos de transacciones: Unwrap, Wrap, Bid, Bid Withdrawn, Claimed, Offer Withdrawn, Offered, Sold, Transfer.

```
db.cryptopunks.distinct("txn_type")
```

| | value |
|---|-----------------|
| 1 | (Unwrap) |
| 2 | (Wrap) |
| 3 | Bid |
| 4 | Bid Withdrawn |
| 5 | Claimed |
| 6 | Offer Withdrawn |
| 7 | Offered |
| 8 | Sold |
| 9 | Transfer |

A su vez, entre las transacciones, la que más se realizó fueron ofertas, y podemos ver que hubo 18.979 ventas, lo cual compone el 11% de las transacciones documentadas en la colección.

```
db.cryptopunks.aggregate( [ { $group: { _id: "$txn_type", count: { $sum: 1 } } },
                           { $sort: { count: -1 } } ] )
```

| | _id | count |
|---|-----------------|----------------|
| 1 | Offered | 65,467 (65.5K) |
| 2 | Bid | 26,808 (26.8K) |
| 3 | Sold | 18,979 (19.0K) |
| 4 | Bid Withdrawn | 16,342 (16.3K) |
| 5 | Transfer | 14,225 (14.2K) |
| 6 | Offer Withdrawn | 10,372 (10.4K) |
| 7 | Claimed | 10,000 (10.0K) |
| 8 | (Wrap) | 2,773 (2.8K) |
| 9 | (Unwrap) | 2,526 (2.5K) |

Ventas

Para la siguiente consulta quiero ver las 10 ventas más caras de todo el periodo. Para eso busco los documentos cuya transacción sea "Sold", ordeno por el campo "eth" de forma descendente y limito la salida a 10 filas para que solo me muestre las 10 ventas más caras.

```
db.cryptopunks.find({ "txn_type": "Sold" },
    { "from_wallet_address": 0, "to_wallet_address": 0, "source": 0 }).sort({ "eth": -1 }).limit(10)
```

| | _id | tx_type | from | to | date | timestamp | eth | punk_id | type | accessories |
|----|-------------------------|---------|-------------|-------------|------------|---------------------|--------------|---------|-----------|-------------|
| 1 | 643edf4f70bbafcfc5df0be | Sold | 0x03911f | 0xf4b4a5 | 2021-03-11 | 2021-03-11T00:00:00 | 4,200 (4.2K) | 7804 | ["Alien"] | ["Glasses"] |
| 2 | 643edf4770bbafcfc5def7 | Sold | 0x6611fe | 0x7b8961 | 2021-03-11 | 2021-03-11T00:00:00 | 4,200 (4.2K) | 3100 | ["Alien"] | ["Glasses"] |
| 3 | 643edf4b70bbafcfc5df00 | Sold | Snowfro | 0x8f7cee | 2021-07-30 | 2021-07-30T00:00:00 | 2,250 (2.3K) | 5217 | ["Alien"] | ["Glasses"] |
| 4 | 643edf5170bbafcfc5df10 | Sold | 0x389747 | 0xf2a70b | 2021-09-11 | 2021-09-11T00:00:00 | 2,000 (2.0K) | 8857 | ["Alien"] | ["Glasses"] |
| 5 | 643edf4e70bbafcfc5df09 | Sold | 0x9fce8e | 0x956d14 | 2021-08-24 | 2021-08-24T00:00:00 | 1,600 (1.6K) | 7252 | ["Alien"] | ["Glasses"] |
| 6 | 643edf4570bbafcfc5def3 | Sold | grunare... | gary.vee... | 2021-07-30 | 2021-07-30T00:00:00 | 1,600 (1.6K) | 2140 | ["Alien"] | ["Glasses"] |
| 7 | 643edf4670bbafcfc5def4 | Sold | 0xfb9b8 | 0x7cd3fc | 2021-08-06 | 2021-08-06T00:00:00 | 1,500 (1.5K) | 2338 | ["Alien"] | ["Glasses"] |
| 8 | 643edf4d70bbafcfc5df04 | Sold | Pranksy | 0xd2c1a0 | 2021-09-04 | 2021-09-04T00:00:00 | 1,320 (1.3K) | 6275 | ["Alien"] | ["Glasses"] |
| 9 | 643edf4e70bbafcfc5df09 | Sold | thebeaut... | 0x9fce8e | 2021-08-04 | 2021-08-04T00:00:00 | 1,000 (1.0K) | 7252 | ["Alien"] | ["Glasses"] |
| 10 | 643edf4d70bbafcfc5df04 | Sold | 0xaff336 | Pranksy | 2021-09-04 | 2021-09-04T00:00:00 | 1,000 (1.0K) | 6275 | ["Alien"] | ["Glasses"] |

Como podemos ver en el resultado, las dos ventas más caras se realizaron el mismo día y fueron por 4.200 eth, que hoy en día equivale a cerca de 7.175.000 euros con el precio de Ethereum estando a 1.708,3 euros por Ether. Estos dos cryptopunks son el tipo "Alien", el más raro de todos. Buscando los ids podemos saber que estas ventas corresponden a los siguientes punks:



A su vez realicé consultas para saber cuántos punks de cada tipo se vendieron. Se puede ver que se vendieron 7 de los 9 punks que son aliens, tiene sentido que la demanda por este tipo sea alta ya que son el tipo más raro de todos y por ende los más valiosos como colecciónables.

```
db.cryptopunks.aggregate( [ { $unwind: "$type" },
    { $match: { tx_type: "Sold" } },
    { $group: { _id: { punk_id: "$punk_id", type: "$type" } } },
    { $group: { _id: "$_id.type", count: { $sum: 1 } } },
    { $sort: { count: -1 } } ] )
```

```

74 db.cryptopunks.aggregate( [ { $unwind: "$type" },
75 | | | | | { $match: { txn_type: "Sold" } },
76 | | | | | { $group: { _id: { punk_id: "$punk_id", type: "$type" } } },
77 | | | | | { $group: { _id: "{$_id.type}", count: { $sum: 1 } } } ,
78 | | | | | { $sort: { count: -1 } } ] )

```

| cryptopunks | | 0.195 s | 5 Docs | | | | | |
|-------------|--------------|---------|--------|--|--|--|--|--|
| _id | count | | | | | | | |
| 1 Male | 4,080 (4.1K) | | | | | | | |
| 2 Female | 2,242 (2.2K) | | | | | | | |
| 3 Zombie | 47 | | | | | | | |
| 4 Ape | 17 | | | | | | | |
| 5 Alien | 7 | | | | | | | |

En la siguiente consulta busco determinar cuáles son los dos aliens que nunca se vendieron. Para hacer esto creé un nuevo campo llamado “sold” que toma el valor 1 si el campo “txn_type” es igual a “Sold”, y el valor 0 de lo contrario. A partir de ahí hice una consulta con aggregate para filtrar solo los aliens y agrupar por punk id para sumar los valores del campo “sold” que creé anteriormente.

```

db.cryptopunks.updateMany({ "txn_type": "Sold"}, { $set: { "sold": 1 } })
db.cryptopunks.updateMany({ "txn_type": { $ne: "Sold" } }, { $set: { "sold": 0 } })
db.cryptopunks.aggregate( [ { $unwind: "$type" },
                           { $match: { type: "Alien" } },
                           { $group: { _id: { punk_id: "$punk_id" }, ventas: { $sum: "$sold" } } },
                           { $sort: { ventas: -1 } } ] )

```

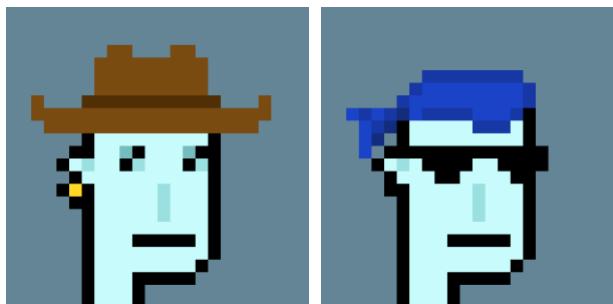
```

82 db.cryptopunks.updateMany({ "txn_type": "Sold"}, { $set: { "sold": 1 } })
83 db.cryptopunks.updateMany({ "txn_type": { $ne: "Sold" } }, { $set: { "sold": 0 } })
84 db.cryptopunks.aggregate( [ { $unwind: "$type" },
85 | | | | | { $match: { type: "Alien" } },
86 | | | | | { $group: { _id: { punk_id: "$punk_id" } , ventas: { $sum: "$sold" } } },
87 | | | | | { $sort: { ventas: -1 } } ] )

```

| cryptopunks | | 1.692 s | 9 Docs | | | | | |
|----------------------|--------|---------|--------|--|--|--|--|--|
| _id | ventas | | | | | | | |
| 1 { punk_id : 2890 } | 2 | | | | | | | |
| 2 { punk_id : 7804 } | 2 | | | | | | | |
| 3 { punk_id : 3100 } | 2 | | | | | | | |
| 4 { punk_id : 7523 } | 1 | | | | | | | |
| 5 { punk_id : 6089 } | 1 | | | | | | | |
| 6 { punk_id : 5822 } | 1 | | | | | | | |
| 7 { punk_id : 5905 } | 1 | | | | | | | |
| 8 { punk_id : 3443 } | 0 | | | | | | | |
| 9 { punk_id : 635 } | 0 | | | | | | | |

De esta consulta podemos ver que hay 3 punks aliens que se vendieron 2 veces y cuatro se vendieron solo una vez. Los que no se vendieron son los que tienen id 3443 y 635.



Python

Para realizar el ejercicio con Python utilicé la interfaz de Jupyter Notebooks lo primero que hice fue instalar al librería “pymongo” que permite comunicarse con el servidor de MongoDB. A continuación, importé los módulos de “pymongo”, “json” y “pandas”.

```
In [2]: !pip install pymongo
import pymongo
import json
import pandas as pd

Collecting pymongo
  Using cached pymongo-4.3.3-cp39-cp39-macosx_10_9_universal2.whl (413 kB)
Collecting dnspython<3.0.0,>=1.16.0
  Using cached dnspython-2.3.0-py3-none-any.whl (283 kB)
Installing collected packages: dnspython, pymongo
Successfully installed dnspython-2.3.0 pymongo-4.3.3
```

Como segundo paso, me conecté al servidor de MongoDB y cree un database llamado “tarea_pymongo” y una colección llamada “cryptopunks”.

```
In [4]: myclient = pymongo.MongoClient("mongodb://localhost:27017/")
db = myclient["tarea_pymongo"]
cryptopunks = db["cryptopunks"]
```

Como último paso del setup, inserté el documento cryptopunks.json en la colección de “cryptopunks”. Una vez hecho este paso ya está todo listo para empezar a hacer consultas.

```
In [5]: with open("cryptopunks.json") as file:
    file_data = json.load(file)

cryptopunks.insert_many(file_data)

Out[5]: <pymongo.results.InsertManyResult at 0x7f860e0fa940>
```

Para realizar el ejercicio en Python utilicé la librería de pandas para poder guardar cada consulta como dataframe y que las salidas sean tablas. Para la primera consulta hice un find de toda la colección, donde se puede ver también la mecánica de los dataframes que expliqué antes. De esta primera consulta podemos ver que esta posee 167.492 documentos, es decir que tiene información de 167.492 transacciones.

```
In [13]: find_all = db.cryptopunks.find()
df_find_all = pd.DataFrame(find_all)
df_find_all
```

| | _id | txn_type | from | to | date | timestamp | source | eth | punk_id | from_wallet_address | to_wallet_address | |
|--------|--------------------------|------------------|----------|----------|------------|---------------------|-----------|--------|---------|---------------------|-------------------|-----|
| 0 | 644a417329549fe3db9c2f92 | Bid Withdrawn | 0xe73a1d | | 2021-09-04 | 2021-09-04T00:00:00 | larvalabs | 321.00 | 0 | None | None | [F] |
| 1 | 644a417329549fe3db9c2f93 | Bid | 0xe73a1d | | 2021-09-01 | 2021-09-01T00:00:00 | larvalabs | 321.00 | 0 | None | None | [F] |
| 2 | 644a417329549fe3db9c2f94 | Bid Withdrawn | 0x2e5e22 | | 2021-09-01 | 2021-09-01T00:00:00 | larvalabs | 320.00 | 0 | None | None | [F] |
| 3 | 644a417329549fe3db9c2f95 | Bid | 0x2e5e22 | | 2021-09-01 | 2021-09-01T00:00:00 | larvalabs | 320.00 | 0 | None | None | [F] |
| 4 | 644a417329549fe3db9c2f96 | Bid Withdrawn | 0x2e5e22 | | 2021-09-01 | 2021-09-01T00:00:00 | larvalabs | 263.00 | 0 | None | None | [F] |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 167487 | 644a417429549fe3db9ebdd1 | Bid | 0xa0a59c | | 2017-07-03 | 2017-07-03T00:00:00 | larvalabs | 0.25 | 9999 | None | None | [F] |
| 167488 | 644a417429549fe3db9ebdd2 | Bid | 0x717403 | | 2017-06-26 | 2017-06-26T00:00:00 | larvalabs | 0.20 | 9999 | None | None | [F] |
| 167489 | 644a417429549fe3db9ebdd3 | Bid | ddavvvee | | 2017-06-23 | 2017-06-23T00:00:00 | larvalabs | 0.20 | 9999 | None | None | [F] |
| 167490 | 644a417429549fe3db9ebdd4 | Bid | 0x5b098b | | 2017-06-23 | 2017-06-23T00:00:00 | larvalabs | 0.15 | 9999 | None | None | [F] |
| 167491 | 644a417429549fe3db9ebdd5 | Claimed | | 0x7e2d1c | 2017-06-23 | 2017-06-23T00:00:00 | larvalabs | NaN | 9999 | None | None | [F] |

167492 rows × 13 columns

Consultas

Como primera consulta quise verificar el número de cryptopunks que existen. La consulta que realicé saca primero una lista con los valores únicos de “punk_id” y luego los cuenta. Podemos ver en el resultado que efectivamente hay 10000 cryptopunks únicos en el mercado.

```
In [24]: len(db.cryptopunks.distinct("punk_id"))
Out[24]: 10000
```

Para iniciar también quise averiguar cuando se realizó la primera transacción que está en la base de datos, y la última. La primera podemos ver que realizó el 23 de junio de 2017 a las 00:00, y fue una venta del punk 0 por 0,98 Ether, que hoy en día serían aproximadamente 1.675 euros. Por otro lado, la última transacción en la colección se realizó el 7 de octubre de 2021 y fue una oferta por el punk número 9039 por 288.88 Ether.

```
In [70]: primera_txn = db.cryptopunks.find({},
    {"_id":0, "from_wallet_address": 0, "to_wallet_address": 0, "source":0}).sort("timestamp",1).limit(1)
df_primera_txn = pd.DataFrame(primera_txn)
df_primera_txn
```

| | txn_type | from | to | date | timestamp | eth | punk_id | type | accessories |
|---|----------|----------|----------|------------|---------------------|------|---------|----------|---|
| 0 | Sold | 0xc352b5 | 0x00d7c9 | 2017-06-23 | 2017-06-23T00:00:00 | 0.98 | 0 | [Female] | [Green Eye Shadow, Earring, Blonde Bob] |

```
In [72]: ultima_txn = db.cryptopunks.find({},
                                         {"_id":0, "from_wallet_address": 0, "to_wallet_address": 0, "source":0}).sort("timestamp",-1).limit(1)
df_ultima_txn = pd.DataFrame(ultima_txn)
```

| | txn_type | from | to | date | timestamp | eth | punk_id | type | accessories |
|---|----------|----------|----|------------|---------------------|--------|---------|--------|--|
| 0 | Offered | 0x38b80a | | 2021-10-07 | 2021-10-07T18:28:15 | 288.88 | 9039 | [Male] | [Messy Hair, Buck Teeth, Regular Shades] |

Tipos

Para conocer los distintos tipos de cryptopunks que existen utilice la siguiente query y podemos ver que hay 5 tipos distintos de cryptopunks: Alien, Ape, Female, Male y Zombie.

```
In [26]: dist_type = db.cryptopunks.distinct("type")
df_dist_type = pd.DataFrame(dist_type)
```

| | 0 |
|---|--------|
| 0 | Alien |
| 1 | Ape |
| 2 | Female |
| 3 | Male |
| 4 | Zombie |

A continuación, hice una consulta para determinar cuántos cryptopunks pertenecen a cada tipo. Para hacer esta consulta utilicé el método de agregación dado que, primero necesitaba realizar un “unwind” del campo “type” porque contiene arrays y necesitaba que queden como valores únicos. Despues agrupé por los campos “punk_id” y “type”, dado que tengo cada punk_id y su tipo repetido por cada transacción que se realizó con ese NFT. Luego realicé otra agrupación por el “type” porque lo que quiero es contar cuantas veces se repite cada tipo, no necesito los punk_id para esto. Finalmente, ordené la salida por orden descendente en la cuenta de los tipos, de forma que quede primero el tipo que más se repite y último el que menos. Como podemos ver en el resultado, el tipo más común es “Male” y el más raro es “Alien”.

```
In [31]: count_type = db.cryptopunks.aggregate( [ { "$unwind": "$type" },
                                                { "$group": { "_id": { "punk_id": "$punk_id", "type": "$type" } } },
                                                { "$group": { "_id": "$_id.type", "count": { "$sum": 1 } } },
                                                { "$sort": { "count": -1 } } ] )
df_count_type = pd.DataFrame(count_type)
```

| | _id | count |
|---|--------|-------|
| 0 | Male | 6039 |
| 1 | Female | 3840 |
| 2 | Zombie | 88 |
| 3 | Ape | 24 |
| 4 | Alien | 9 |

Accesorios

En la siguiente consulta vemos los accesorios que más se repiten y los que menos. Hay 87 accesorios distintos y como se puede ver el accesorio más común es el “Earring”, habiendo 2.459 punks que lo tienen. A este le sigue el “Cigarette” con 961 punks que lo tienen. Entre los menos comunes están el “Choker” con 48 punks y el “Beanie” con 44 punks que lo tienen.

```
In [61]: acc = db.cryptopunks.aggregate( [ { "$unwind": { "path": "$accessories" } },
                                         { "$group": { "_id": { "punk_id": "$punk_id", "accessories": "$accessories" } } },
                                         { "$group": { "_id": "$_id.accessories", "count": { "$sum": 1 } } },
                                         { "$sort": { "count": -1 } } ] )
df_acc = pd.DataFrame(acc)
```

Out[61]:

| | _id | count |
|-----|-----------------|-------|
| 0 | Earring | 2459 |
| 1 | Cigarette | 961 |
| 2 | Hot Lipstick | 696 |
| 3 | Purple Lipstick | 655 |
| 4 | Mole | 644 |
| ... | ... | ... |
| 82 | Orange Side | 68 |
| 83 | Tiara | 55 |
| 84 | Pilot Helmet | 54 |
| 85 | Choker | 48 |
| 86 | Beanie | 44 |

87 rows × 2 columns

Para la siguiente consulta busqué la cantidad de punks que tienen cada cantidad de accesorios. Como podemos ver Hay 333 punks que tienen un solo accesorio, la mayoría de los punks tienen 3 accesorios y solo 1 tiene 7 accesorios.

```
In [62]: cant_acc = db.cryptopunks.aggregate( [ { "$unwind": { "path": "$accessories" } },
                                              { "$group": { "_id": { "punk_id": "$punk_id", "accessories": "$accessories" } } },
                                              { "$group": { "_id": "$_id.punk_id", "count_accessories": { "$sum": 1 } } },
                                              { "$group": { "_id": "$count_accessories", "count": { "$sum": 1 } } },
                                              { "$sort": { "_id": 1 } } ] )
df_cant_acc = pd.DataFrame(cant_acc)
```

Out[62]:

| | _id | count |
|---|-----|-------|
| 0 | 1 | 333 |
| 1 | 2 | 3560 |
| 2 | 3 | 4501 |
| 3 | 4 | 1420 |
| 4 | 5 | 166 |
| 5 | 6 | 11 |
| 6 | 7 | 1 |

Realicé una última consulta para saber cuál es el punk con 7 accesorios y es el número 8348.

```
In [78]: acc7 = db.cryptopunks.aggregate([{"$project": { "punk_id": 1, "count": { "$size": "$accessories" } } },
                                         { "$sort": { "count": -1 } },
                                         { "$limit": 1 } ])
df_acc7 = pd.DataFrame(acc7)
```

Out[78]:

| | _id | punk_id | count |
|---|--------------------------|---------|-------|
| 0 | 644a417429549fe3db9e44c1 | 8348 | 7 |



Transacciones

A continuación, hice una consulta para saber los tipos de transacciones que contiene la colección. De aquí sale que hay 9 tipos distintos de transacciones: Unwrap, Wrap, Bid, Bid Withdrawn, Claimed, Offer Withdrawn, Offered, Sold, Transfer.

```
In [25]: txn_type = db.cryptopunks.distinct("txn_type")
df_txn_type = pd.DataFrame(txn_type)
```

```
Out[25]:
0
0      (Unwrap)
1      (Wrap)
2        Bid
3   Bid Withdrawn
4     Claimed
5 Offer Withdrawn
6     Offered
7       Sold
8     Transfer
```

A su vez, entre las transacciones, la que más se realizó fueron ofertas, y podemos ver que hubo 18.979 ventas, lo cual compone el 11% de las transacciones documentadas en la colección.

```
In [63]: count_txn = db.cryptopunks.aggregate( [ { "$group": { "_id": "$txn_type", "count": { "$sum": 1 } } },
                                              { "$sort": { "count": -1 } } ] )
df_count_txn = pd.DataFrame(count_txn)
```

```
Out[63]:
_id  count
0    Offered  65467
1        Bid  26808
2       Sold  18979
3  Bid Withdrawn  16342
4    Transfer  14225
5 Offer Withdrawn  10372
6     Claimed  10000
7      (Wrap)  2773
8    (Unwrap)  2526
```

Ventas

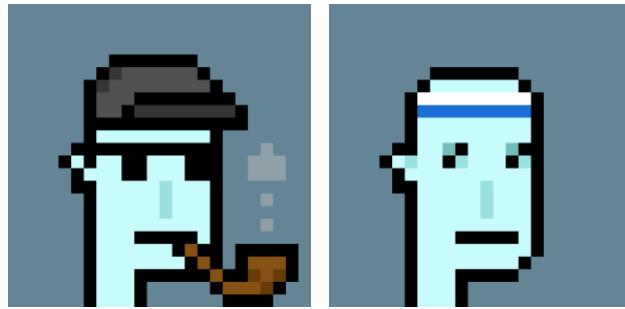
Para la siguiente consulta quise ver las 10 ventas más caras de todo el periodo. Para eso busco los documentos cuya transacción sea “Sold”, ordeno por el campo “eth” de forma descendente y limito la salida a 10 filas para que solo me muestre las 10 ventas más caras.

```
In [41]: top10_ventas = db.cryptopunks.find({"txn_type": "Sold"}, {"from_wallet_address": 0, "to_wallet_address": 0, "source": 0}).sort("eth", -1).limit(10)
df_top10_ventas = pd.DataFrame(top10_ventas)
```

Out[41]:

| | _id | txn_type | from | to | date | timestamp | eth | punk_id | type | accessories |
|---|--------------------------|----------|-------------|-------------|------------|---------------------|--------|---------|----------|-----------------------------------|
| 0 | 644a417429549fe3db9e20be | Sold | 0x03911f | 0xf4b4a5 | 2021-03-11 | 2021-03-11T00:00:00 | 4200.0 | 7804 | [Alien] | [Cap Forward, Pipe, Small Shades] |
| 1 | 644a417329549fe3db9cd57 | Sold | 0x6611fe | 0x7b8961 | 2021-03-11 | 2021-03-11T00:00:00 | 4200.0 | 3100 | [Alien] | [Headband] |
| 2 | 644a417329549fe3db9d6a1e | Sold | Snowfro | 0x8f7ce8 | 2021-07-30 | 2021-07-30T00:00:00 | 2250.0 | 5217 | [Ape] | [Gold Chain, Knitted Cap] |
| 3 | 644a417429549fe3db9e6a59 | Sold | 0x389747 | 0xf2a70b | 2021-09-11 | 2021-09-11T00:00:00 | 2000.0 | 8857 | [Zombie] | [Wild Hair, 3D Glasses] |
| 4 | 644a417429549fe3db9d9fdb | Sold | 0x9fce8e | 0x956d14 | 2021-08-24 | 2021-08-24T00:00:00 | 1600.0 | 7252 | [Zombie] | [Chinstrap, Earring, Crazy Hair] |
| 5 | 644a417329549fe3db9c9819 | Sold | grunare... | gary.vee... | 2021-07-30 | 2021-07-30T00:00:00 | 1600.0 | 2140 | [Ape] | [Knitted Cap, Small Shades] |
| 6 | 644a417329549fe3db9ca54d | Sold | 0xfab9b8 | 0x7cd3fc | 2021-08-06 | 2021-08-06T00:00:00 | 1500.0 | 2338 | [Zombie] | [Mohawk Thin] |
| 7 | 644a417329549fe3db9db0d4 | Sold | Pranksy | 0xd2c1a0 | 2021-09-04 | 2021-09-04T00:00:00 | 1320.0 | 6275 | [Zombie] | [Shadow Beard, Mohawk Dark] |
| 8 | 644a417429549fe3db9df9e0 | Sold | thebeaut... | 0x9fce8e | 2021-08-04 | 2021-08-04T00:00:00 | 1000.0 | 7252 | [Zombie] | [Chinstrap, Earring, Crazy Hair] |
| 9 | 644a417329549fe3db9db0db | Sold | 0xaaff336 | Pranksy | 2021-09-04 | 2021-09-04T00:00:00 | 1000.0 | 6275 | [Zombie] | [Shadow Beard, Mohawk Dark] |

Como podemos ver en el resultado, las dos ventas más caras se realizaron el mismo día y fueron por 4.200 eth, que hoy en día equivale a cerca de 7.175.000 euros con el precio de Ethereum estando a 1.708,3 euros por Ether. Estos dos cryptopunks son el tipo “Alien”, el más raro de todos. Buscando los ids podemos saber que estas ventas corresponden a los siguientes punks:



A su vez realicé consultas para saber cuántos punks de cada tipo se vendieron. Se puede ver que se vendieron 7 de los 9 punks que son aliens, tiene sentido que la demanda por este tipo sea alta ya que son el tipo más raro de todos y por ende los más valiosos como colecciónables.

```
In [43]: tipo_ventas = db.cryptopunks.aggregate( [ { "$unwind": "$type" },
{ "$match": { "txn_type": "Sold" } },
{ "$group": { "_id": { "punk_id": "$punk_id", "type": "$type" } },
{ "$group": { "_id": "$_id.type", "count": { "$sum": 1 } } },
{ "$sort": { "count": -1 } } ] )
df_tipo_ventas = pd.DataFrame(tipo_ventas)
```

Out[43]:

| _id | count |
|----------|-------|
| 0 Male | 4080 |
| 1 Female | 2242 |
| 2 Zombie | 47 |
| 3 Ape | 17 |
| 4 Alien | 7 |

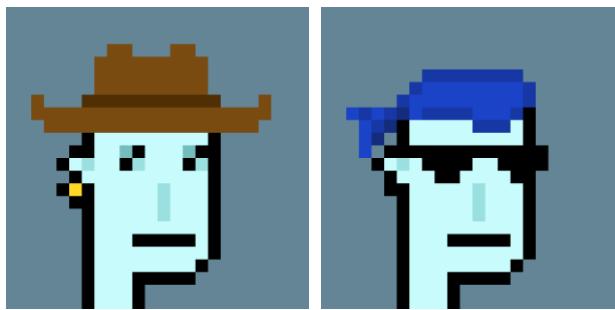
En la siguiente consulta busco determinar cuáles son los dos aliens que nunca se vendieron. Para hacer esto creé un nuevo campo llamado “sold” que toma el valor 1 si el campo “txn_type” es igual a “Sold”, y el valor 0 de lo contrario. A partir de ahí hice una consulta con aggregate para filtrar solo los aliens y agrupar por punk id para sumar los valores del campo “sold” que creé anteriormente.

```
In [77]: db.cryptopunks.update_many({"tx_type": "Sold"}, {"$set": {"sold": 1}})
db.cryptopunks.update_many({"tx_type": "Sne": "Sold"}, {"$set": {"sold": 0}})
alien_ventas = db.cryptopunks.aggregate([
    { "$match": { "type": "Alien" } },
    { "$group": { "_id": { "punk_id": "$punk_id" }, "ventas": { "$sum": "sold" } } },
    { "$sort": { "ventas": -1 } }
])
df_alien_ventas = pd.DataFrame(alien_ventas)
```

Out[77]:

| | _id | ventas |
|---|-------------------|--------|
| 0 | {'punk_id': 3100} | 2 |
| 1 | {'punk_id': 7804} | 2 |
| 2 | {'punk_id': 2890} | 2 |
| 3 | {'punk_id': 7523} | 1 |
| 4 | {'punk_id': 5905} | 1 |
| 5 | {'punk_id': 5822} | 1 |
| 6 | {'punk_id': 6089} | 1 |
| 7 | {'punk_id': 3443} | 0 |
| 8 | {'punk_id': 635} | 0 |

De esta consulta podemos ver que hay 3 punks aliens que se vendieron 2 veces y cuatro se vendieron solo una vez. Los que no se vendieron son los que tienen id 3443 y 635.



Conclusiones

Del ejercicio anterior se concluye que hay 10.000 Cryptopunks únicos. Hay cinco tipos de Cryptopunks: Male, Female, Zombie, Ape y Alien. De estos tipos el más común es Male, siendo que 6.039 Cryptopunks pertenecen a este tipo, y el más raro es Alien, ya que solo 7 Cryptopunks son aliens. A su vez, hay 87 accesorios distintos, el que más se repite es el “Earring” y hay 2.459 punks que lo tienen. El segundo que más se repite es el “Cigarette” con 961 punks que lo tienen. Entre los menos comunes están el “Choker” con 48 punks y el “Beanie” con 44 punks que lo tienen. Además, hay 8 punks que no tienen ningún accesorio. Por otro lado, los punks pueden tener varios accesorios al mismo tiempo, y hay 333 punks que tienen un solo accesorio, la mayoría tienen 3 accesorios y solo 1 tiene 7 accesorios.

En cuanto a las transacciones que están en la colección, hay 9 tipos distintos de transacciones: Unwrap, Wrap, Bid, Bid Withdrawn, Claimed, Offer Withdrawn, Offered, Sold, Transfer. La transacción que más se realizó fue la oferta, y hubo 18.979 ventas, con lo cual el 11% de las transacciones documentadas en la colección fueron ventas. Además, la primera transacción documentada en la colección se realizó el 23 de junio de 2017 a las 00:00, y fue una venta del punk 0 por 0,98 Ether. Por otro lado, la última transacción en la colección se realizó el 7 de octubre de 2021 y fue una oferta por el punk número 9039 por 288.88 Ether.

Finalmente, del análisis de las ventas, se vio que dentro del top 10 de las ventas más caras, las dos de mayor valor se vendieron por el mismo monto, 4.200 Ether. Estos dos punks son aliens, y revisando los cuántos punks se vendieron de cada tipo, se puede ver que 7 de los 9 aliens se vendieron.