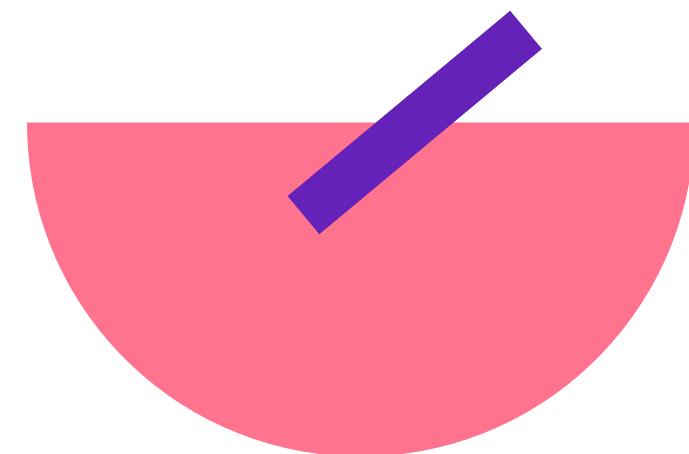




Анализ рынка nft

+ web scrapping Steam

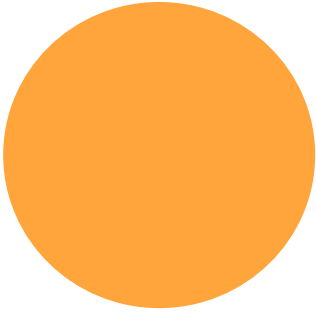
Выполнили:
Степашкина В.П.
Фортов Е.К.
СВИТКОВСКИЙ Я.М.





Что такое NFT?

«NFT-токен (невзаимозаменяемый токен) — уникальный цифровой сертификат, который хранится в блокчейне, гарантирует оригинальность предмета и даёт эксклюзивные права на него».

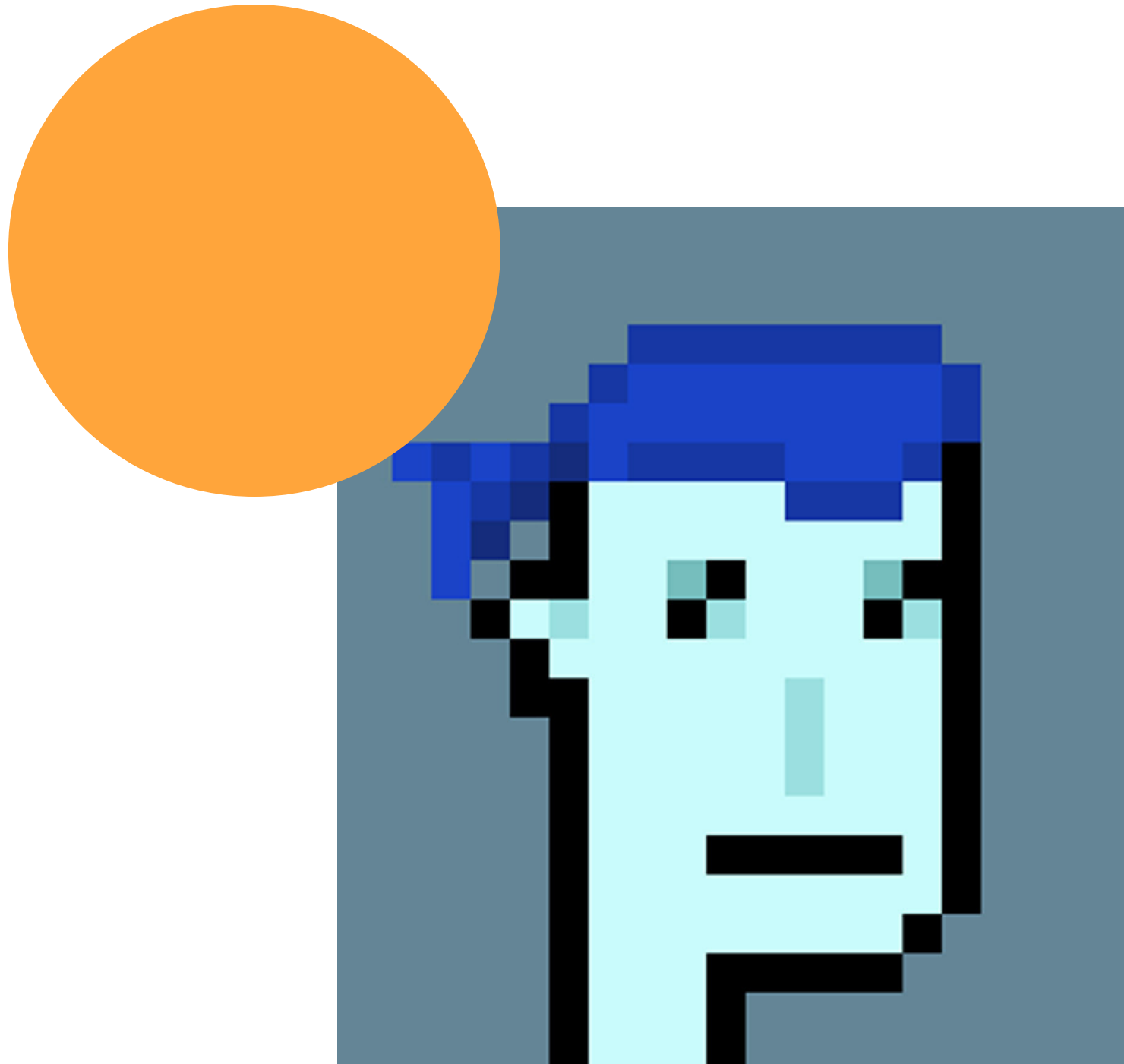


Гипотезы

- 1) Люди чаще покупают предметы искусства, коллекционные и цифровые предметы.
- 2) Объем продаж определенной коллекцией прямо пропорционален доле владения продуктов в ней
- 3) Чем больше кол-во предметов, тем меньше минимальная цена
- 4) Art, Collectibles and Digital - самые популярные виды NFT объектов на рынке NFT, так как именно их очень много в начале колонки "Category"
- 5) В среднем, чем короче название коллекции, тем больше средняя стоимость всех итемов в ней

Переменные

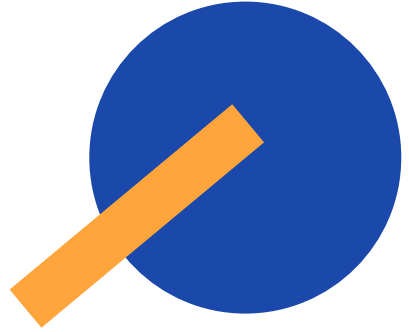
Name: название коллекции NFT.
Volume: Объем продаж коллекции NFT (SOL).
Volume_USD: объем продаж коллекции NFT (USD).
Market_Cap: рыночная капитализация (SOL).
Market_Cap_USD: рыночная капитализация (USD).
Sales: количество продаж товаров из коллекции NFT.
Floor_Price: самая низкая цена NFT в коллекции (SOL).
Floor_Price_USD: самая низкая цена NFT в коллекции (USD).
Average_Price: средняя цена NFT в коллекции (SOL).
Average_Price_USD: средняя цена NFT в коллекции (USD).
Owners: количество владельцев NFT в коллекции.
Assets: количество айтемов в коллекции.
Owner_Asset_Ratio: процент владения предметами в коллекции.
Category: Категории коллекции NFT.
Website: веб-сайт коллекции NFT.
Logo: логотип коллекции NFT.



Преобразования

- Изменение типа данных
- Разбиение переменной
- Добавление нового столбца, отвечающего за новую категориальную переменную
- Срез + сортировки





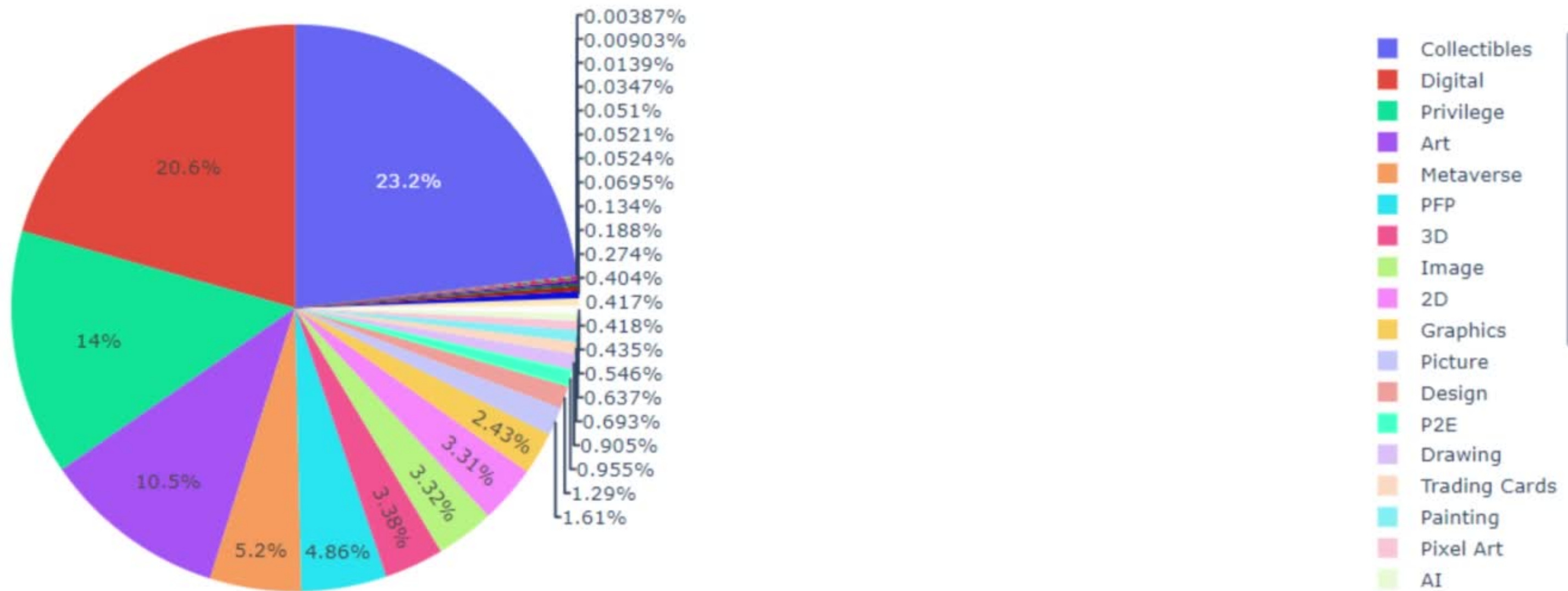
Категориальный процент владения



Вывод:

По данной таблице мы можем заметить, что большая часть коллекций не выкуплена. даже на 60%. Что говорит нам о редких случаях большого спроса на проекты.

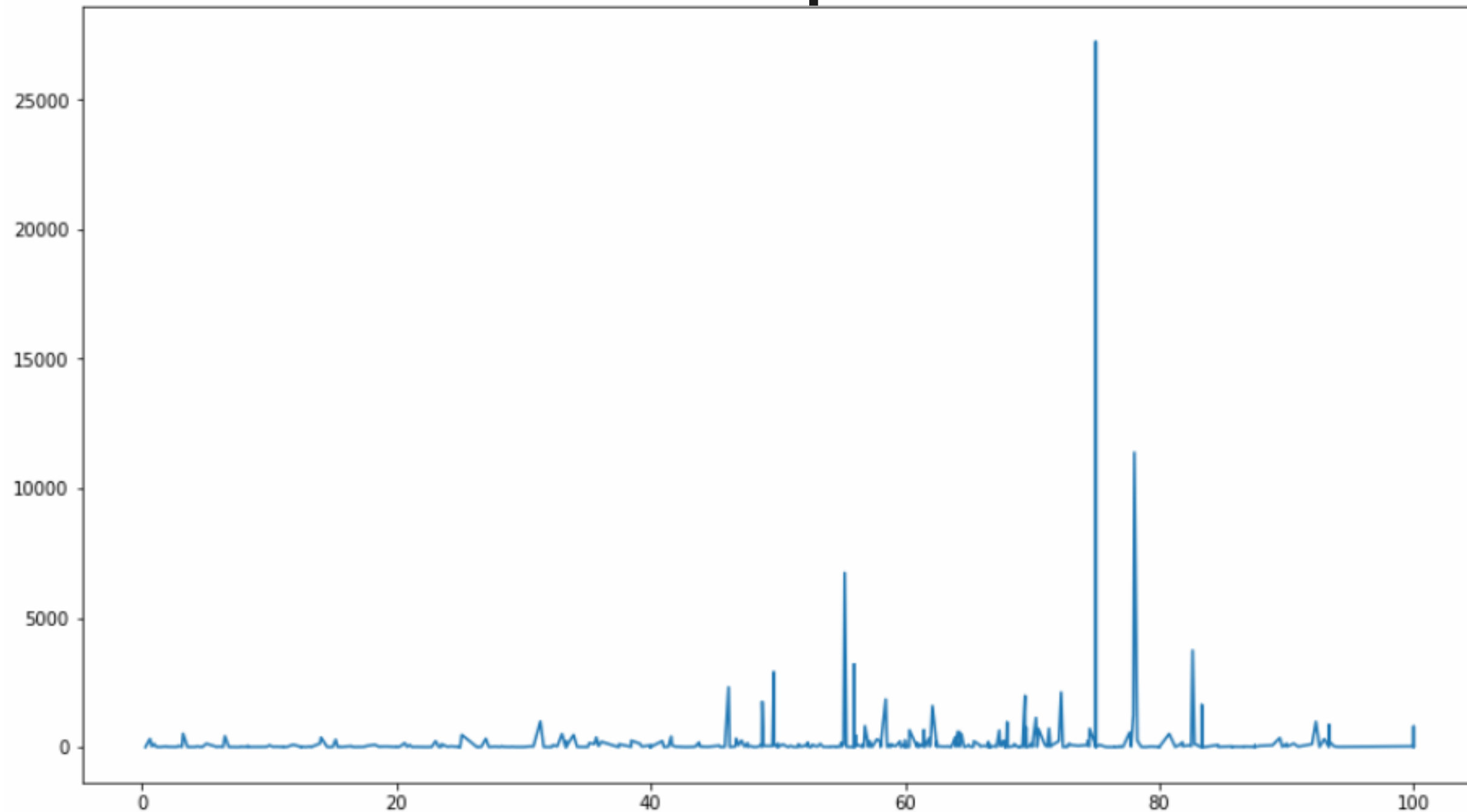
Категории NFT по объему



Вывод:

Здесь мы видим, что объем продаж имеет максимальные значения по следующим категориям: коллекционные, цифровые, искусство и дающие привелегии NFT. Дальнейшие подгруппы имеют намного меньший процент объема продаж по всему рынку(менее 5.3 процента). А эти "топ-4" категории занимают 68,3% объема всех продаж.

Зависимость объема рынка от процента владения коллекцией

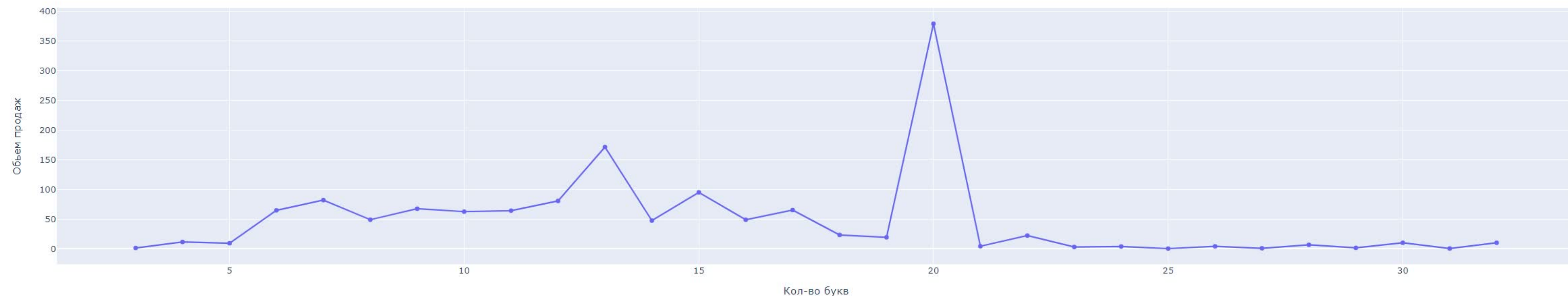


Вывод:

По графику можно сделать вывод, что напрямую эти две переменные никак не связаны, так как большие показатели по объему продаж показывают коллекции не с максимальным количеством владения, а с 50% по 80%. А после 80% объем продаж идет на спад с несколькими исключениями.

Зависимость между кол-вом букв и продажами

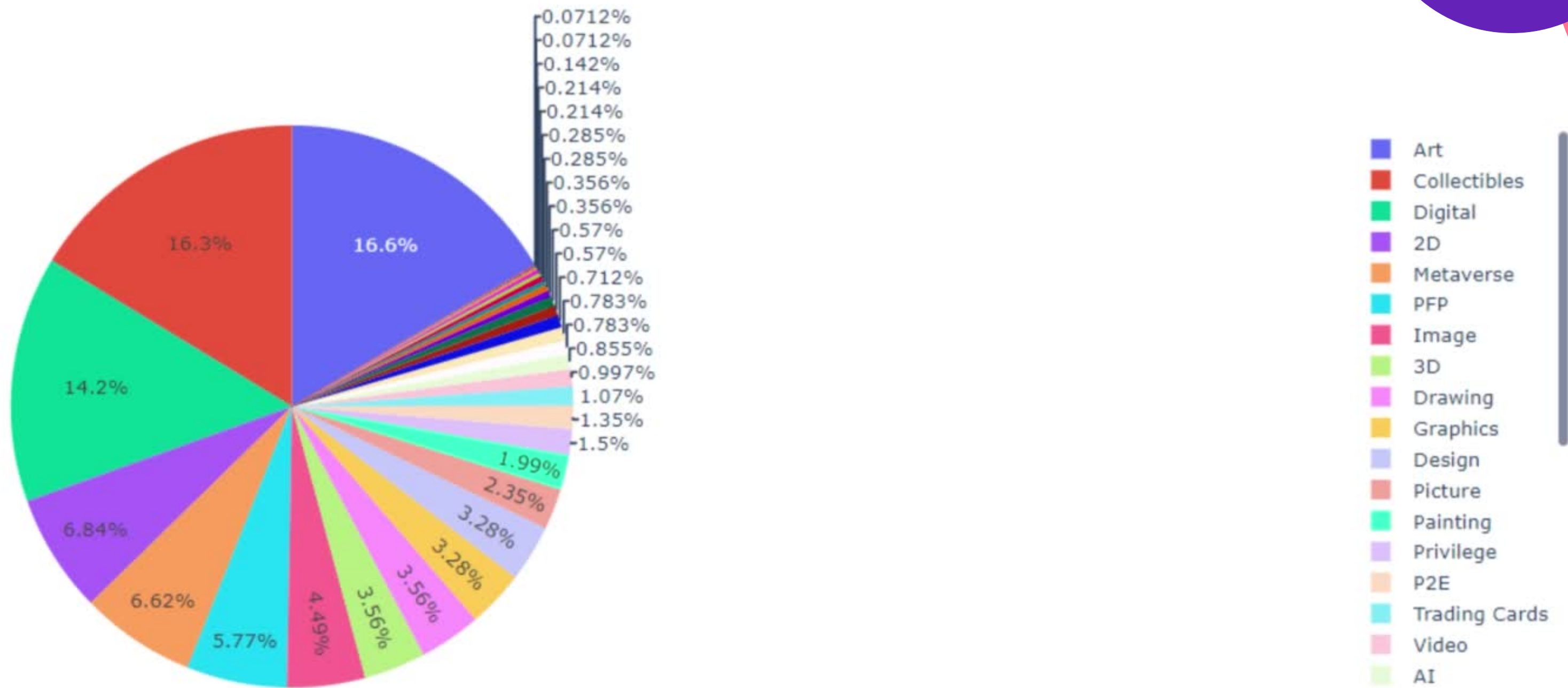
Зависимость между количеством букв и продажами



Вывод:

Заметим, что показатели всех промежутков до 20 букв идут равномерно, после небольшого спада с 19 на 21 букву продолжают идти аналогично. При этом есть большой скачок при 20 буквах в названии.

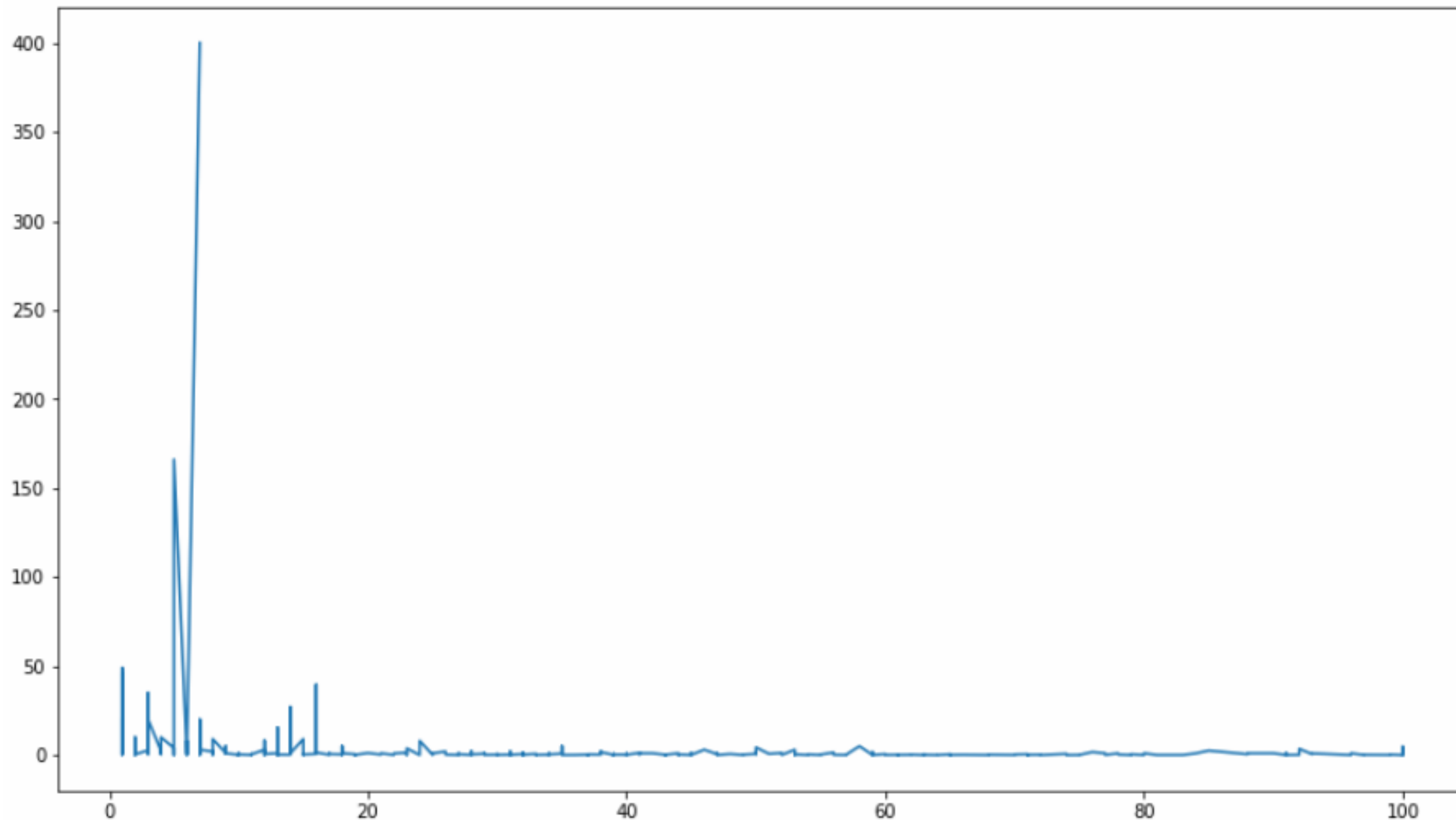
Количество NFT по категориям



Вывод:

Здесь видны 3 самые популярные категории NFT: искусство, коллекционные, цифровые NFT. Они занимают 47,1% рынка, то есть почти половину всех NFT существующих на данной платформе. А последующие категории занимают менее 6,9%.

Зависимость минимальной цены NFT от количества предметов в коллекции



Вывод:

Мы можем наблюдать зависимость коллекций, в которых кол-во продуктов до 17, за исключением некоторых проектов, которая представляет собой параболу.

Описательная статистика

	Index	Volume	Volume_USD	Market_Cap	Market_Cap_USD	Sales	Floor_Price	Floor_Price_USD	Average_Price	Average_Price_USD	Owners	Assets	Owner_Asset_Ratio
count	592.000000	592.000000	5.920000e+02	592.000000	592.000000	592.000000	544.000000	544.000000	592.000000	592.000000	592.000000	592.000000	543.000000
mean	296.547297	197.548885	2.900413e+04	45.134817	6626.466216	164.491554	2.423015	355.747019	2.373566	348.486991	53.731419	116.694257	54.385672
std	172.520622	1283.031435	1.883747e+05	179.327166	26328.741804	451.321395	18.982148	2786.959001	15.035229	2207.472329	133.893752	251.598765	27.632770
min	0.000000	0.000000	0.000000e+00	0.000000	0.000000	0.000000	0.020000	2.936400	0.000000	0.000000	0.000000	0.000000	0.260000
25%	147.750000	7.300000	1.071786e+03	0.000000	0.000000	11.000000	0.110000	16.150200	0.306029	44.931155	4.000000	10.000000	33.330000
50%	295.500000	18.595000	2.730118e+03	0.000000	0.000000	35.000000	0.300000	44.046000	0.720412	105.770928	15.000000	38.000000	57.250000
75%	444.250000	80.825000	1.186673e+04	21.841974	3206.750000	107.250000	1.000000	146.820000	1.596750	234.434771	44.000000	102.250000	72.845000
max	599.000000	27256.630000	4.001818e+06	2530.877143	371583.000000	5229.000000	400.000000	58728.000000	347.500000	51019.950000	2041.000000	3320.000000	100.000000

Выводы:

- 1) Можно заметить очень большую разницу в объеме, что 75% рынка меньше 80 SOL, а максимальная составляет 27000 SOL, что говорит нам о негармоничном ценообразовании(трудно поддается расчётам).
- 2) Видно, что 50% коллекций NFT имеют 0 рыночную капитализацию, что говорит нам о нерентабельности данного товара.
- 3) 50% NFT стоят менее 45\$.
- 4) Не смотря на нестабильные распределения показателей, Owner_Asset_Ratio имеет правильное распределение с точки зрения статистики.

Общие выводы:

- Чаще всего люди покупают следующие категории NFT: искусство, коллекционные и цифровые. Самый большой объем продаж занимают следующие категории: коллекционные, цифровые и дающие привилегии.
- Объем продаж никак не зависит от процента владения коллекцией. А максимальное значение достигается при 75% владения.
- Зависимость минимальной цены и количества была обнаружена на маленьких значениях в виде параболы. На больших значениях её нет.
- Между количеством букв в названии и средней ценой NFT коллекции - нет зависимости. А максимум достигается при 20 буквах в названии.
- Видна гигантская разница в объеме, что 75% рынка меньше 80 SOL, а максимальная составляет 27000 SOL, что говорит нам о негармоничном ценообразовании(трудно поддается расчетам)
- Несмотря на нестабильные распределения показателей, Owner_Asset_Ratio(процент владения коллекцией) имеет правильное распределение с точки зрения статистики.
- Полученные показатели мало нам о чем говорят в связи с тем, что мы рассматриваем NFT, которые по сути основаны на искусстве и привилегиях, на что играют большую роль рекламные кампании и современные тенденции, нежели другие параметры. Даже те корреляции, которые были выявлены в жизни себя не могут оправдать, тк имеют слишком сильный разброс параметров.



Веб-скрапинг bestseller-ов в сфере игровой IT индустрии.

Сайт со списком популярных игр:

[https://store.steampowered.com/search/?
sort_by=Price_DESC&filter=topsellers](https://store.steampowered.com/search/?sort_by=Price_DESC&filter=topsellers)

Из него получены ссылки на соответствующие страницы игр.

Количество анализируемых игр - 622.



Проблемы, возникшие во время составления web-scraper-a

- Некоторые url адреса, полученные с анализируемой страницы, редиректились на рекламные страницы, что затрудняло процесс извлечения данных (все параметры были NaN или аварийное завершение программы). Исправлено: составлен "черный" список этих "лже-адресов", и их мы не анализируем.
- На одних страницах информация об ОЗУ приведена в ГБ(чаще всего), на других же - в МБ. Исправлено: МБ переводятся в ГБ.
- На разных страницах приводилось разное число параметров игры, в связи с чем шло обращение к недействительным адресам массива параметров (и программа падала). Исправлено: использование исключений и вставка значения NaN в соответствующий параметр.

Параметры dataset-a

- 1) Название
- 2) Жанр
- 3) Разработчик
- 4) Издатель
- 5) Серия
- 6) Дата
- 7) Цена базового пакета
- 8) Цена VIP-пакета
- 9) Средняя цена пакетов
- 10) Минимальная ОС
- 11) Минимальный процессор
- 12) Мин. кол-во ОЗУ(в ГБ)
- 13) Мин. видеокарта
- 14) Мин. место на диске
- 15) Минимальная ОС
- 16) Минимальный процессор
- 17) Мин. кол-во ОЗУ(в ГБ)
- 18) Мин. видеокарта
- 19) Мин. место на диске
- 20) Оценки
- 21) Вебсайт

#1 Частотный анализ
парметра minRAM(GB)

8.0	173
4.0	156
6.0	55
2.0	46
1.0	14
3.0	10
12.0	9
10.0	5
0.5	3
1.5	2
5.0	2
16.0	2
2.5	1

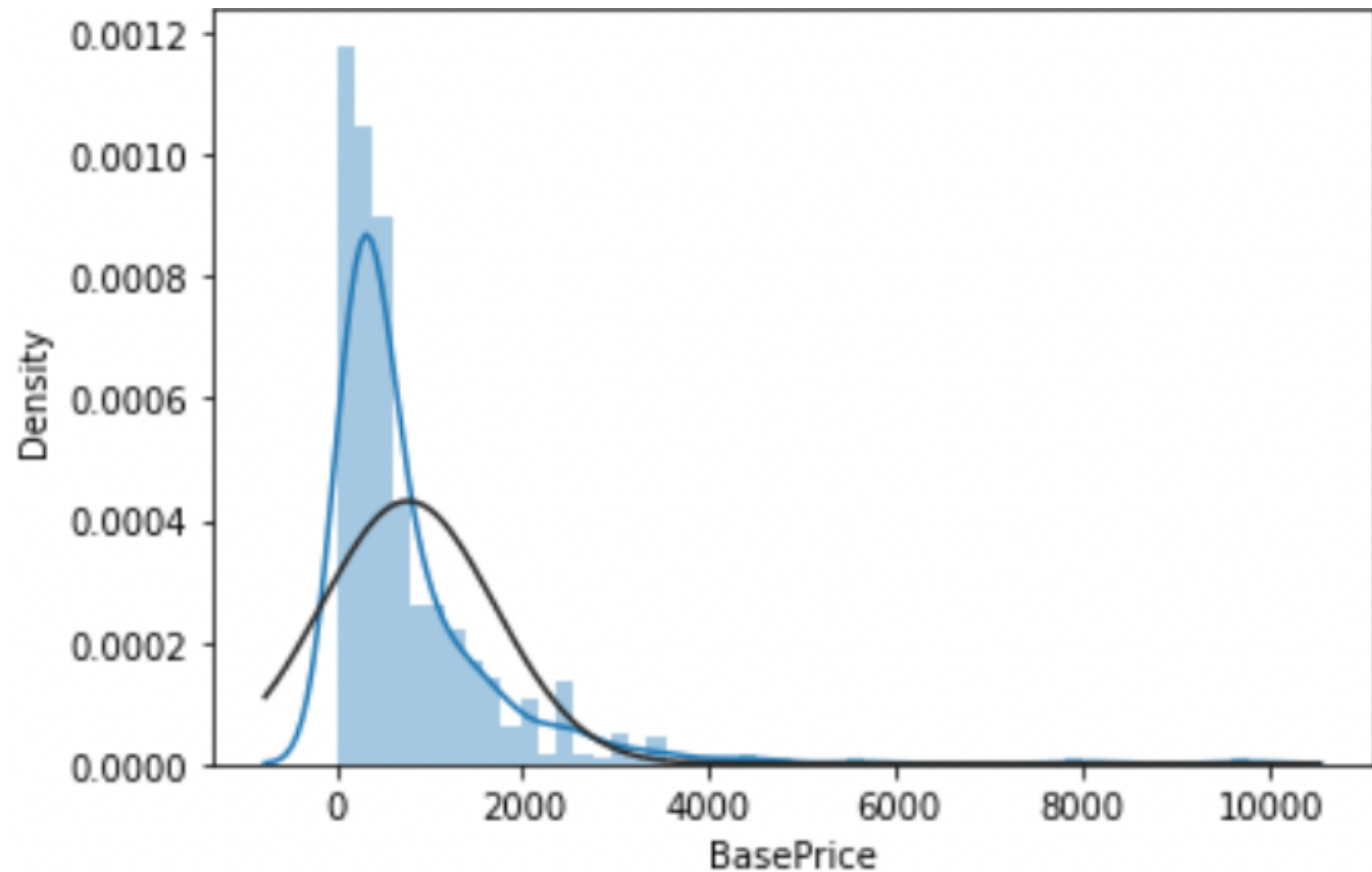
Вывод: 8 ГБ ОЗУ - самое популярное
требование разработчиков для
пользователей своей игры.

#2 Корреляционный анализ по всем
числовым параметрам датасета

	BasePrice	PremiumPrice	MiddlePrice	minRAM(GB)	minStorage(GB)
BasePrice	1.000000	0.174929	0.269885	0.149364	0.401000
PremiumPrice	0.174929	1.000000	0.984235	0.094604	0.073139
MiddlePrice	0.269885	0.984235	1.000000	0.136942	0.130587
minRAM(GB)	0.149364	0.094604	0.136942	1.000000	0.304166
minStorage(GB)	0.401000	0.073139	0.130587	0.304166	1.000000
recRAM(GB)	0.144032	0.112690	0.150334	0.790634	0.319228
recStorage(GB)	0.432266	0.149129	0.209882	0.335051	0.993334

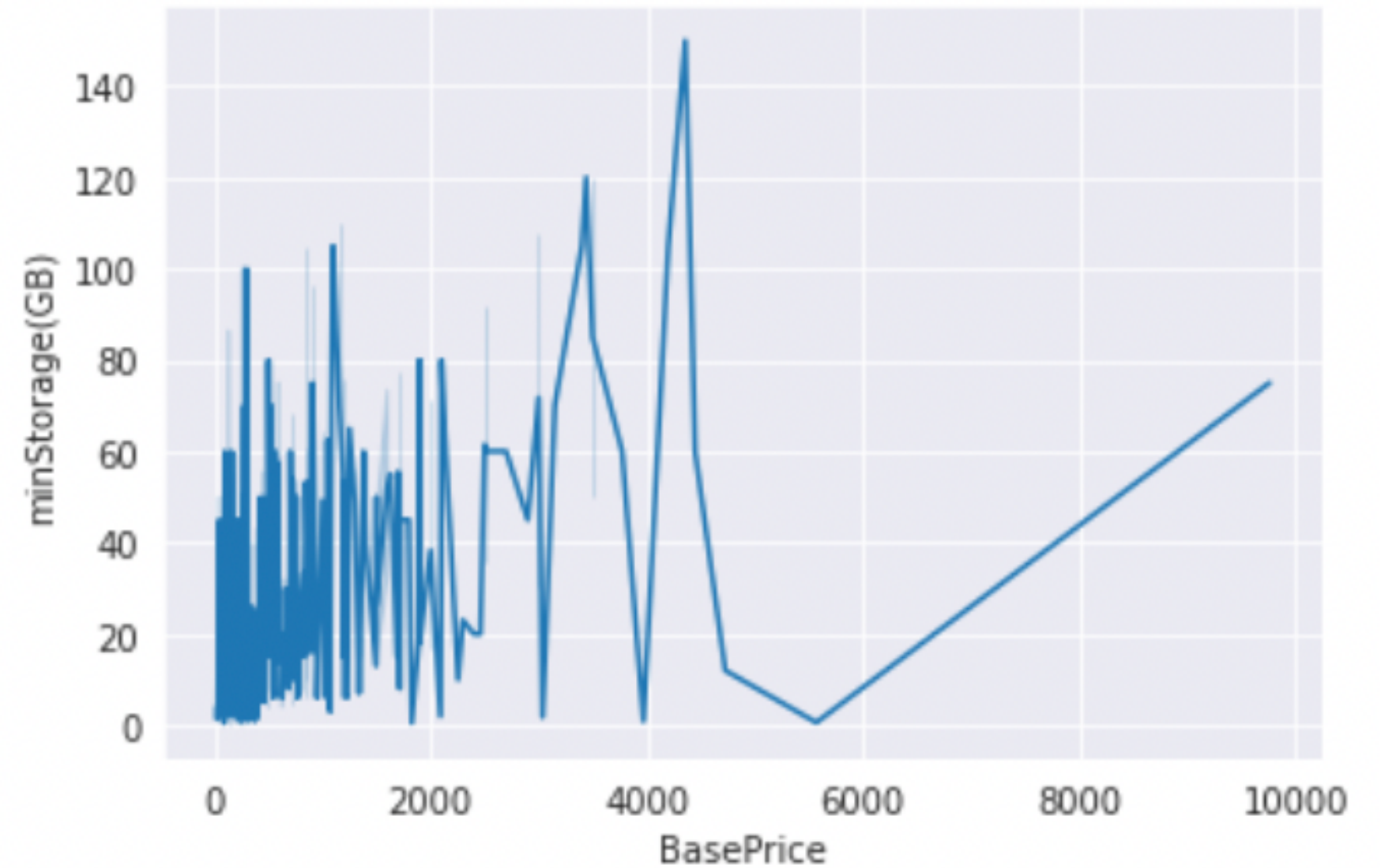
Вывод: минимальные и рекомендованные
требования по ГБ отлично коррелируют.

#3 Описательная статистика цен базовых пакетов приложений



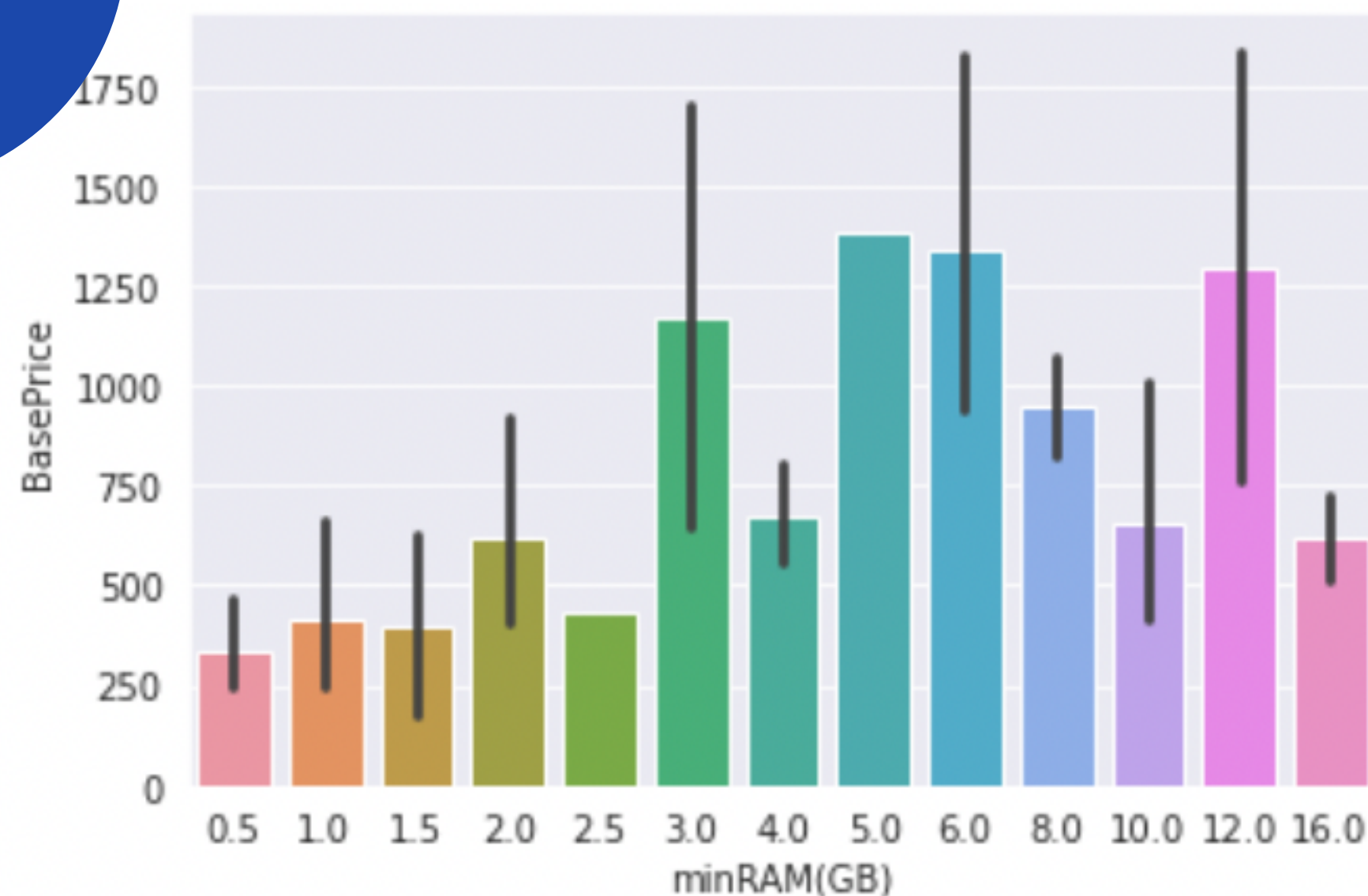
Вывод: большинство приложений стоят в пределах [0, 2000].

#4 Зависимость цены и размера игры

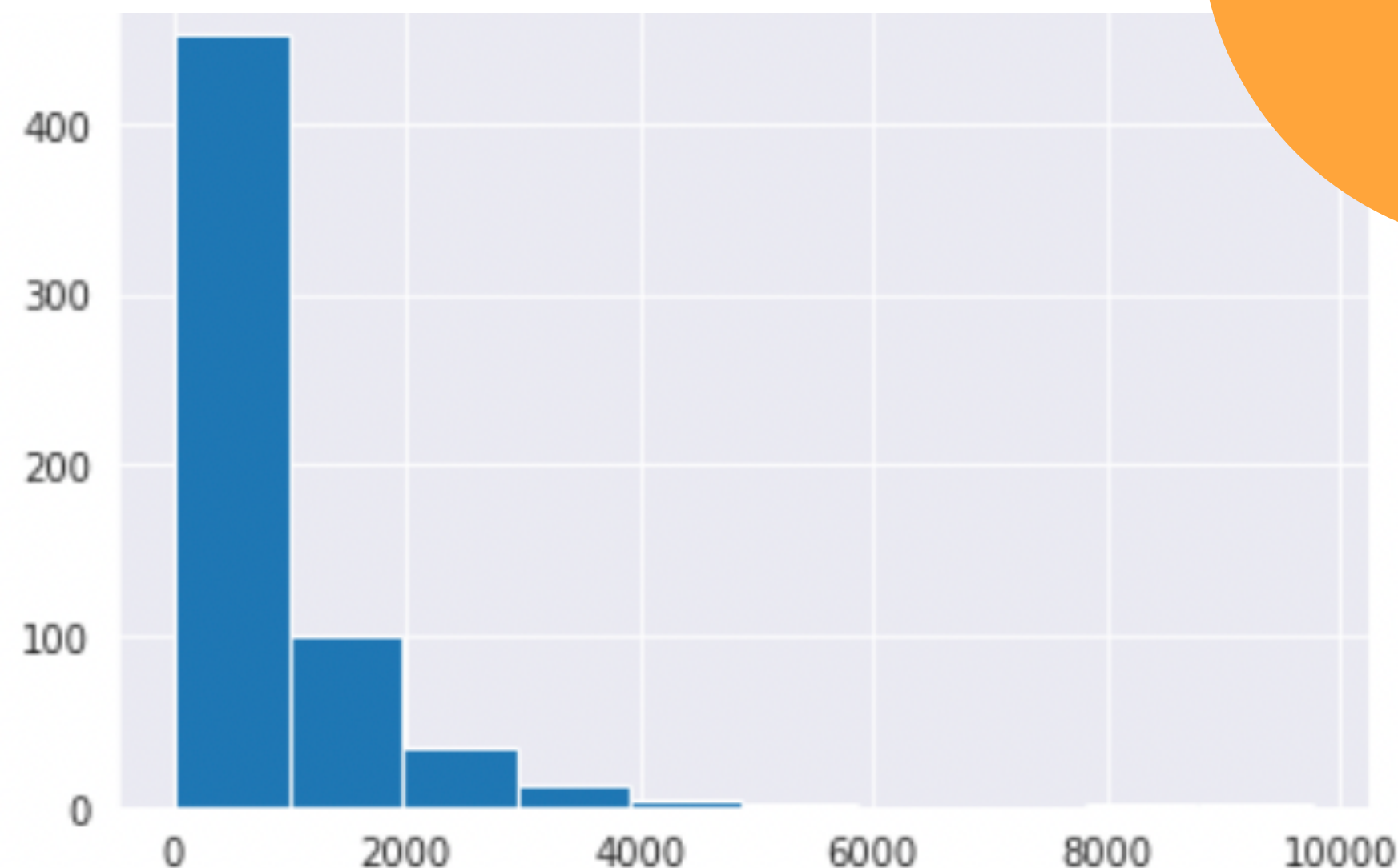


Вывод: таковой зависимости нет, т к волатильность графика очень большая.

#5 Сравнение цены игр с различными требованиями ОЗУ



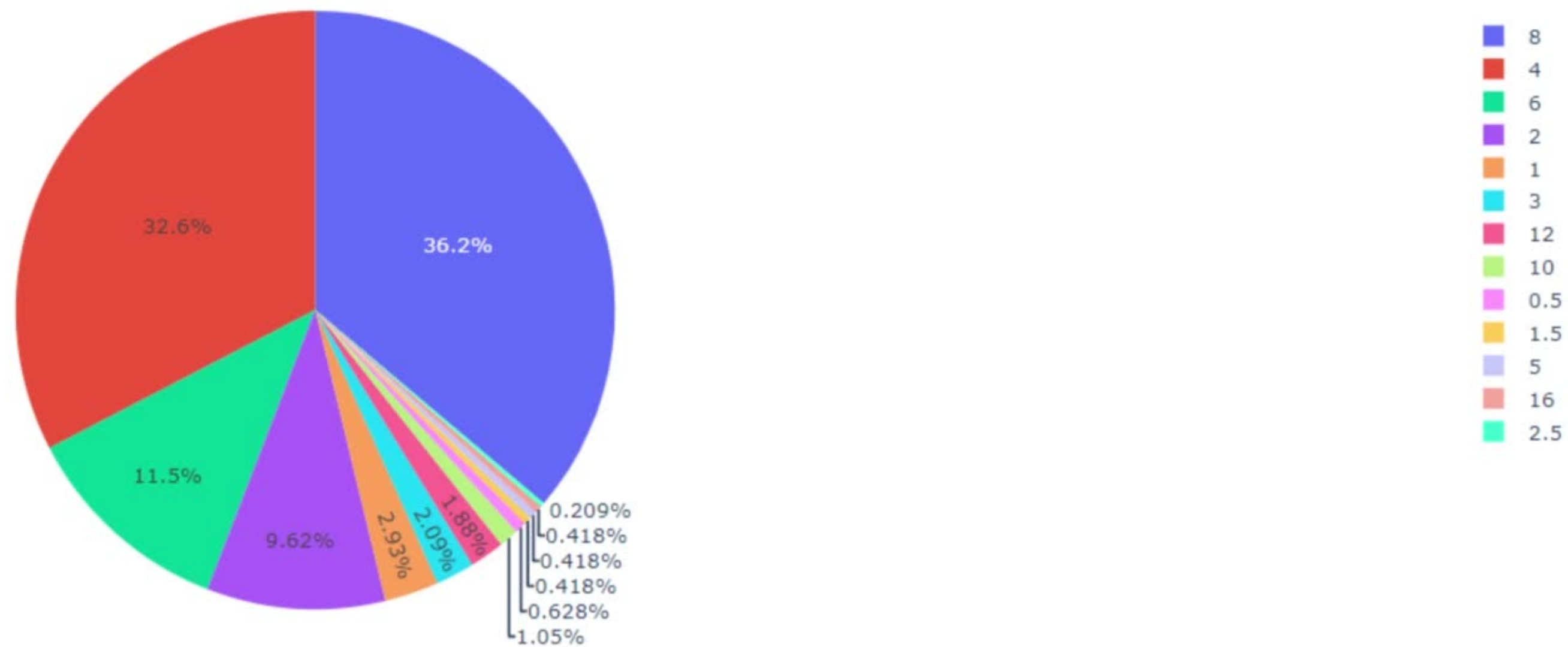
#6 Гистограмма мин цен игр



Вывод: чем меньше у тебя ОЗУ тем дешевле игра и, следовательно, тем более простую игру пользователь может себе позволить с т.з. "железа" своей машины.

Вывод: большинство приложений стоят в пределах [0, 2000].

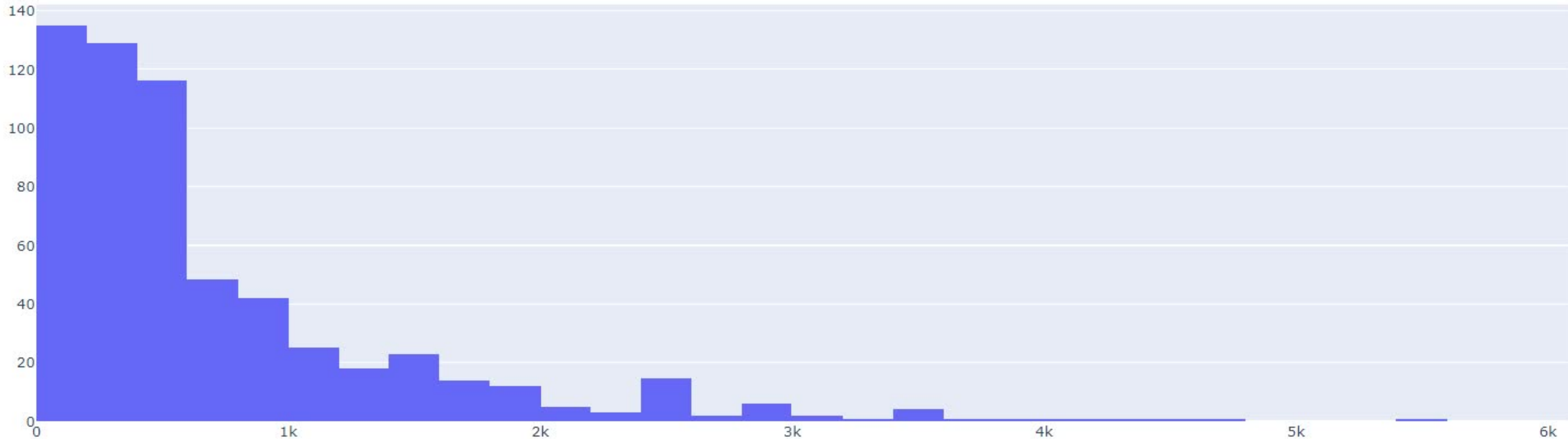
Круговая диаграмма по минимальному объему ОЗУ



Вывод:

4.0 и 8.0 ГБ ОЗУ по популярности преобладают перед другими размерами ОЗУ

Зависимость количества игр от цены



Вывод:

Количество игр с низкой стоимостью значительно преобладает над премиум сегментом.

Спасибо за десятку!

