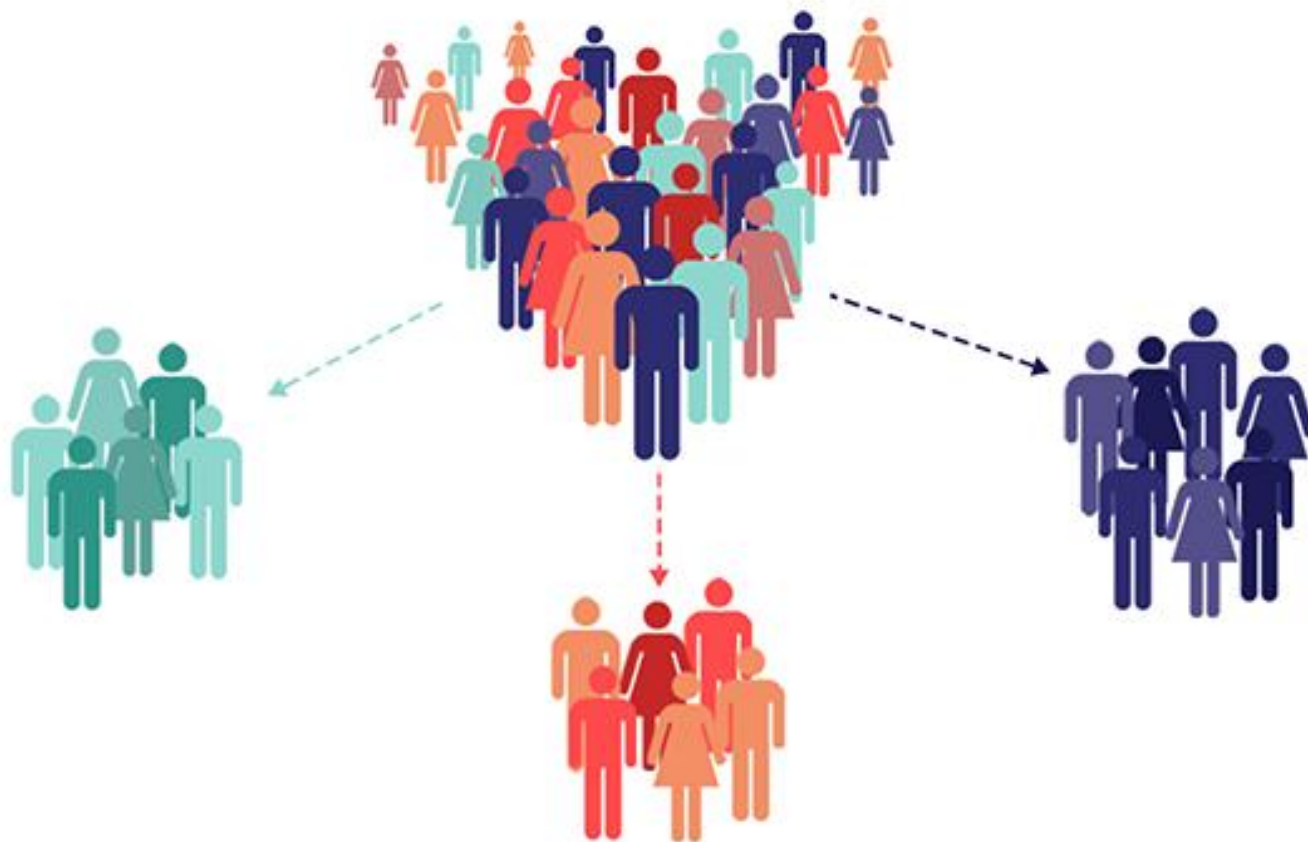


Выявление профилей потребления



Цель исследования:

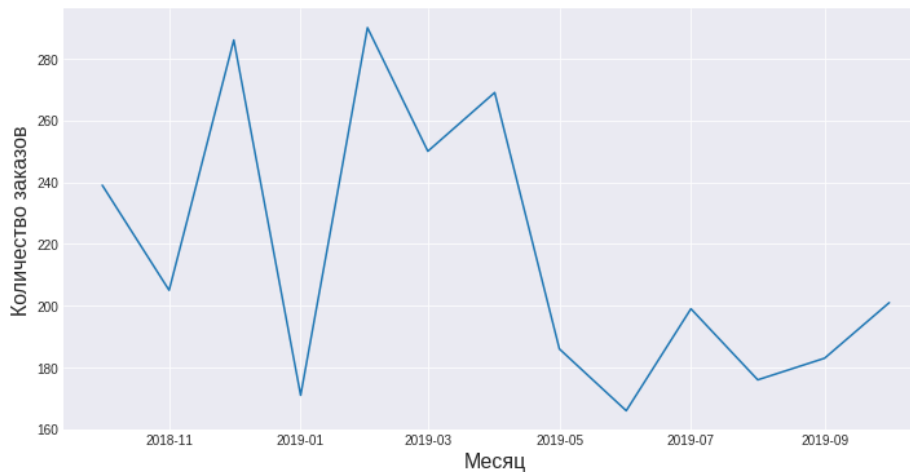
- выявление профилей покупателей на основании истории их покупок
 - анализ товарного ассортимента
 - разработка персонализированных предложений для клиентов.
-
- Исследование выполнила: Ефимова Анна (г.Москва / efimova.anna@yandex.ru)
 - Дата проведения: 15.07.2022 г.

Содержание

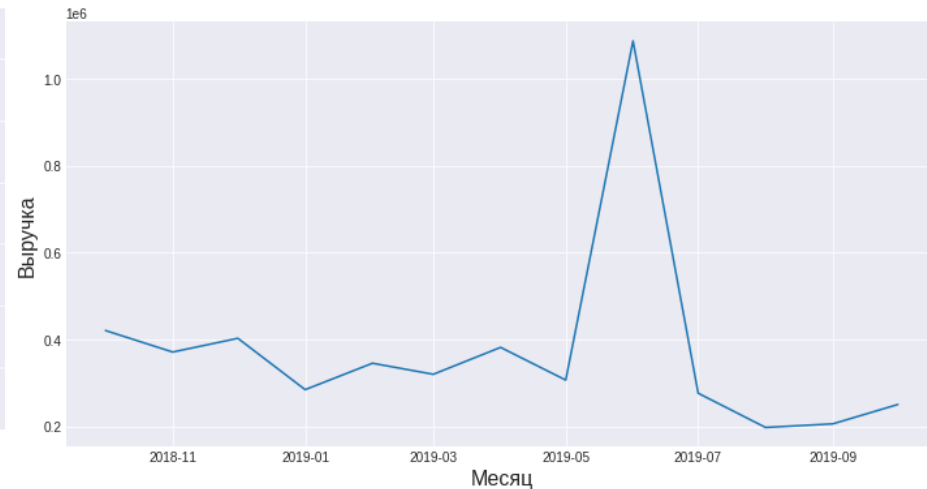
1. Динамика продаж по месяцам
2. Распределение по количеству наименований и стоимости заказа
3. Средний чек
4. Распределение по количеству заказов
5. Выручка по категориям
6. Продажи по месяцам в разрезе каждой категории
7. Кластеризация
8. Распределения признаков для кластеров
9. Выручка и количество клиентов по кластерам
10. Распределением категорий по кластерам
11. Общие выводы
12. Рекомендации
13. Источники

Динамика продаж по месяцам

Количество заказов по месяцам



Выручка по месяцам

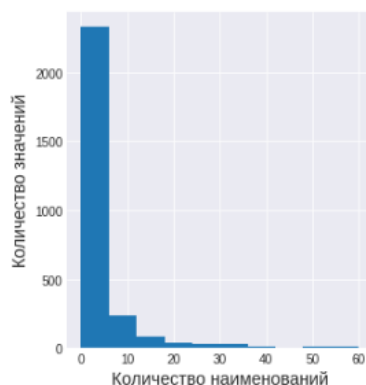


- Наблюдается спад оформления заказов в январе, июне-сентябре.
- Выручка в июне достигает пикового значения, скорее всего это связано с выбросами.

Распределение по количеству наименований и стоимости заказа

Количество наименований

Распределение по признаку



Точечный график

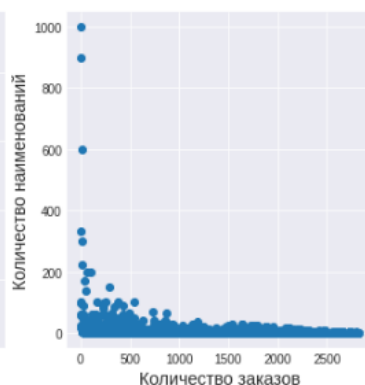
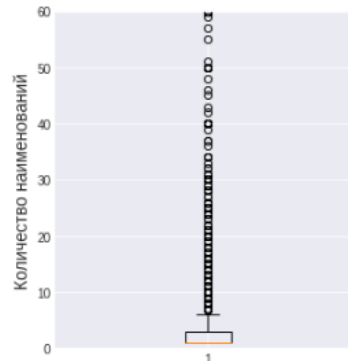
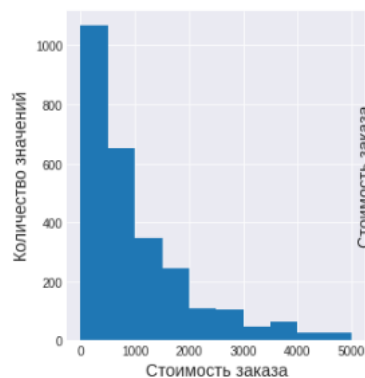


Диаграмма размаха



Стоимость заказа

Распределение по признаку



Точечный график

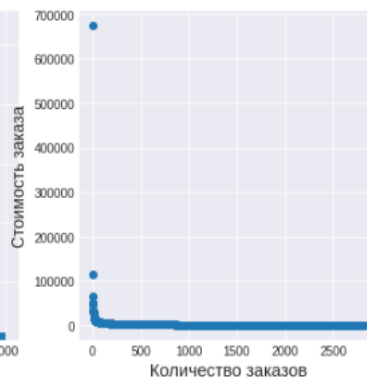
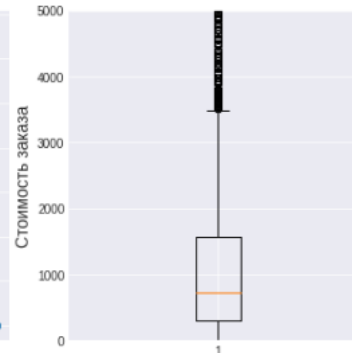


Диаграмма размаха



➤ Всего покупателей 2451
Оформлено заказов 2784

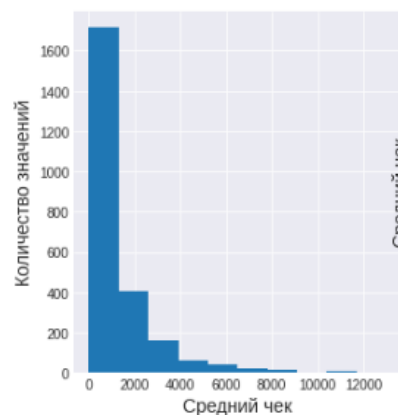
➤ Среднее количество
позиций в заказе 1 шт.

➤ Средняя стоимость заказа
727 руб.

Средний чек

Средний чек

Распределение по признаку



Точечный график

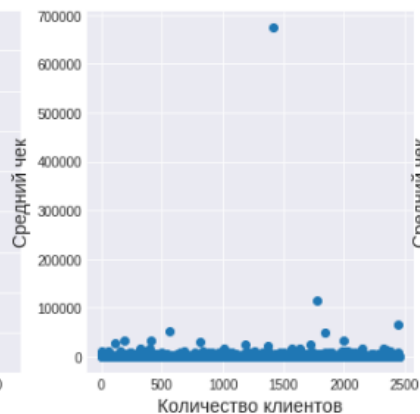
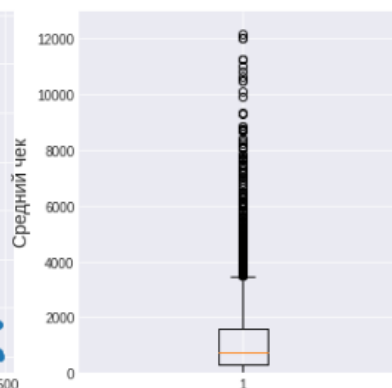
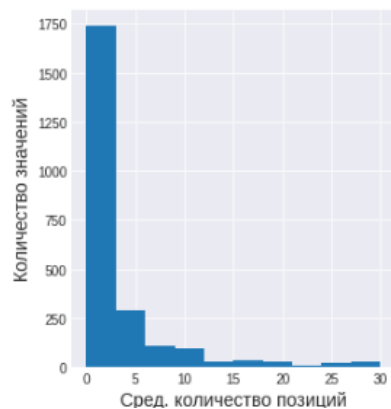


Диаграмма размаха



Сред. количество позиций

Распределение по признаку



Точечный график

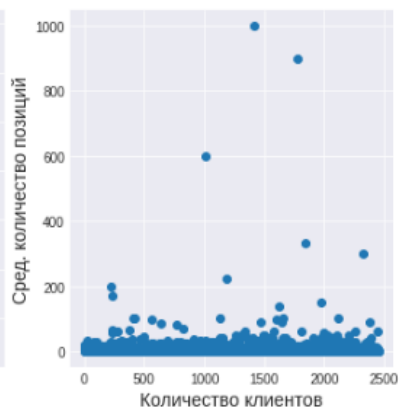
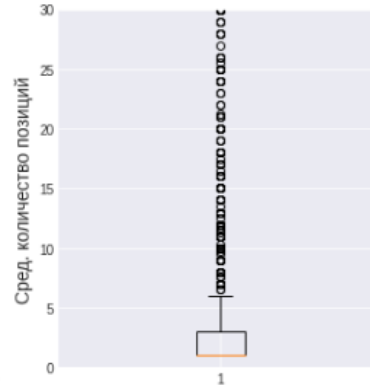
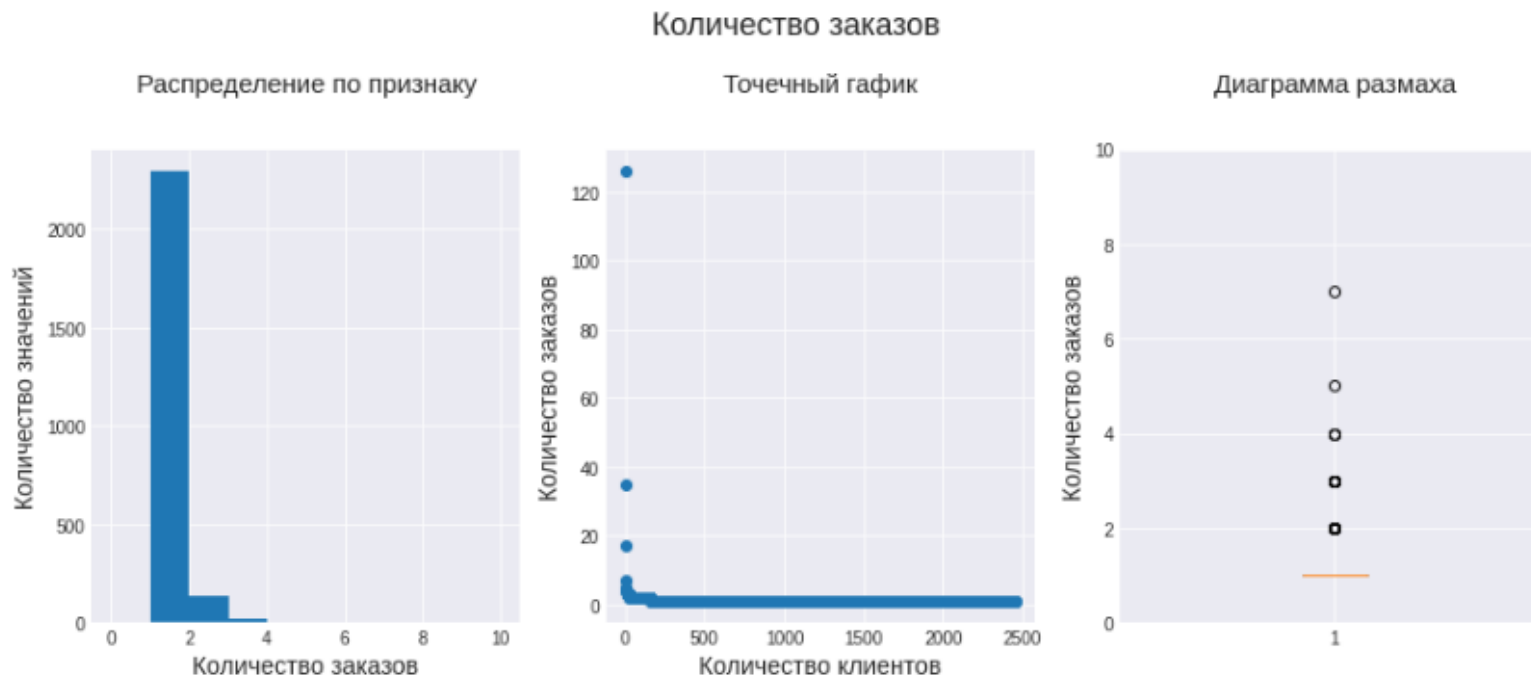


Диаграмма размаха



Распределение по количеству заказов



- В среднем клиенты делают всего 1 заказ. Средний чек 727 руб.
- Количество клиентов с 1 заказом и 1 наименованием в заказе – 1353
- Отбросили все выбросы по количеству заказов (> 5), среднему чеку ($> 10\,000$) и среднему количеству наименований в заказе (> 60).

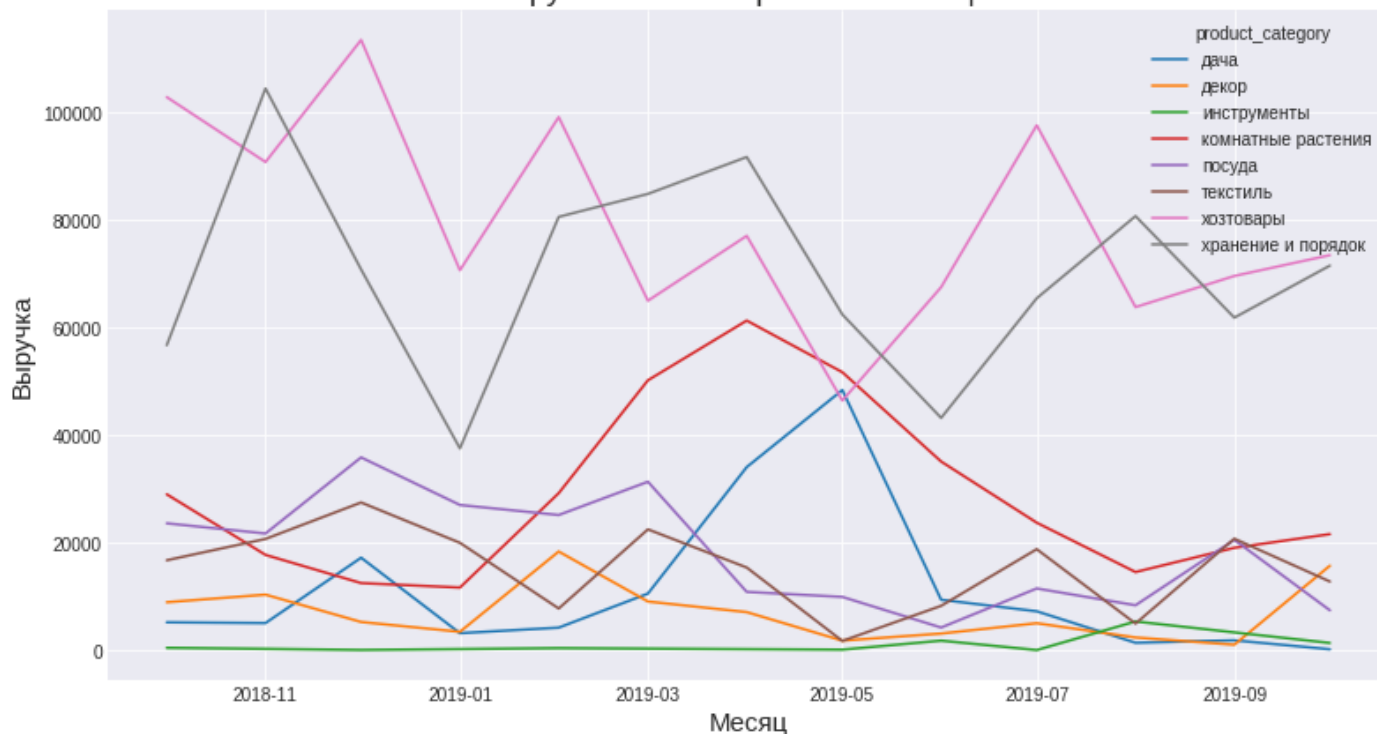
Выручка по категориям



- На основе категоризации товаров выделили 8 категорий товаров.
- Больше всего выручка в категориях "Хозтовары" и "Хранение и порядок".

Продажи по месяцам в разрезе каждой категории

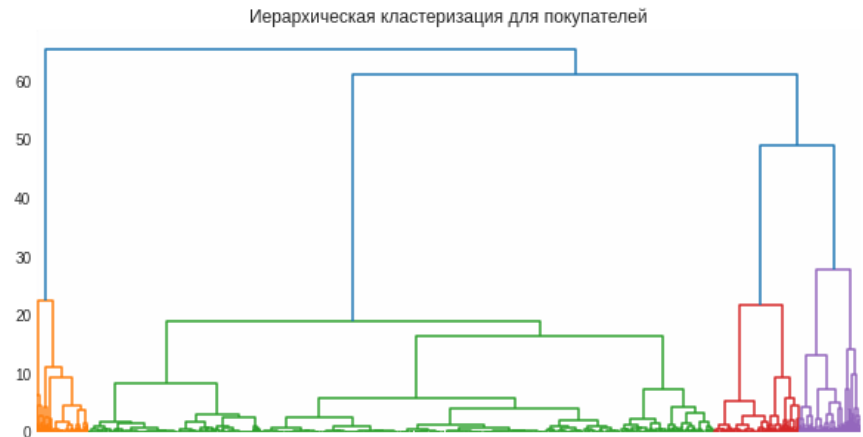
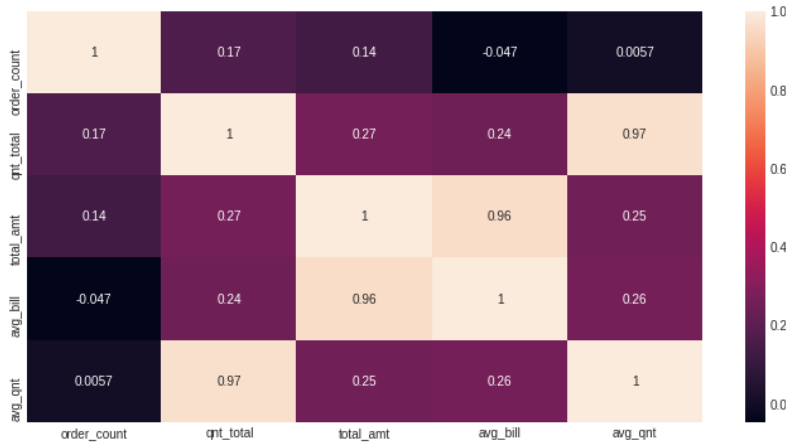
Выручка по категориям по месяцам



Продажа в некоторых категориях носит сезонный характер:

- пик продаж для категории "комнатные растения" в апреле, для категории "дача" - в мае
- категории "хозтовары" и "хранение и порядок" всегда востребованы, продажи по ним проседают в январе и июне
- "посуда" популярна в декабре и марте - время праздников

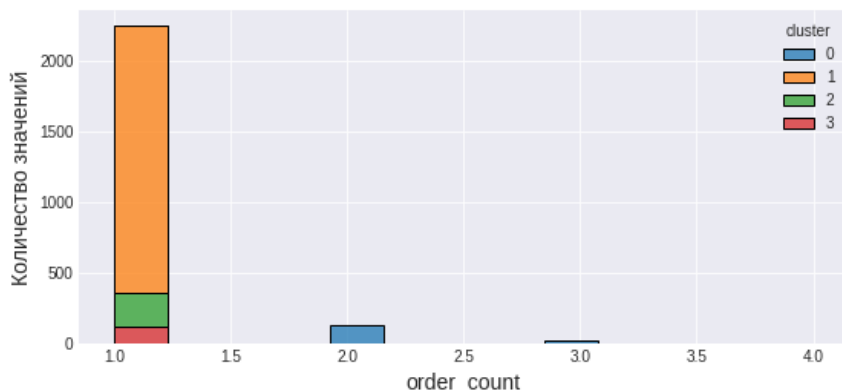
Кластеризация



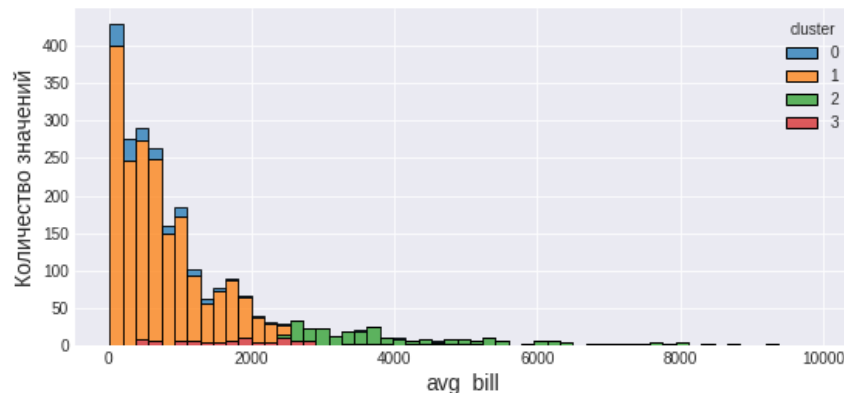
- На основании матрицы корреляции удалили попарно коррелирующие признаки total_amt, avg_bill и qnt_total, avg_quantity
- С помощью иерархической кластеризации выделили оптимальное количество кластеров – 4.

Распределения признаков для кластеров

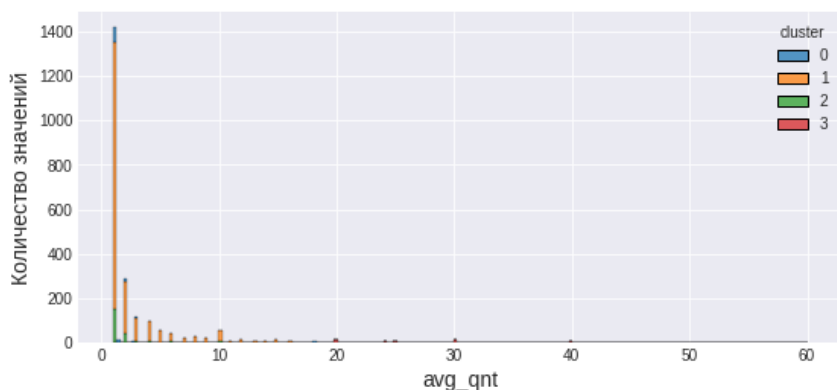
Распределение по признаку order_count



Распределение по признаку avg_bill



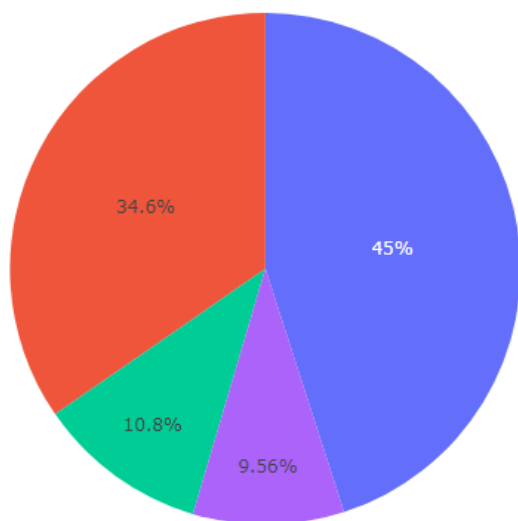
Распределение по признаку avg_qnt



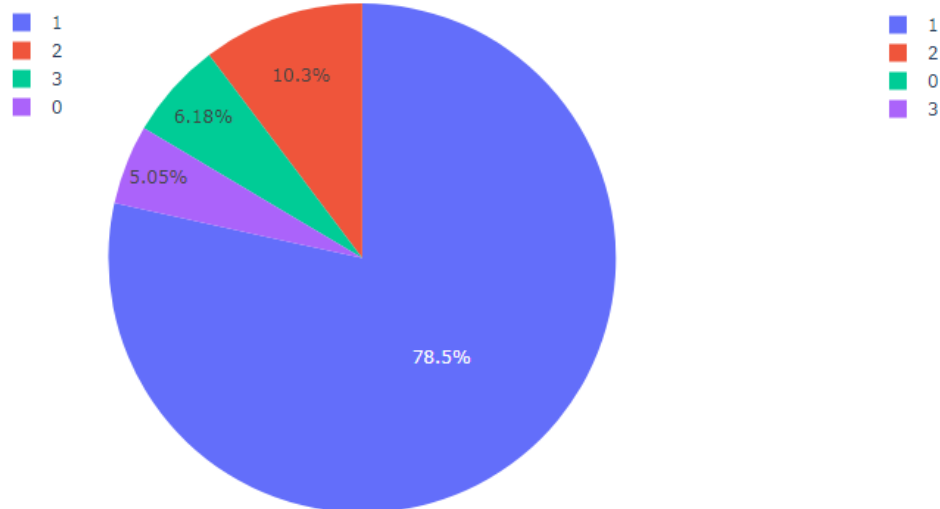
- В 1 и 2 кластере клиенты делают в среднем 1 заказ со средним количеством наименований 1 шт., но средний чек в 1 кластере маленький, а во 2-м высокий
- В 0 кластере клиенты делают больше 1 заказа, с небольшим средним чеком
- В 3 кластере клиенты делают 1 заказ с большим количеством наименований и с большим за счет этого средним чеком

Выручка и количество клиентов по кластерам

Выручка по кластерам

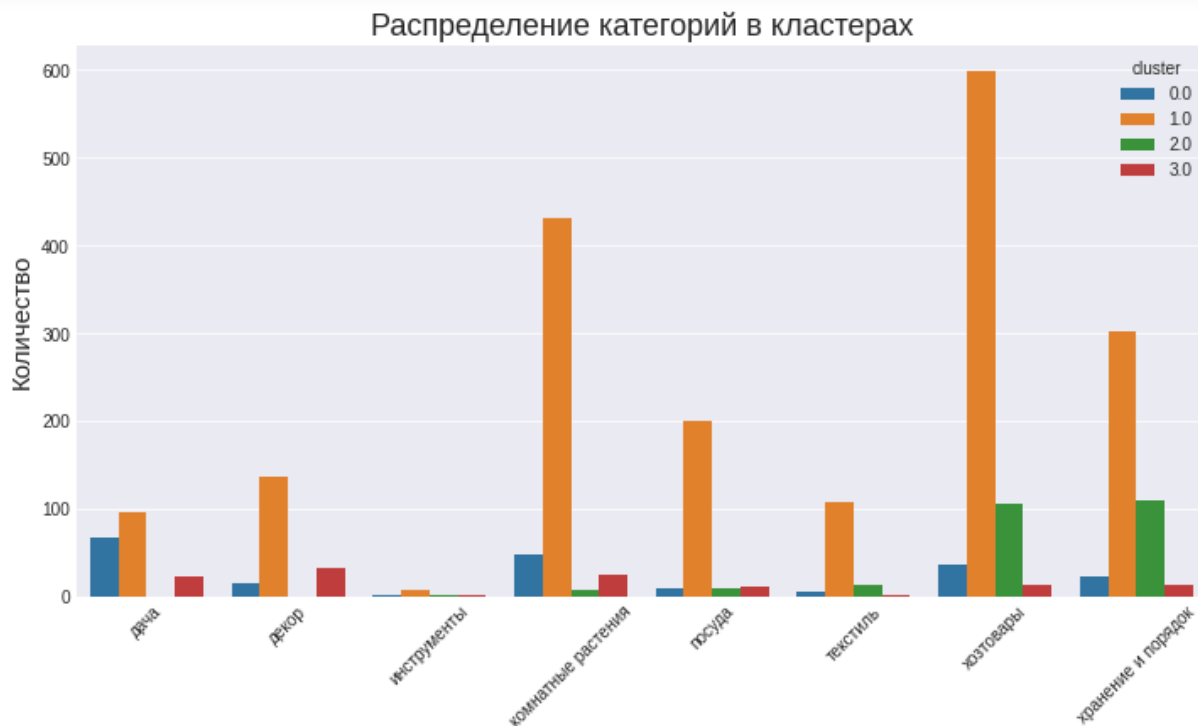


Число покупателей по кластерам



➤ Самый большой по количеству клиентов - 1 кластер и он приносит больше всего выручки, почти половину от общей выручки

Распределение категорий по кластерам



- «0» кластер предпочитает комнатные растения и рассаду
- «1» кластер - недорогие хозтовары и комнатные растения
- «2» кластер - дорогие хозтовары и товары из категории "хранение и порядок"
- «3» кластер - покупает много дешевых позиций из категорий декор, дача и растения.

Общие выводы

- ❖ Больше всего выручка в категориях "Хозтовары" и "Хранение и порядок".
- ❖ В некоторых категориях продажи носят сезонный характер
- ❖ Выделили 4 кластера покупателей. В 1 и 2 кластере клиенты делают в среднем 1 заказ со средним количеством наименований 1 шт., но средний чек в 1 кластере маленький, а во 2-м высокий. В 0 кластере клиенты делают больше 1 заказа, с небольшим средним чеком. В 3 кластере клиенты делают 1 заказ с большим количеством наименований и с большим за счет этого средним чеком
- ❖ Самый многочисленный 1 кластер и он приносит больше всего выручки
- ❖ Нашли приоритетные категории для каждого клиента и посмотрели на распределение приоритетных категорий в каждом кластере. 0 кластер предпочитает комнатные растения и рассаду, 1 кластер - недорогие хозтовары и комнатные растения, 2 кластер - дорогие хозтовары и товары из категории "хранение и порядок", 3 кластер - покупает много дешевых позиций из категорий декор, дача и растения

Рекомендации

- ❖ Подготовить персонализированные предложения для клиентов, учитывая их предпочтения по товарам и ценовом сегменте: отправлять письма с подборками товаров из интересующей категории, со скидками на эти товары
- ❖ Клиентам с 1 заказом и 1 позицией в заказе (в основном 1 и 2 кластеры) - отправить письмо с напоминанием о магазине и промокодом на скидку
- ❖ Клиентам 0 кластера, у которых больше 1 заказа, рассылать письма с новыми товарами из приоритетных для них категорий
- ❖ Клиентам из 3 кластера с большим количеством наименований в заказах предлагать скидки на крупные заказы, в зависимости от определенного количества.
- ❖ Учитывать сезонность некоторых категорий товаров и рекомендовать их в подходящее время, как например товары из категорий "дача" и "комнатные растения"

Источники

- Источники данных: транзакции интернет-магазина товаров для дома и быта «Пока все ещё тут». (таблица *ecommerce_dataset.csv*).
- Данные предоставлены компанией Яндекс (Яндекс.Практикум / курс «Аналитик данных»)