

Дипломный проект

Анализ и предсказание объема
продаж



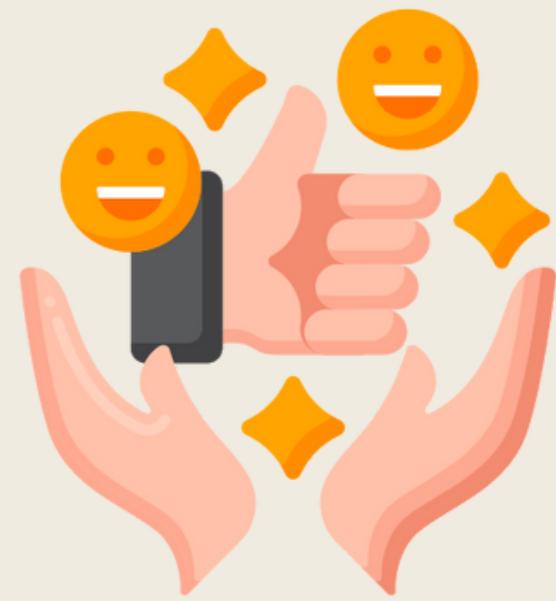
Входные данные

| Название переменной | Описание |
|--------------------------|---|
| date | Дата |
| store_nbr | Идентификационный номер магазина |
| family | Семейство товаров |
| sales | Общее количество проданных товаров для семейства в конкретном магазине в указанную дату |
| onpromotion | Общее количество товаров в семействе, которые были по акции в данный день |
| city/state | Город/Провинция, где находится магазин с соответствующим store_nbr |
| type | Тип магазина (A, B, C, D, E) |
| cluster | Кластер (группа похожих магазинов), к которому принадлежит магазин |
| transactions | Количество транзакций в конкретном магазине в указанную дату |
| type | Тип дня (рабочий/ праздник/ перенос/ доп/ мост/ событие) |
| locale | Локальный/ региональный/ национальный праздник |
| locale_name, description | Имя праздника, описание |
| transferred | Был ли перенос праздничного дня |
| dcoilwtico | Цена на нефть в указанную дату |

Цели предсказания:



Сокращение
продовольственных отходов



Повышение
удовлетворенности
покупателей



Увеличение выручки за счёт
оптимизации объёма
закупок

Задачи проекта:

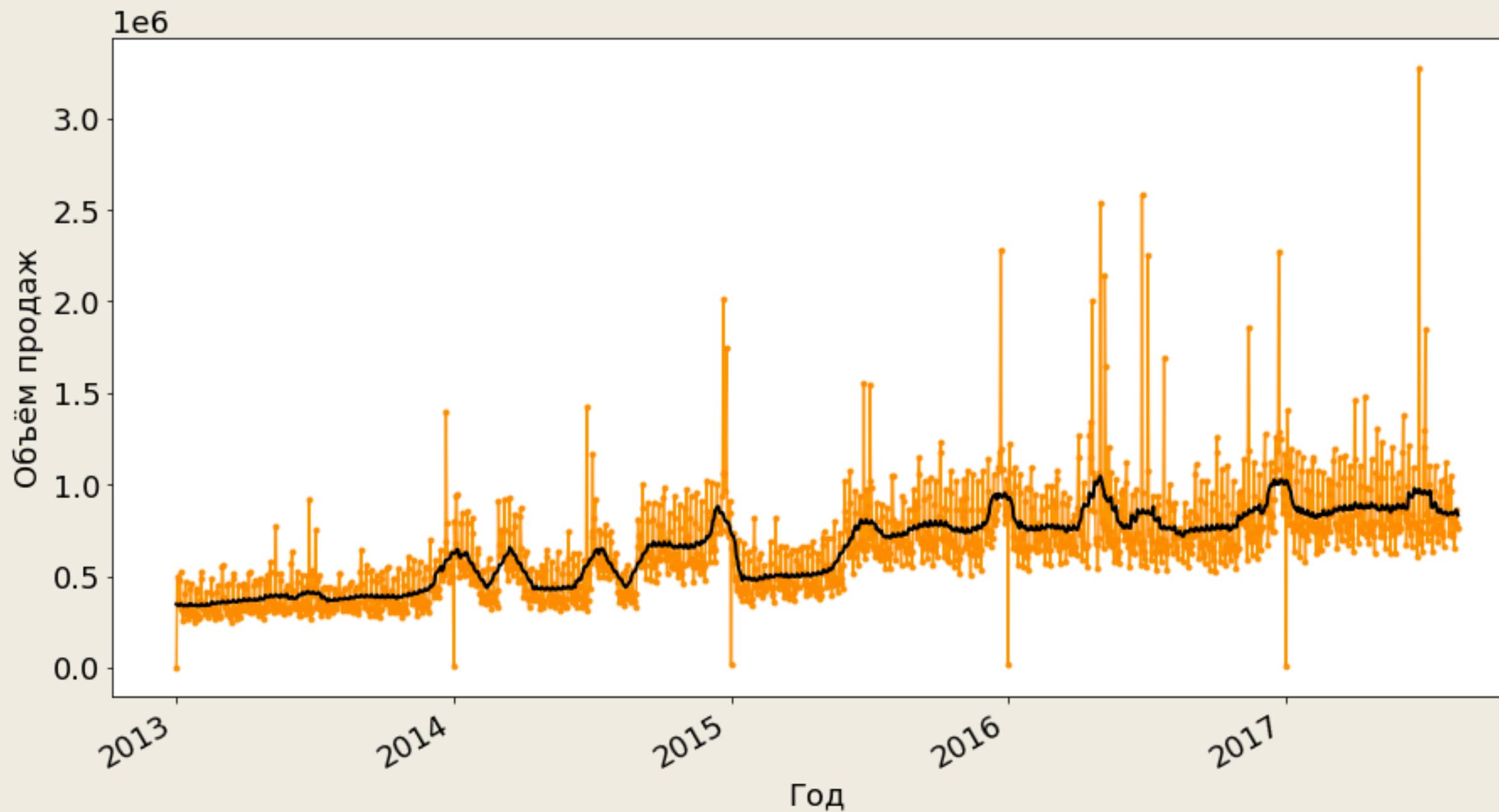
- Изучить объёмы продаж
- Изучить признаки
- Изучить влияние признаков на объёмы продаж
- Сгенерировать новые признаки
- Обучить модель и посчитать метрики



График изменения продаж

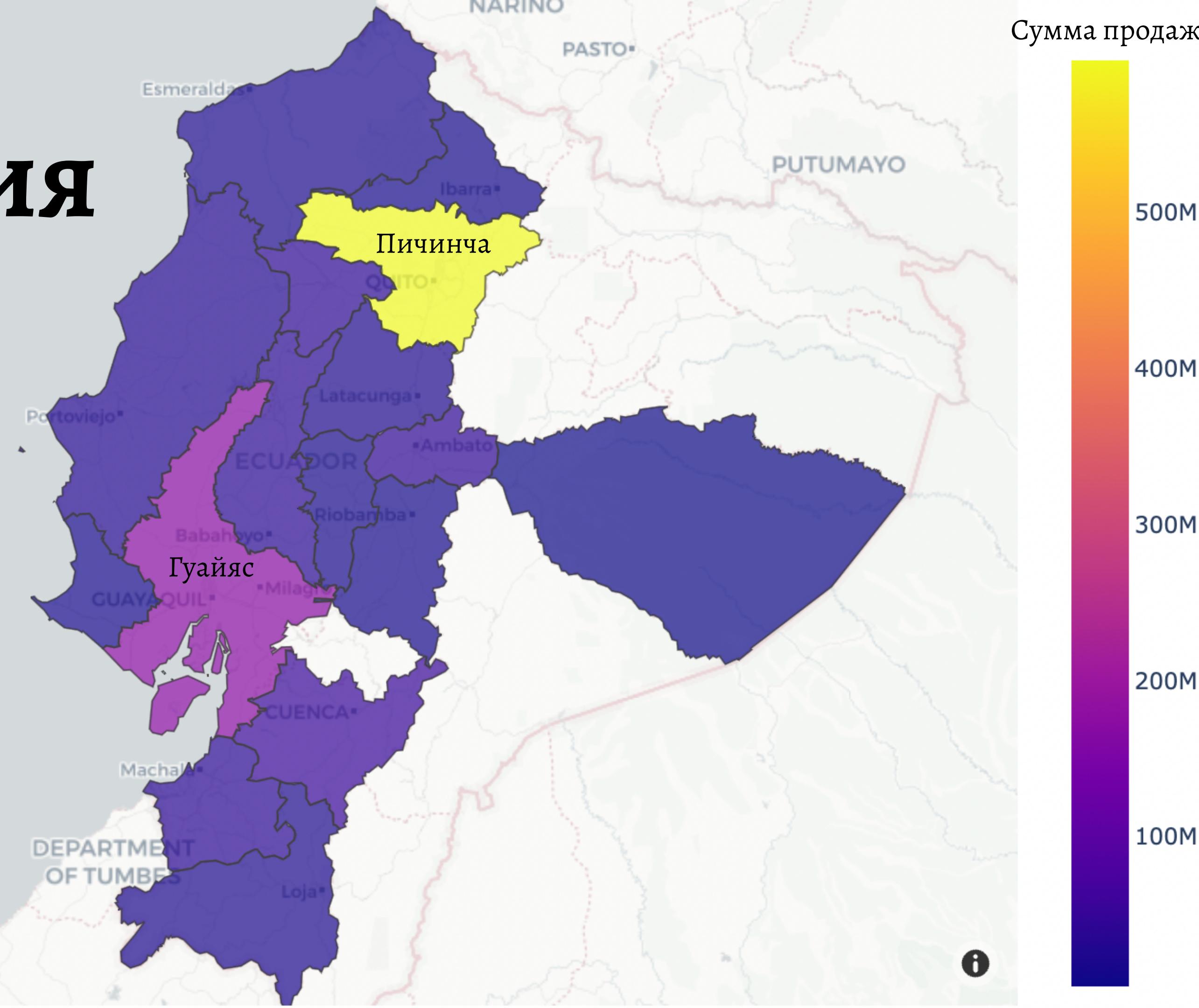
с выводом скользящего окна по среднему значению

Видно, что суммарный объём продаж с течением времени растёт и за рассматриваемый период с 2013 по 2017 гг. объём увеличился примерно в 2 раза.

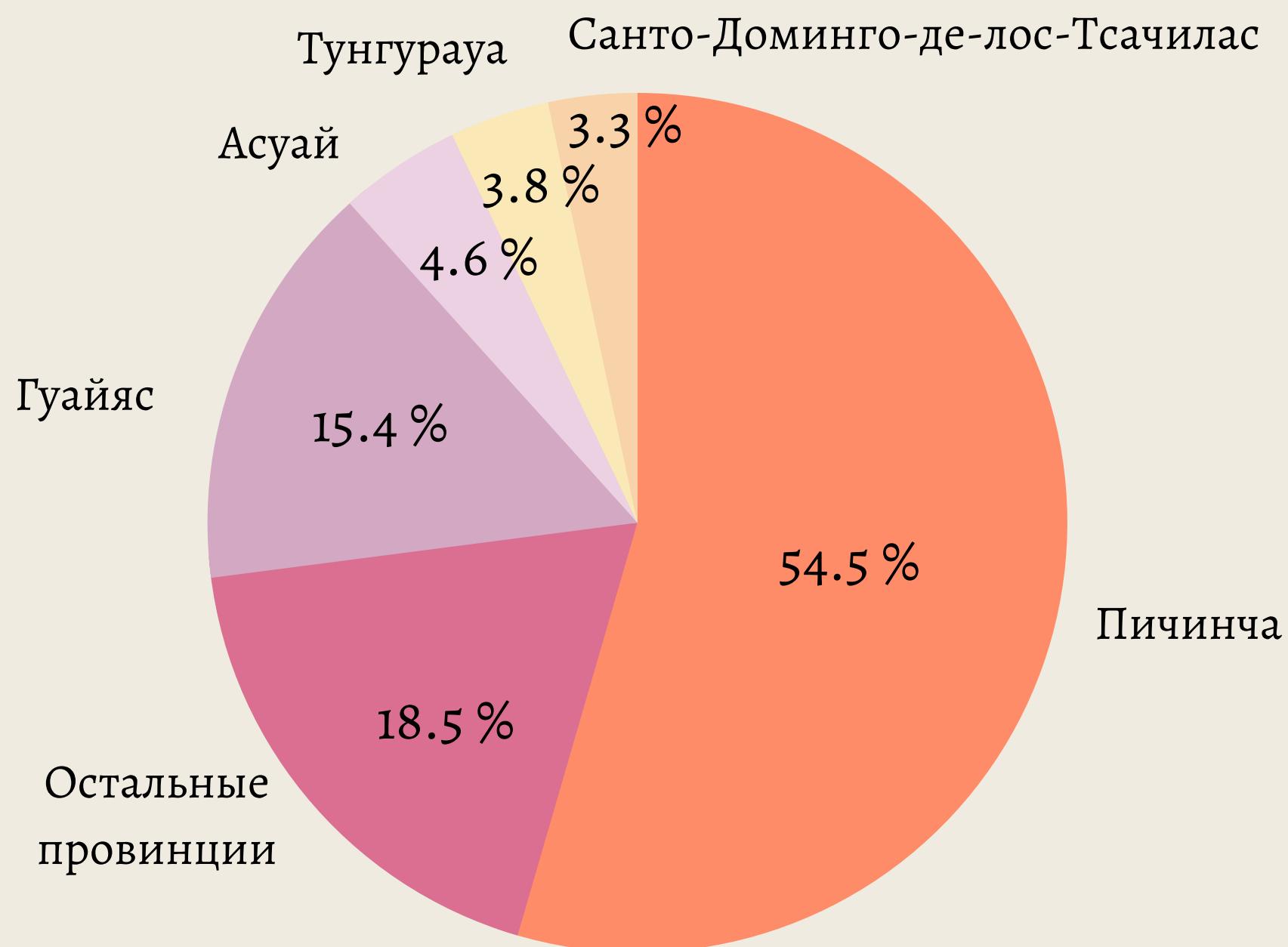


Сумма продаж

География продаж

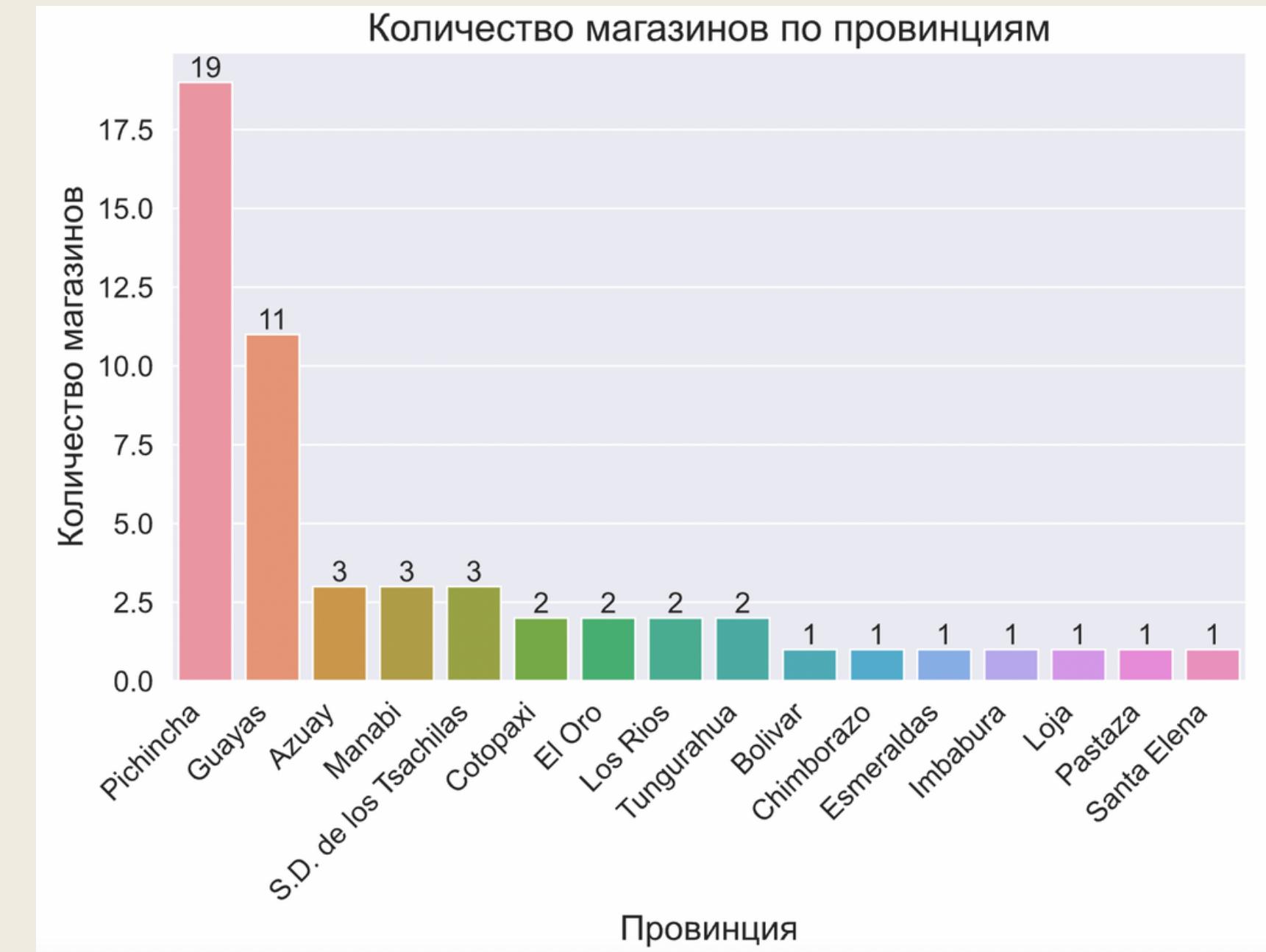
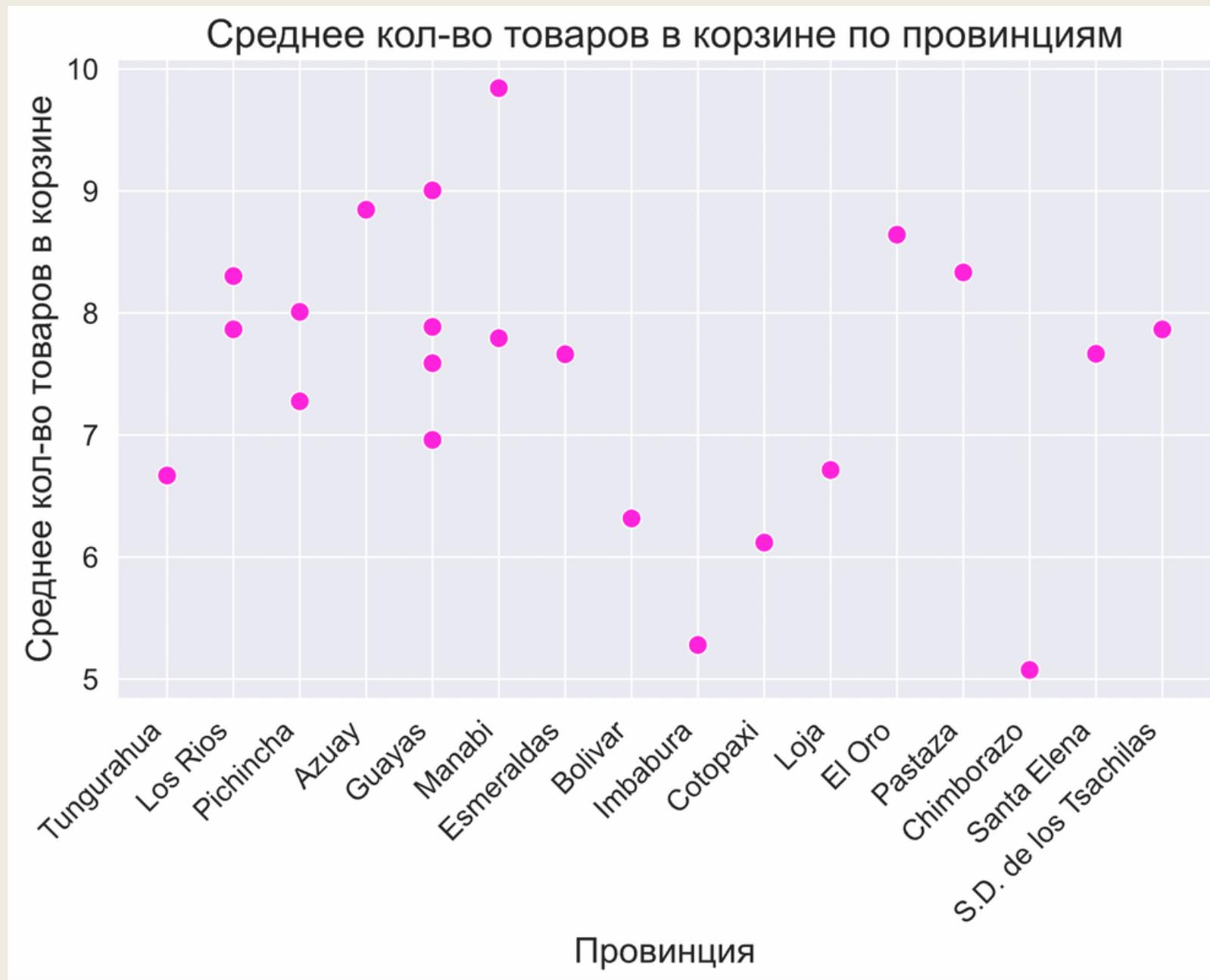


География в цифрах



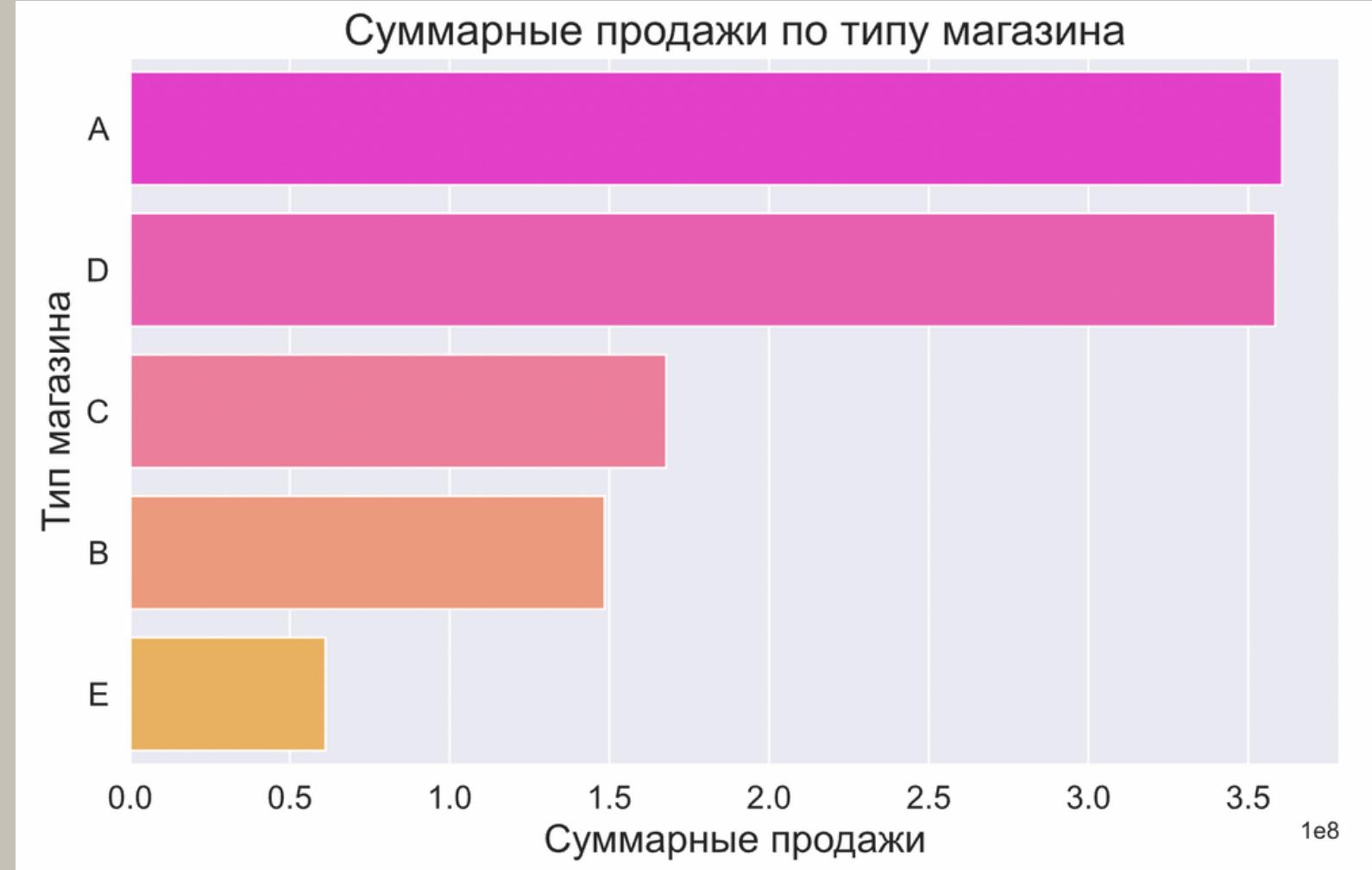
| Провинция (город) | Объёмы продаж | Доля продаж в админ. городе | Число магазинов |
|-------------------------------|---------------|-----------------------------|-----------------|
| Пичинча (Кито) | 597 585 883 | 95.2 % | 19 |
| Остальные провинции | 202 367 485 | - | 16 |
| Гуаяс (Гуаякиль) | 168 649 985 | 74.5 % | 11 |
| Асуай (Куэнка) | 50 194 046 | 100 % | 3 |
| Тунгурахуа (Амбато) | 41 159 773 | 100% | 2 |
| Санто-Доминго-де-лос-Тсачилас | 36 617 572 | 100 % | 3 |

Анализ продаж по провинциям

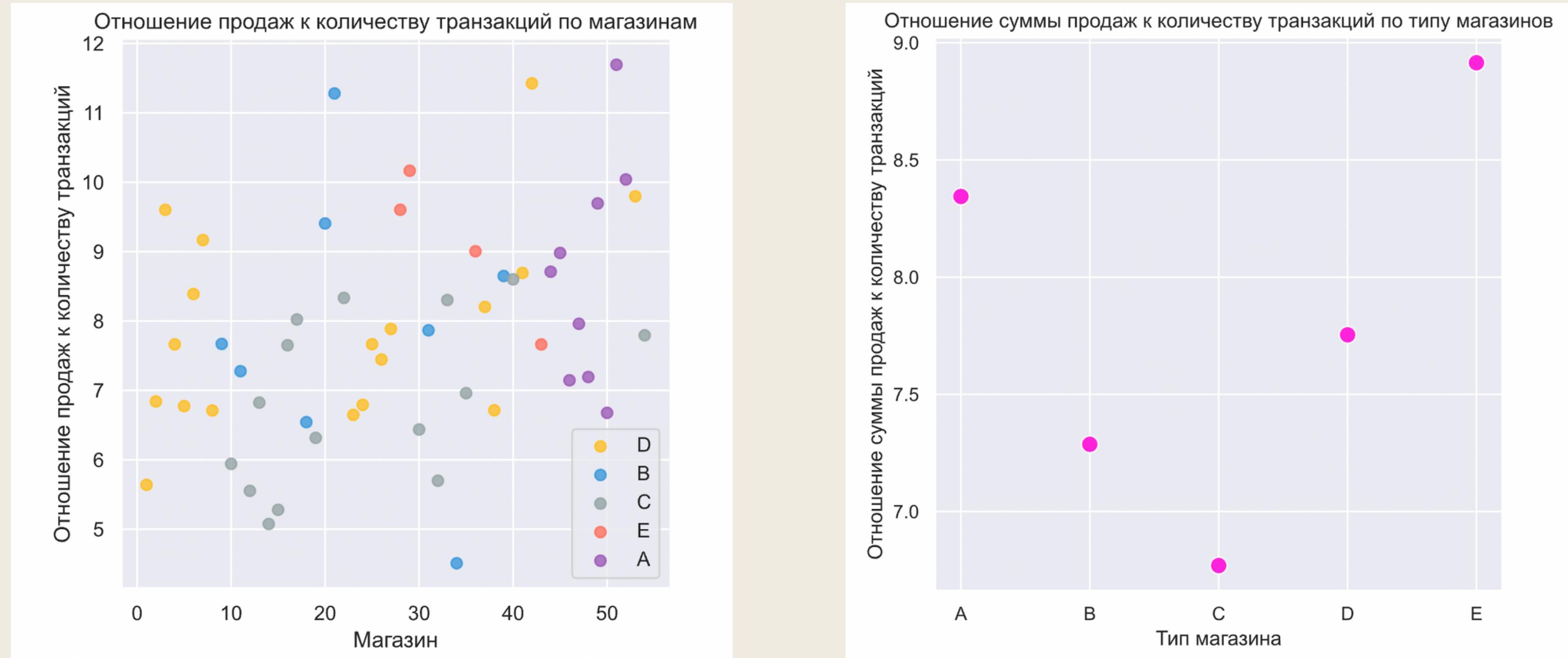


Анализ продаж по типу магазина

- Магазины типов А и D лидируют по сумме продаж
- Магазины типа D преобладают по общему количеству
- При этом число магазинов типа А (лидеров по сумме продаж) в 2 раза меньше, чем типа D
- Наименьшая сумма продаж у магазинов типа Е, число которых также минимально



Среднее количество товаров в чеке



- У магазинов типа Е наибольшее среднее количество товаров в чеке. Однако к данному типу относится всего 4 магазина.
- У магазинов типа С наихудший средний показатель - ниже 7.

Суммарные продажи по годам

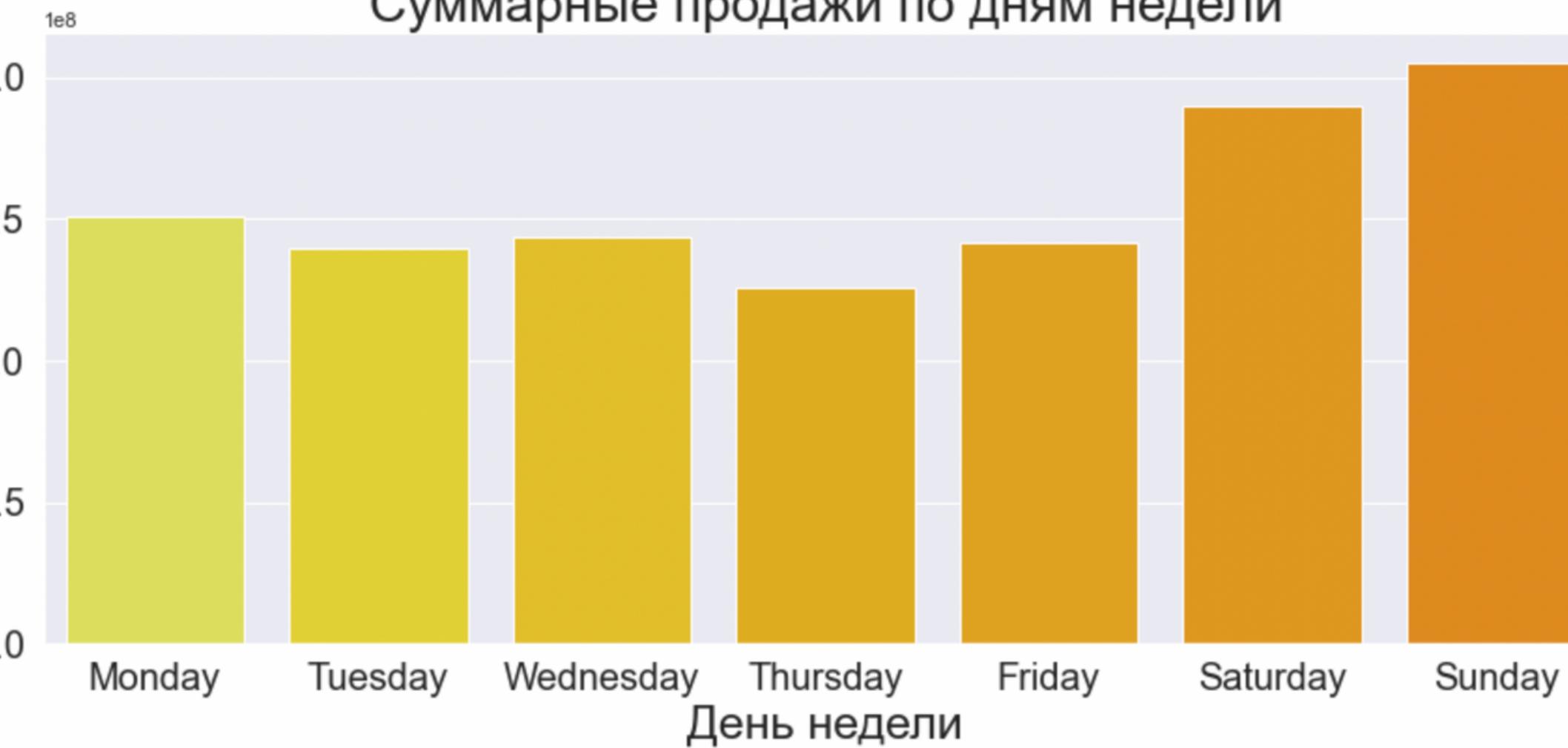
Сумма продаж



Суммарные продажи растут от года к году.
Объем за 2017 год меньше, так как и данных за
этот год всего половина.

Суммарные продажи по дням недели

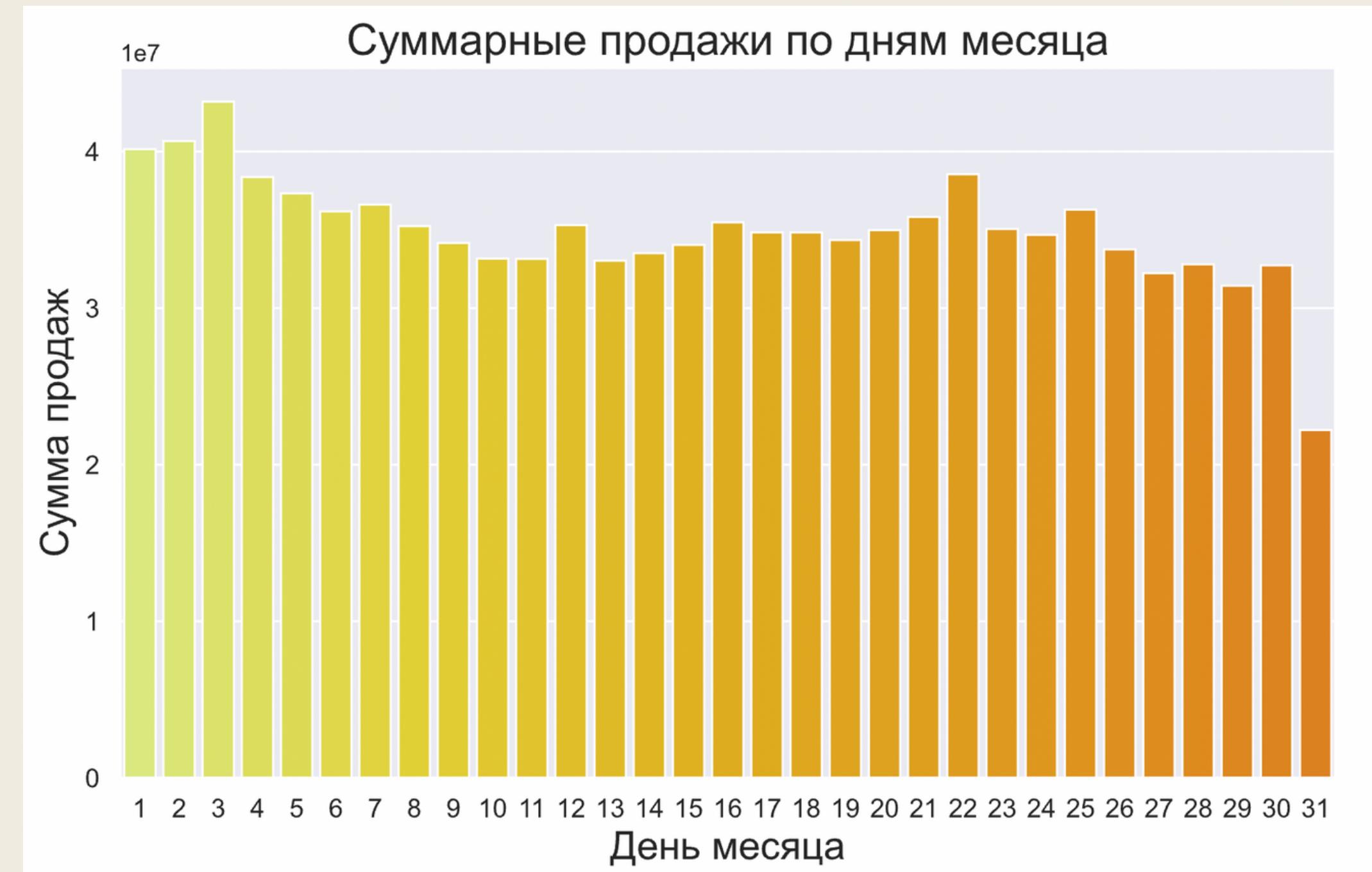
Сумма продаж



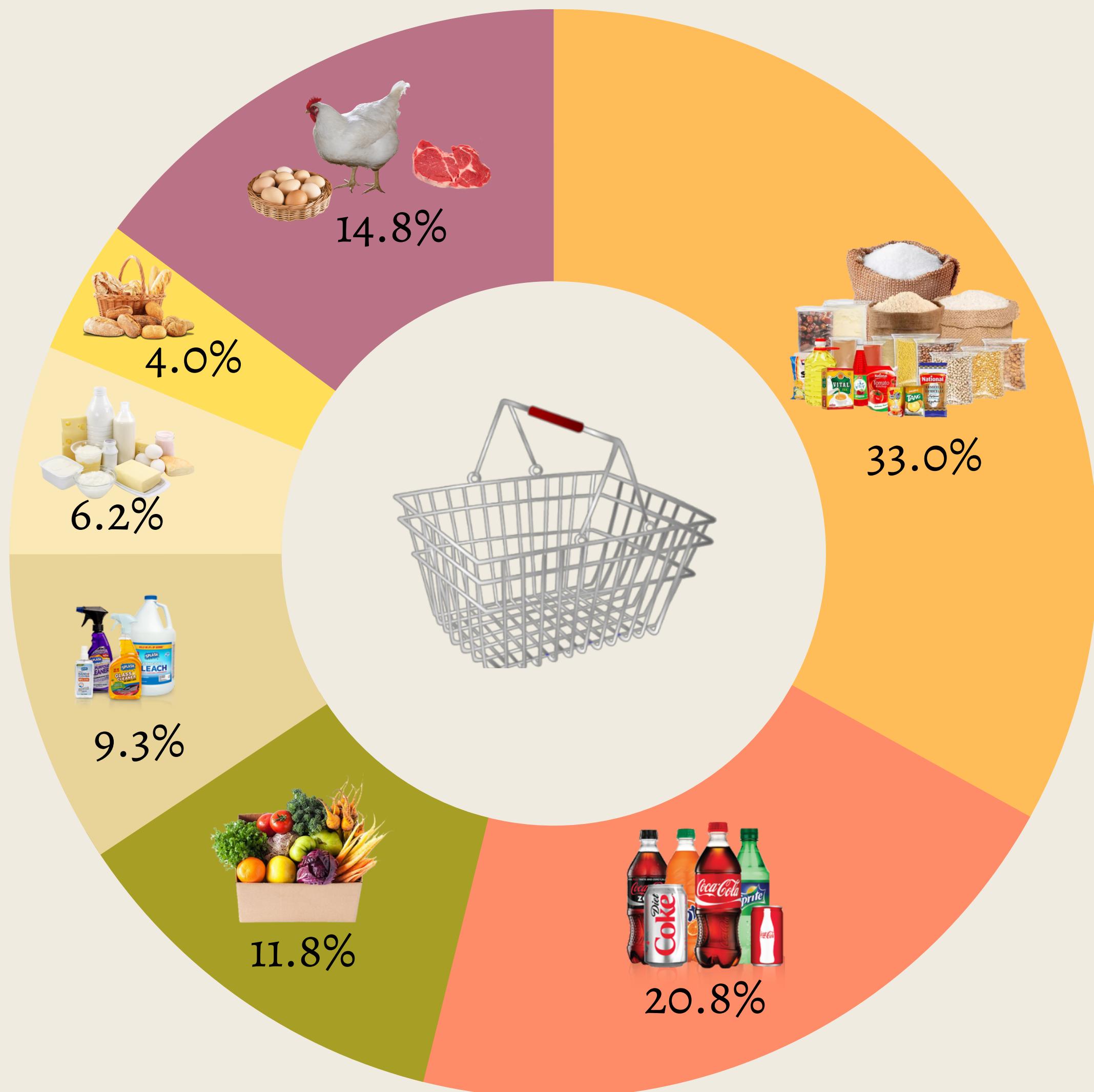
Наибольшие объемы продаж приходятся на
выходные.

Анализ продаж по дням месяца

- Продажи в первых числах месяца выглядят наибольшими.
- В целом выделяются из ряда 3, 12, 22 и 25 числа месяца.
- Зарплаты согласно данным приходят в 15-й и последний день месяца. По графику в начале и середине месяца как раз наблюдаются всплески.
- Выплаты зарплат, вероятнее всего, провоцируют увеличенный объем продаж в эти дни.



Анализ продуктовой корзины



Выводы:

- Анализ имеющихся данных позволяет сделать вывод о наличии сезонности в продажах: наибольший объём продаж наблюдается на выходных и в начале месяца.
- Результаты предсказаний модели позволяют оптимизировать объем закупок и сократить количество отходов
- Это, в свою очередь, позволит увеличить выручку за счёт оптимизации объема закупок: закупка товаров в нужном объеме и по мере необходимости.
- Информация о сроках годности товаров позволила бы учесть реализацию товаров по времени и скорректировать продажи путем дополнительных акций и скидок.

**Спасибо
за внимание!**