

Прогнозирование цен на монеты

Команда

Ноль, целковый, полушка

Задача

- Изучение рынка исторических монет для поиска возможностей сбережения личного капитала
- Нахождение альтернативы фондовому рынку и банковским услугам

Задача

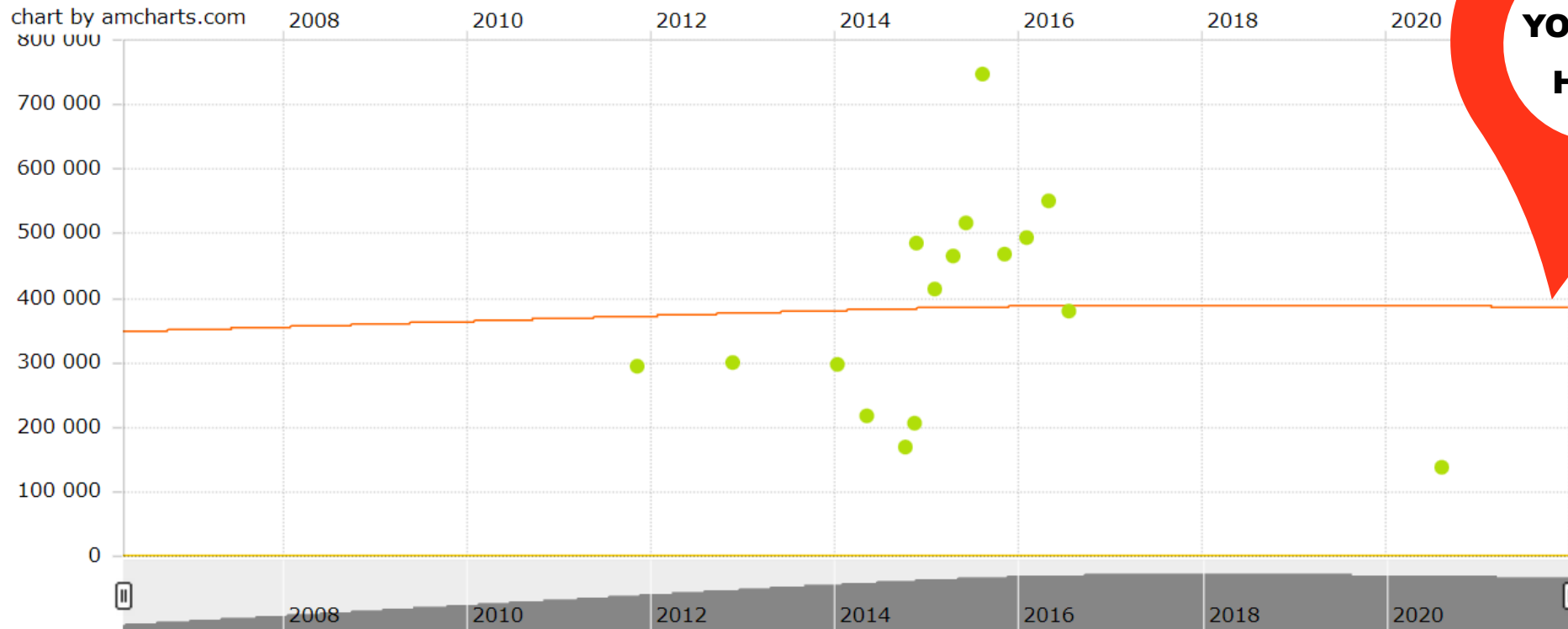
Никто:

Абсолютно никто:

Мы: анализ монет из гаража деда

Задача

Научиться оценивать рыночную стоимость монет,
для которых сделки на аукционах происходят в среднем реже,
чем раз в месяц



Data

Auction trades dataset

- `state` — сохранность
 - G — Good
 - VG — Very Good
 - F — Fine
 - VF — Very Fine
 - XF — Extra Fine
 - AU — Almost Uncirculated
 - UNC — Uncirculated
 - Proof — Идеальное
- `auction` — название аукциона
- `buyer` — ник покупателя
- `n_bid` — количество заявок на аукционе
- `price` — цена сделки (руб.)
- `date` — дата сделки
- `coin_name` — название монеты

info dataset

- `catalog_num` — каталожный номер
- `letters` — буквы на монете
- `mint` — монетный двор
- `circulation` — тираж
- `metal` — металл
- `weight` — вес монеты (г.)
- `pure_metal_content` — содержание металла в монете (г.)
- `diameter` — диаметр (мм.)
- `obverse` — надпись на лицевой стороне
- `reverse` — надпись на обратной стороне
- `coin_name` — название монеты



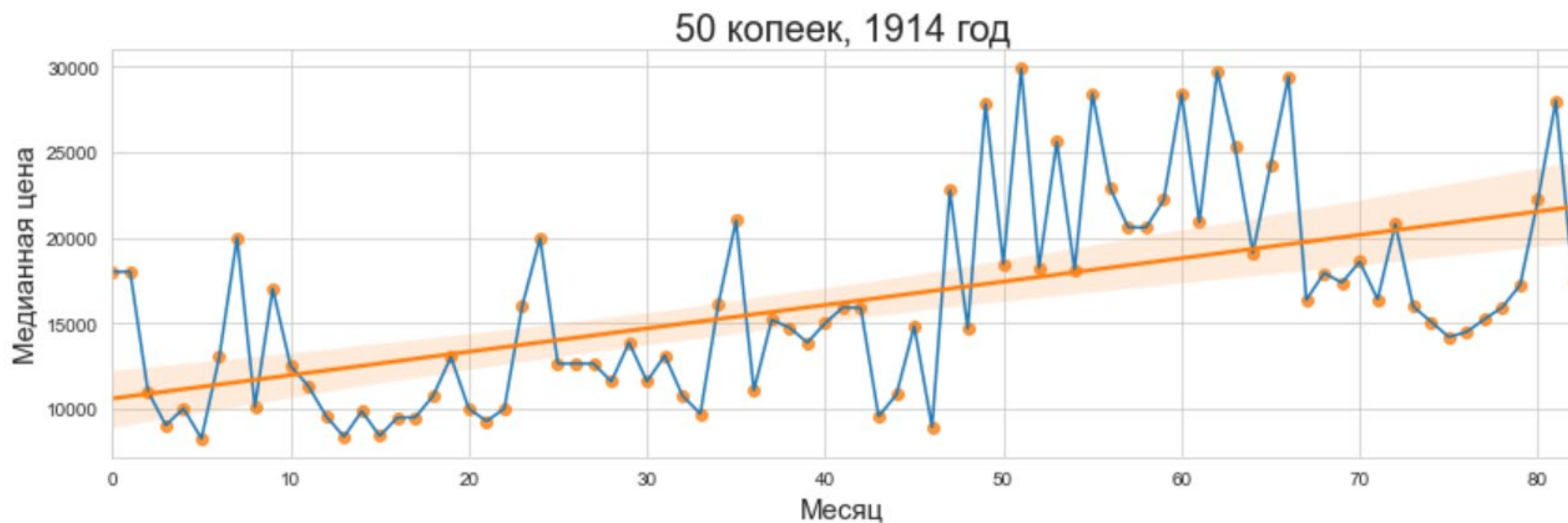
**монеты Николая II
(1894-1917)**

Нормирование

Зачем оно нужно?

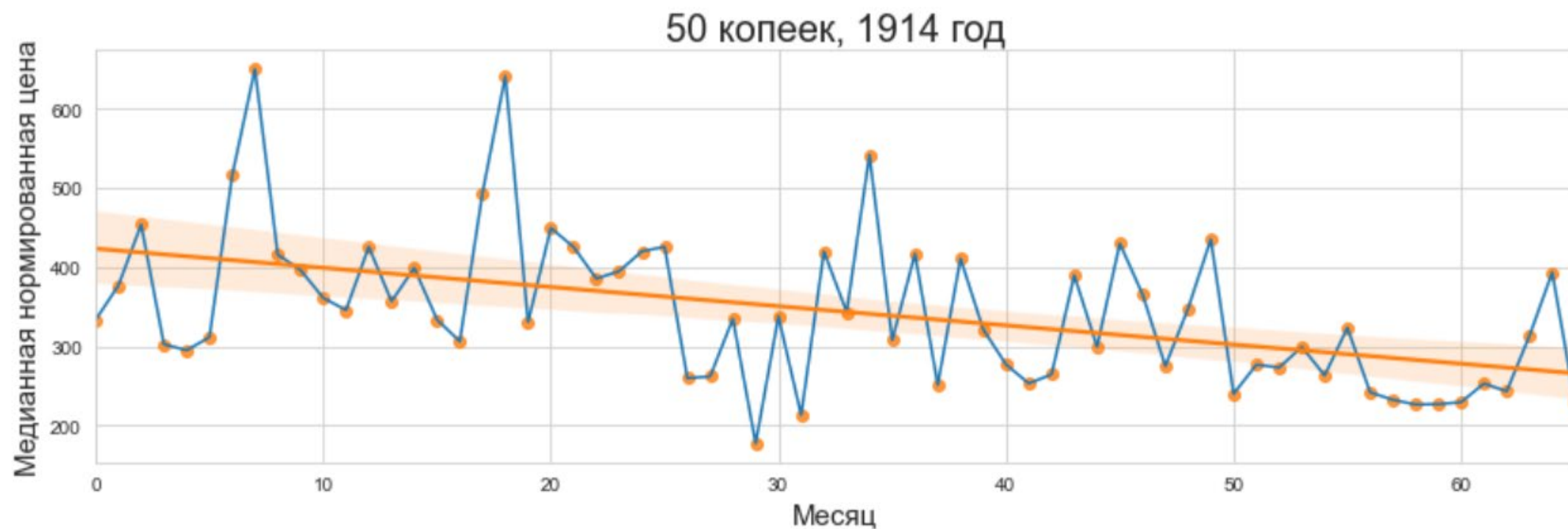
Ожидание

+150% за 7 лет, то есть 14% годовых



Реальность

-30% за 7 лет, то есть -5% годовых



Понятие коллекционной ценности

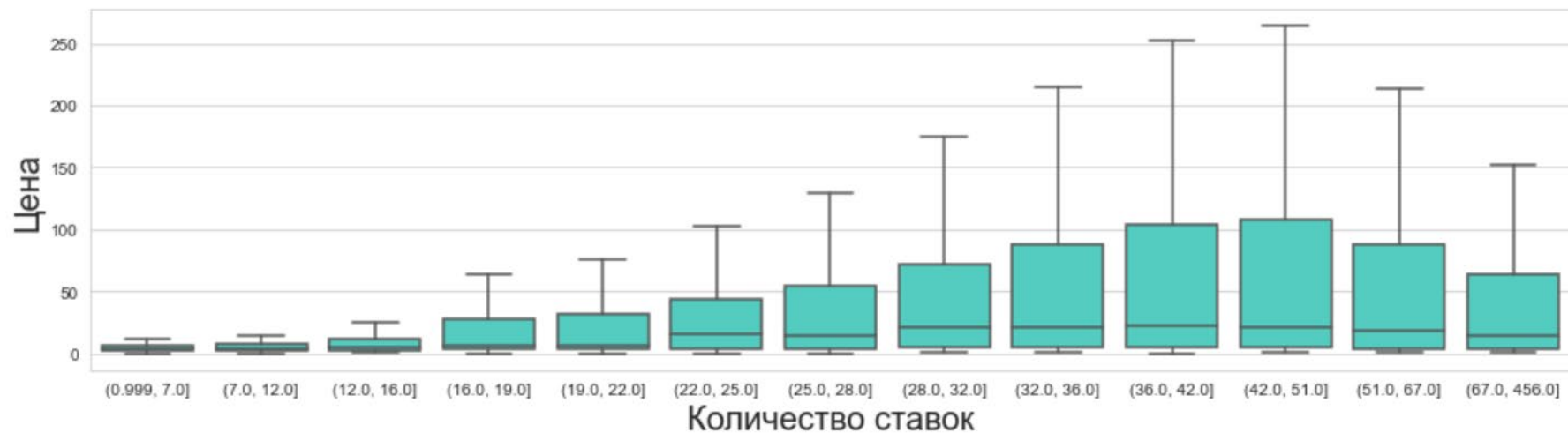
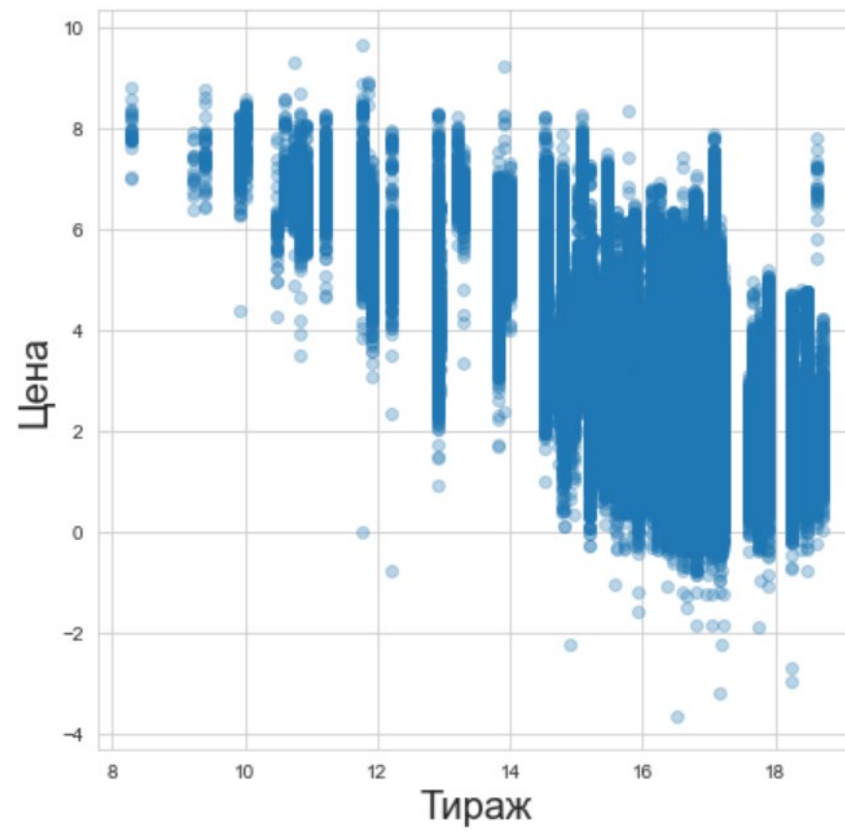
На цену монеты влияют внешние факторы:

- Валюта
- Инфляция
- Цена металла
- Сезонность

Если изъять из цены эти факторы, она станет более стационарной и легко предсказываемой

Анализ признаков

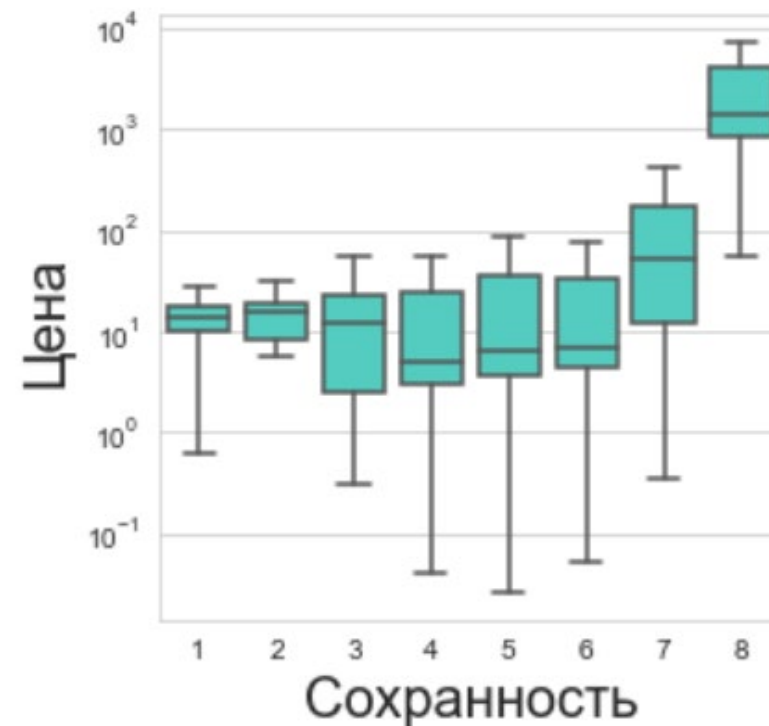
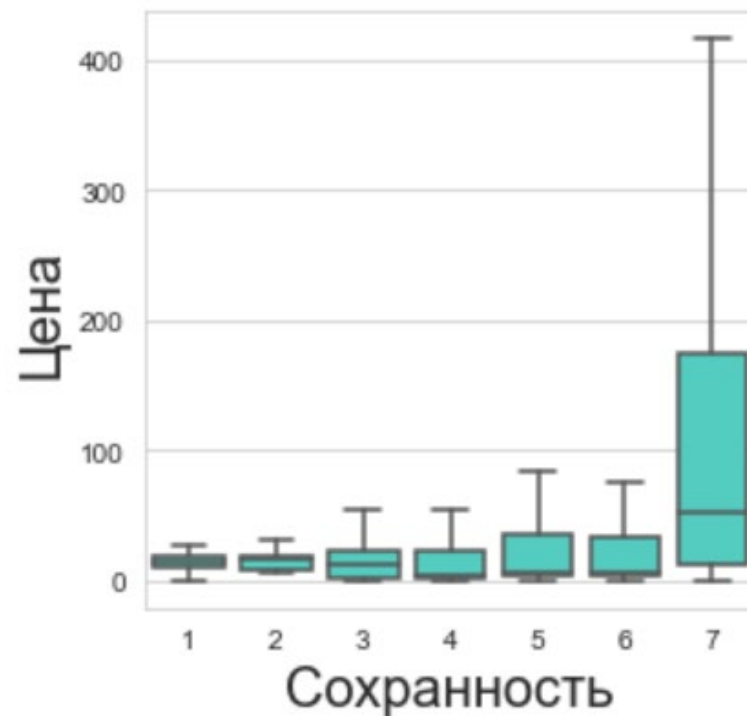
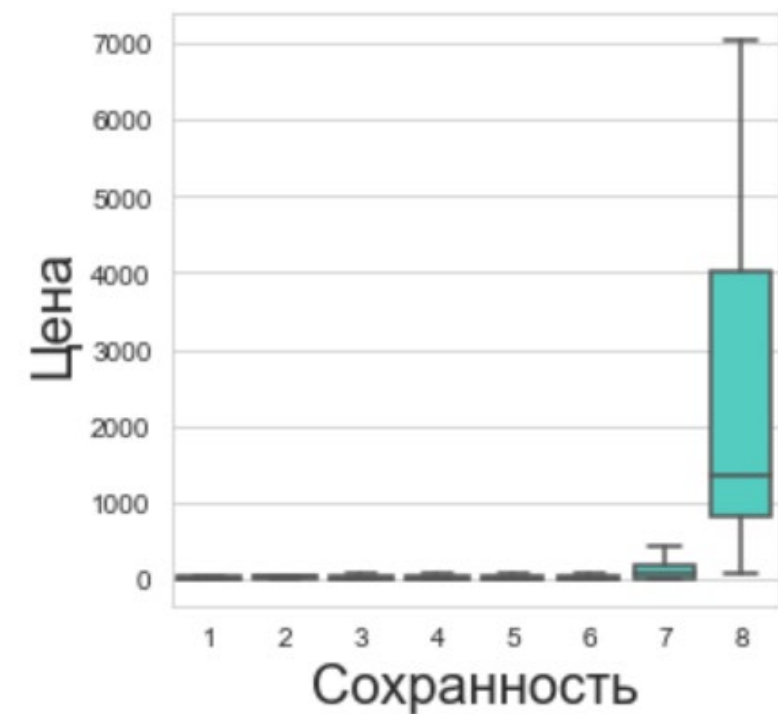
для поиска полезных фичей



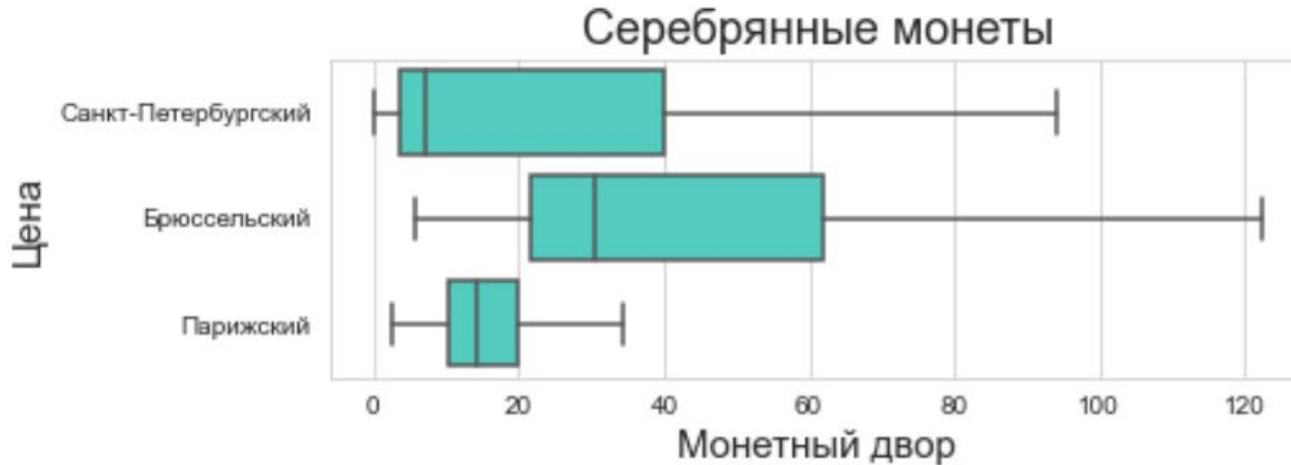
Анализ признаков

Без монет
идеального качества

log шкала

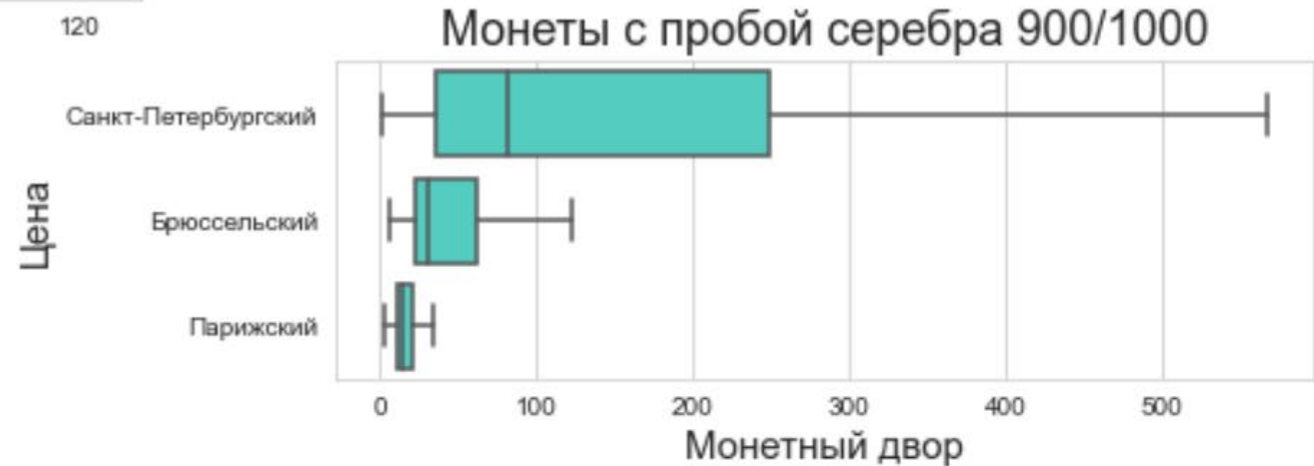


Анализ признаков



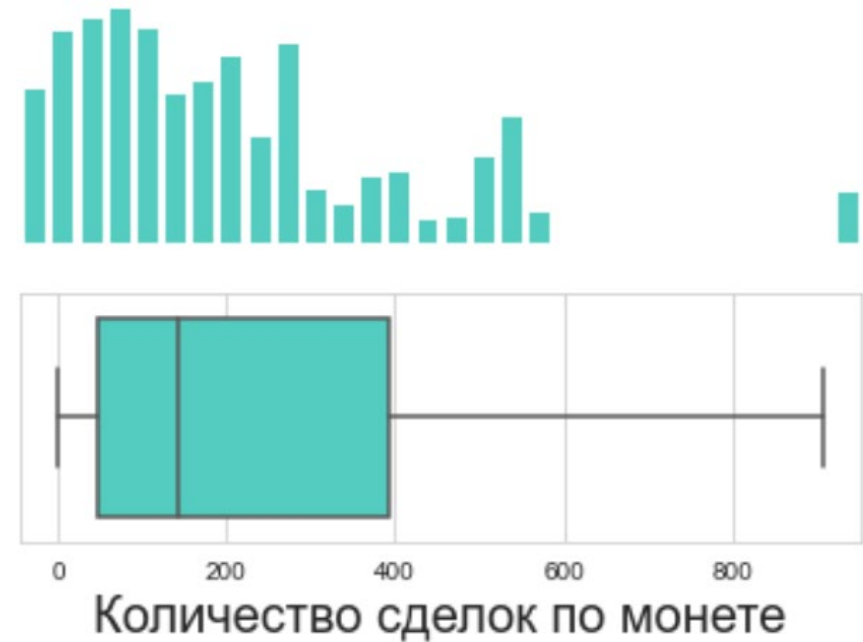
- Все монеты с пробой 500 чеканились на Санкт-Петербургском монетном дворе
- Отсюда иллюзия, что монеты с него в целом дешевле

- Серебряные монеты бывают с пробой 500 или 900
- Монеты с пробой 500 в среднем дешевле



Неликвидные монеты

- Интереснее всего оценивать стоимость монет, которые редко выставляются на аукционы
- Определим неликвидные монеты как те, которые торгуются в среднем реже раза в месяц
- Для них одной информации из временного ряда не хватит



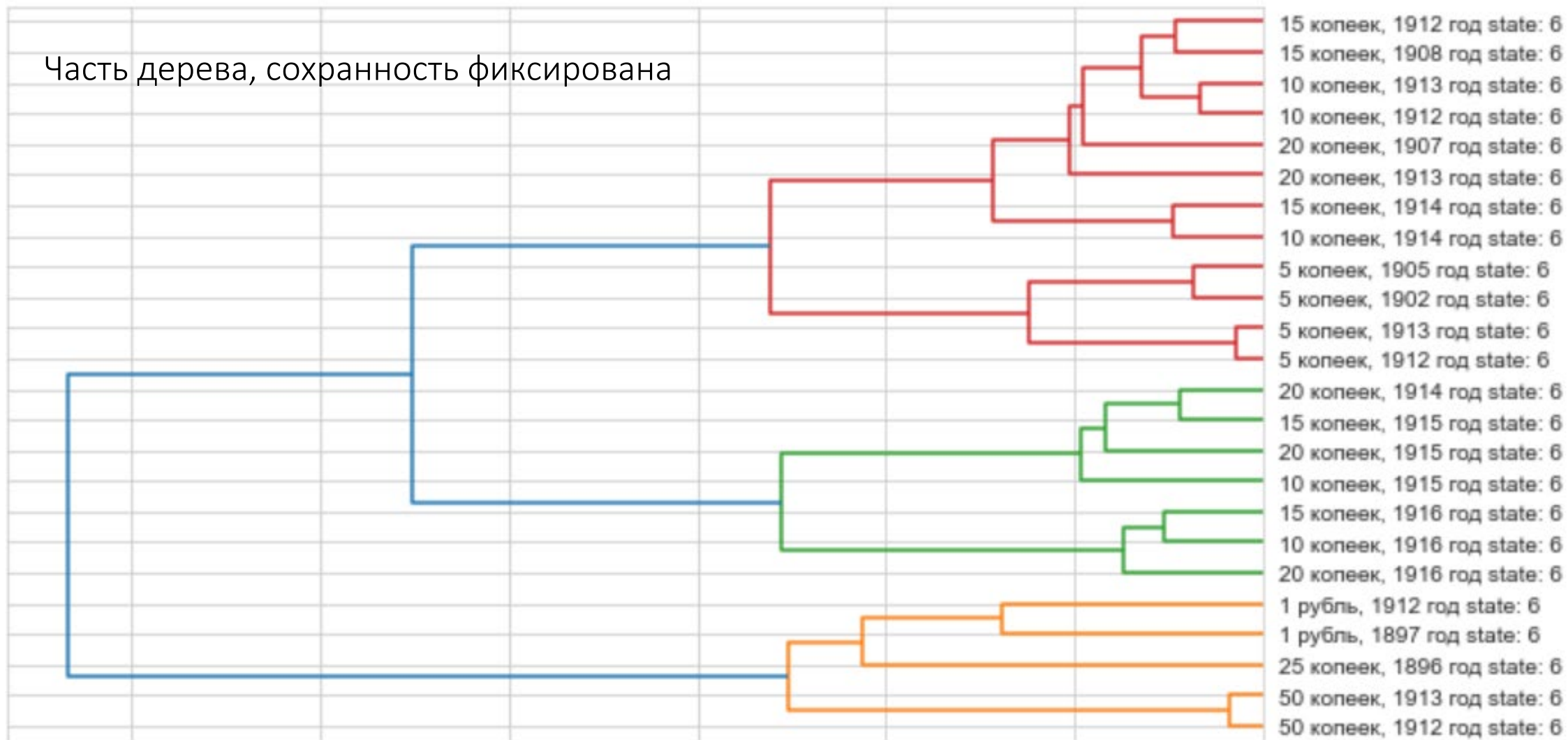
Кластеризация

- Чтобы предсказывать цену на монеты, для которых мало данных о сделках, используем информацию о «похожих» ликвидных монетах
- Идея: кластеризировать монеты
- Признаки:
 - Сохранность
 - Среднее количество ставок
 - Год выпуска
 - Номинал
 - Вес
 - Тираж
 - Средняя цена
 - Количество сделок

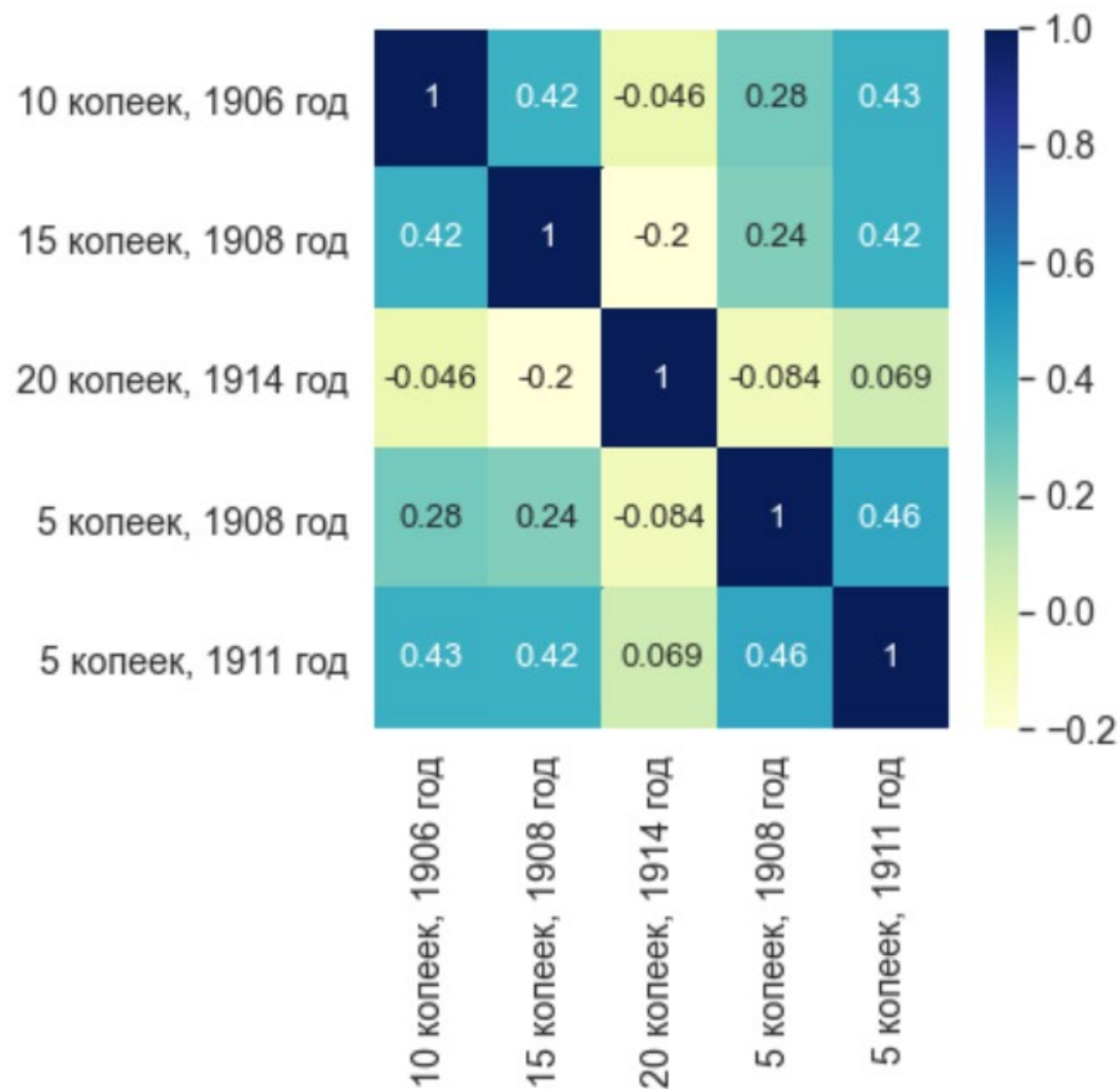


Кластеризация

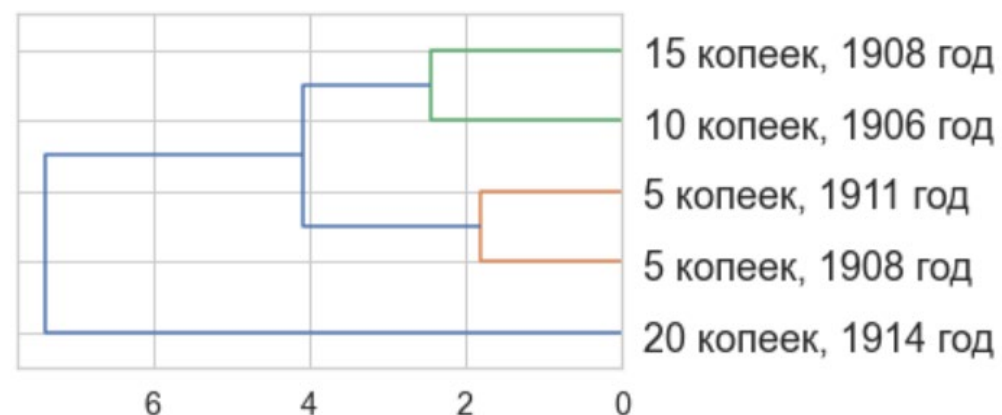
Часть дерева, сохранность фиксирована



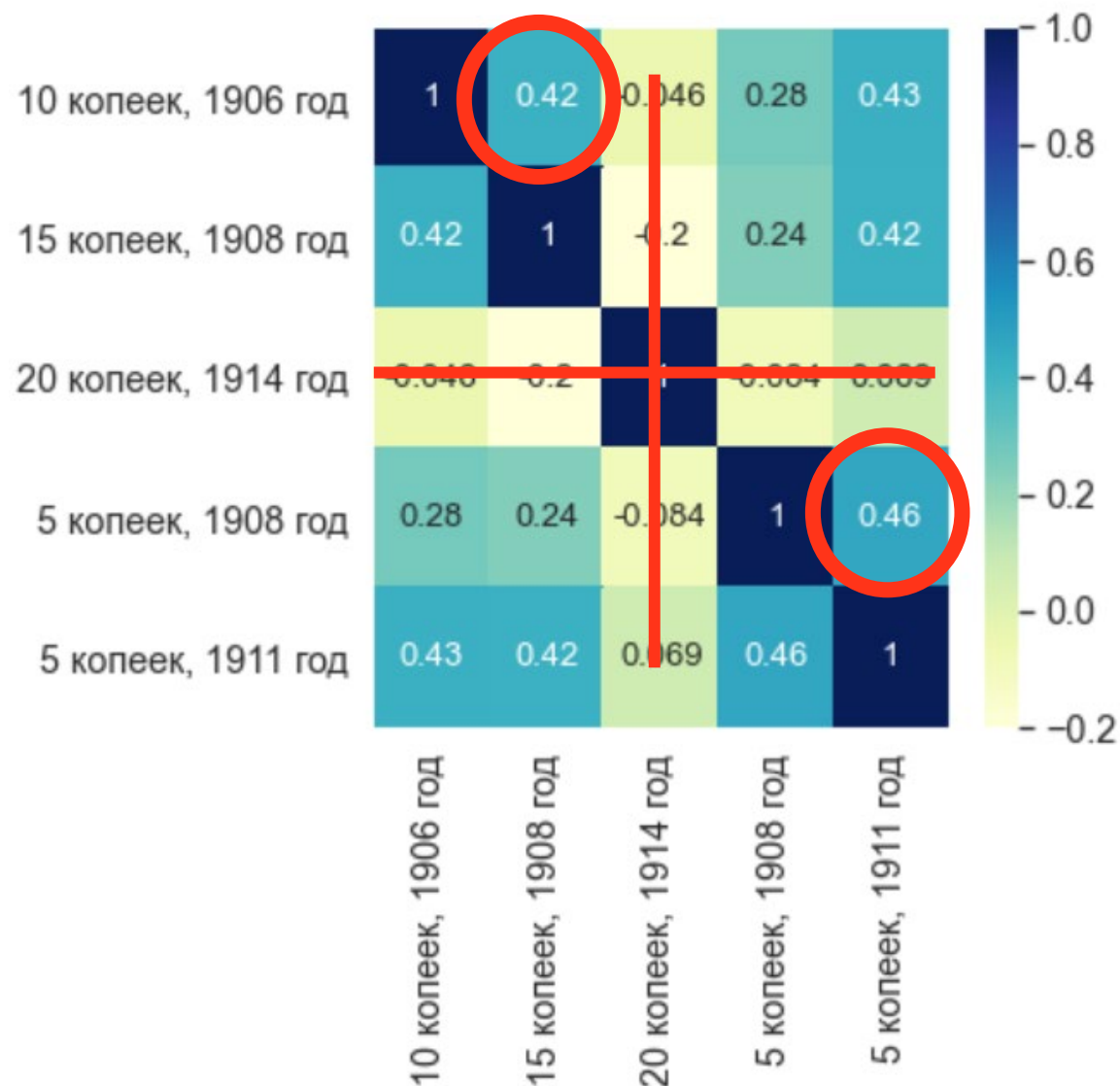
Кластеризация и корреляции



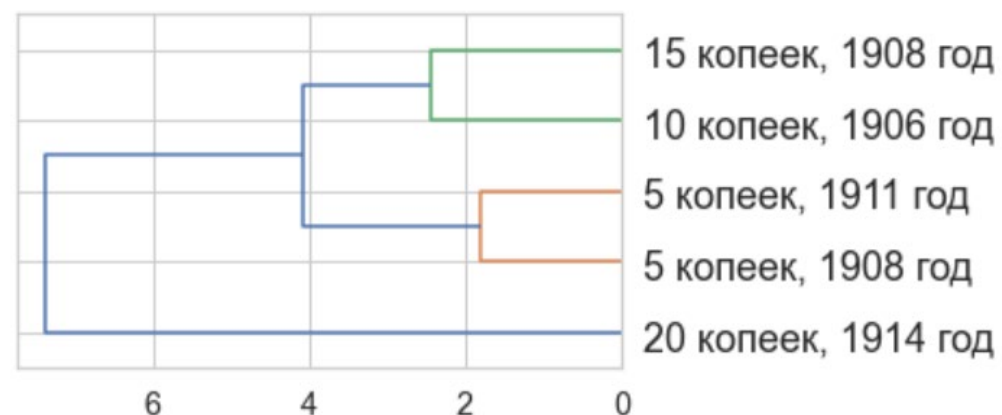
Правдоподобность кластеризации оценим с помощью корреляции (Спирмена) между ценами монет



Кластеризация и корреляции



Основные паттерны дерева прослеживаются в матрице корреляций



Дальнейшие планы

- Сформировать датасет для предсказания стоимости ликвидных монет на основе данных временных рядов
- Сформировать датасет для предсказания стоимости неликвидных монет на основе данных о прошлых ценах и стоимости «похожих» более ликвидных монет
- Подобрать и обучить модели