

# Анализ туристической индустрии российских отелей

Проект выполнили:

Воронина Дана Дмитриевна, ББИ221

Лисов Александр Владимирович, ББИ221

Смирнов Артём Дмитриевич, ББИ221

Хинтуба Кира Мурмановна, ББИ221



## Бизнес-цель

Анализ данных отелей, чтобы помочь бизнесу принимать более эффективные решения для получения высоких рейтингов на сайтах.

Целевая переменная - rating



---

---

Git Hub

там есть логирование



## Использование API с токеном. OpenTripMap

- API OpenTripMap дает доступ к данным о достопримечательностях, но содержит и информацию об отелях.
- Бесплатный API ключ предоставляет 5 000 запросов в день и 10 запросов в секунду.

>

2k  
строк

## Использование API в открытом доступе. Hotellook

- Обнаружили API-запрос к [https://search.hotellook.com/api/wl\\_search/result](https://search.hotellook.com/api/wl_search/result)  
Изучили параметры и воссоздали запрос через requests в Python.

>

40k  
строк

> 10k строк

## Использование Parsing Otello

- На первом этапе используется API, чтобы получить список URL-адресов веб-страниц.
- Рассмотрели отдельные страницы отелей.
- Нашли удобный способ парсить их.
- Воспользовались Selenium для получения URL-адресов. Сравнили методы.

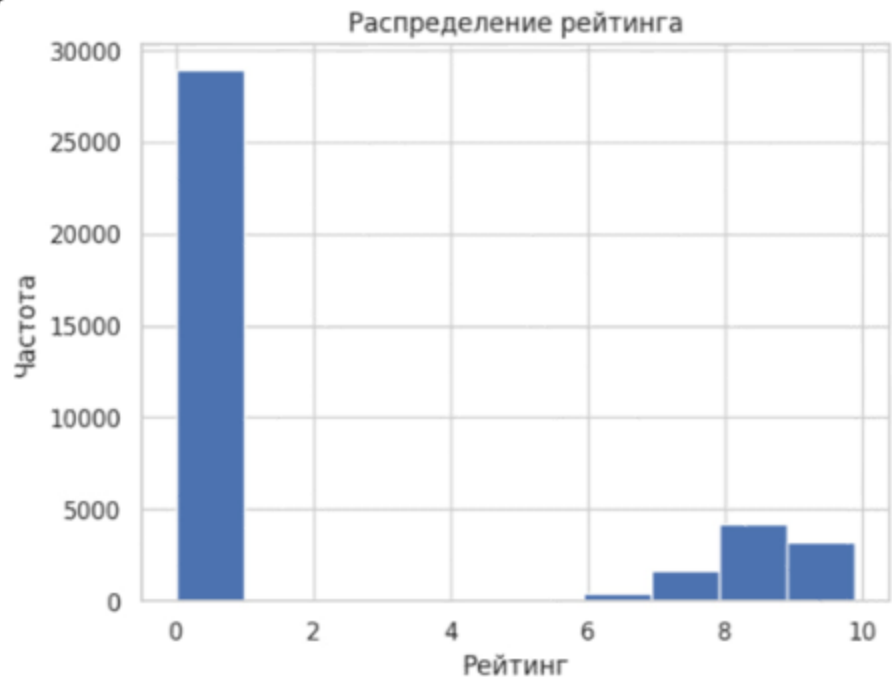
---

# EDA: предобработка

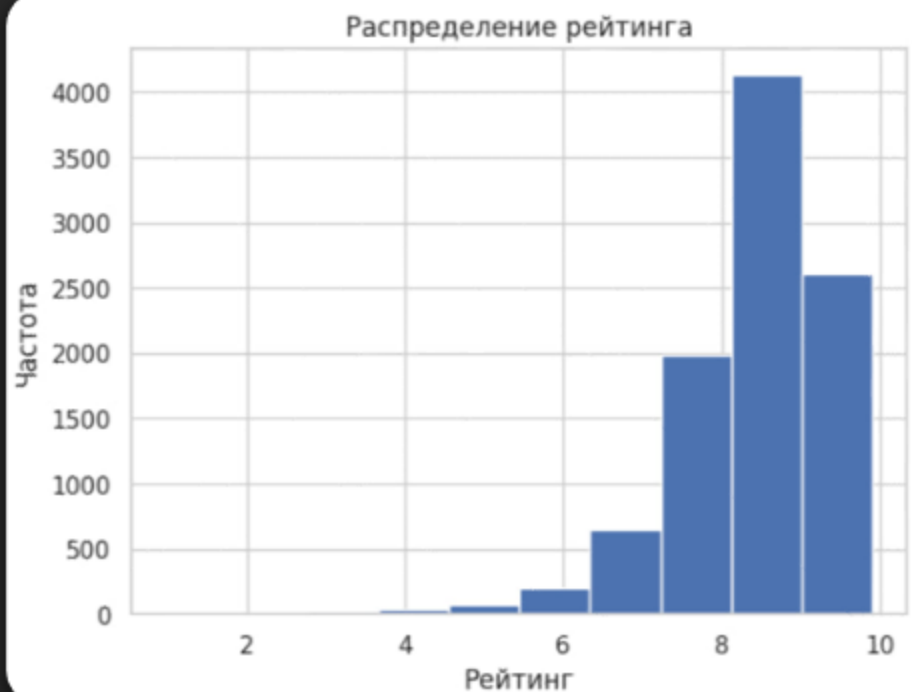
- Перекодировали значения
- Вытаскивали их из словарей и списков
- Удаляли/заполняли nan
- Удаляли нерелевантные признаки
- Формировали новые признаки
- Выявляли корреляции между признаками и target

# EDA

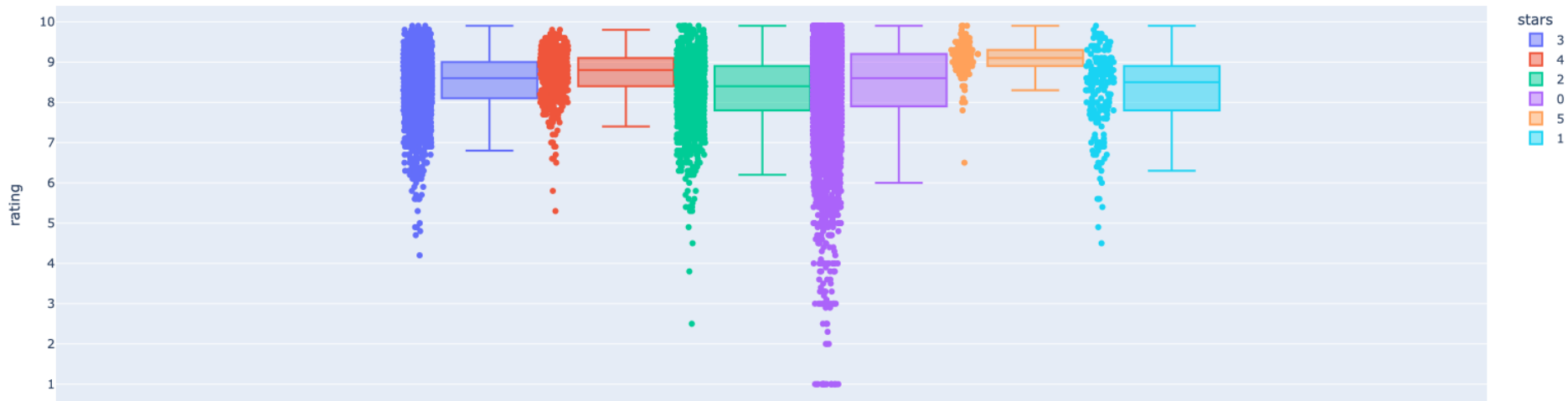
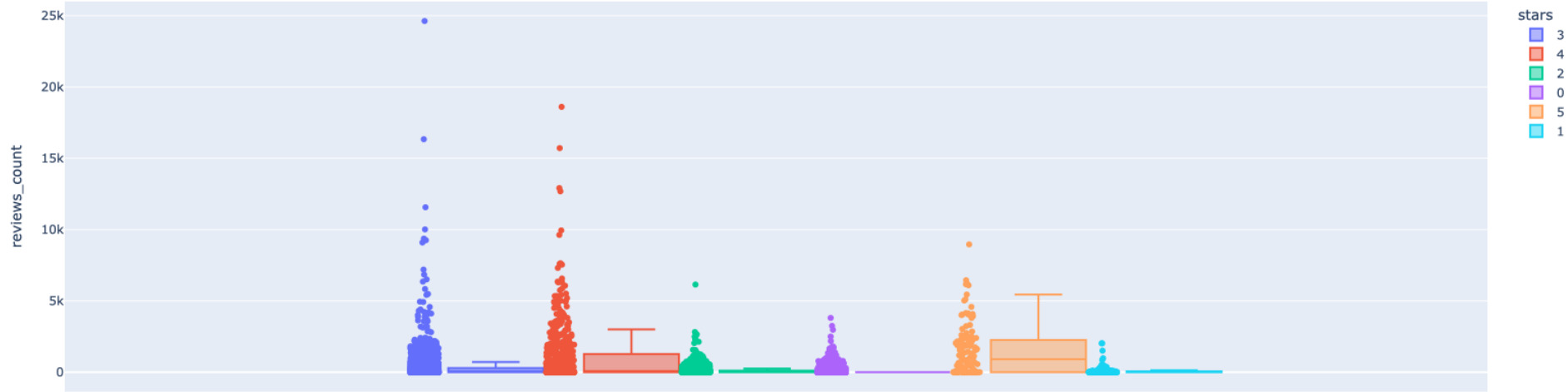
Обработали выбросы  
Удалили строки с пустыми name  
Удалили нерелевантные строки



Было



Стало





## График зависимости количества услуг от рейтинга, учитывая звезды

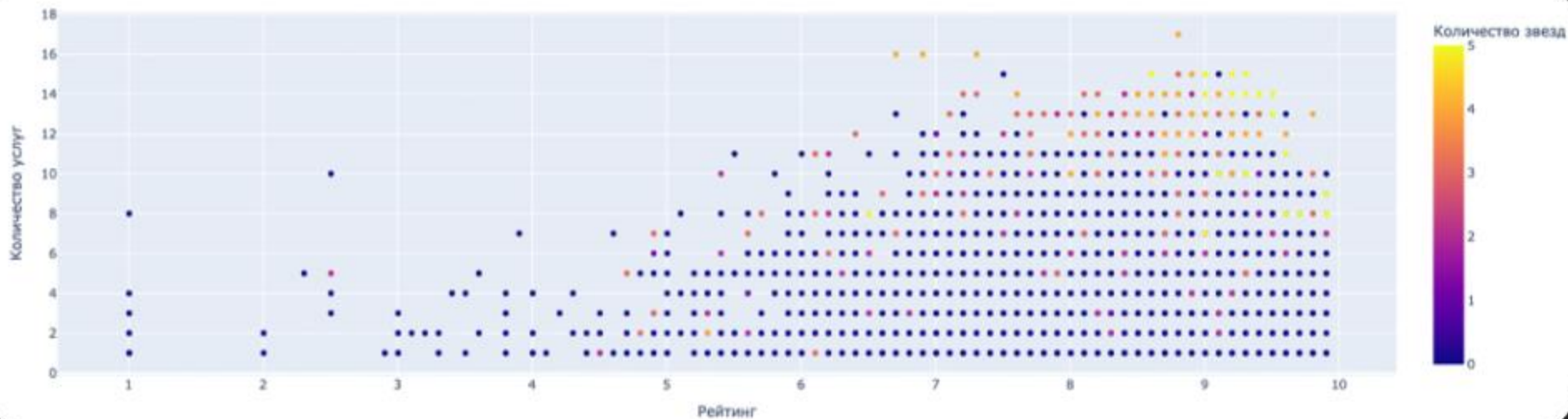


График распределения среднего рейтинга по услугам с учетом частотности услуги

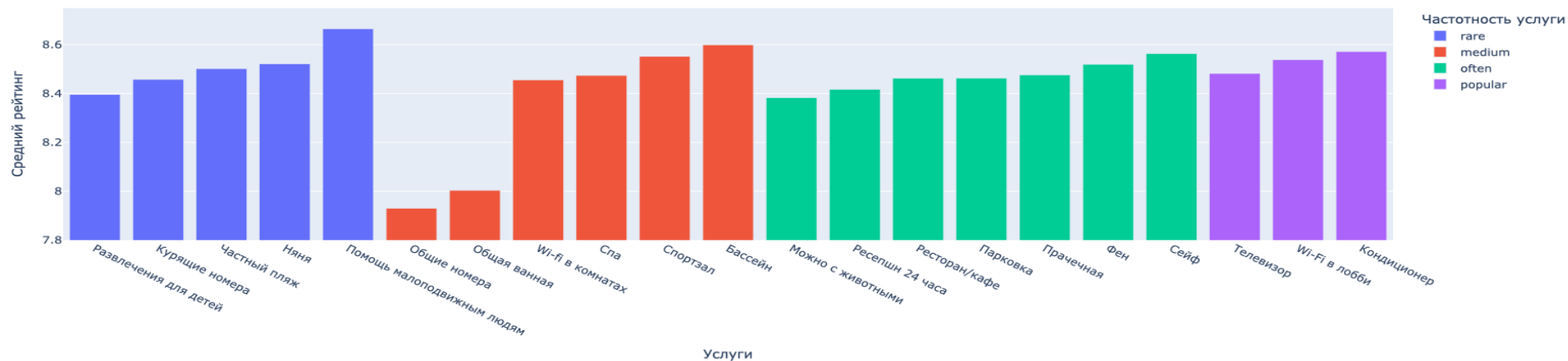
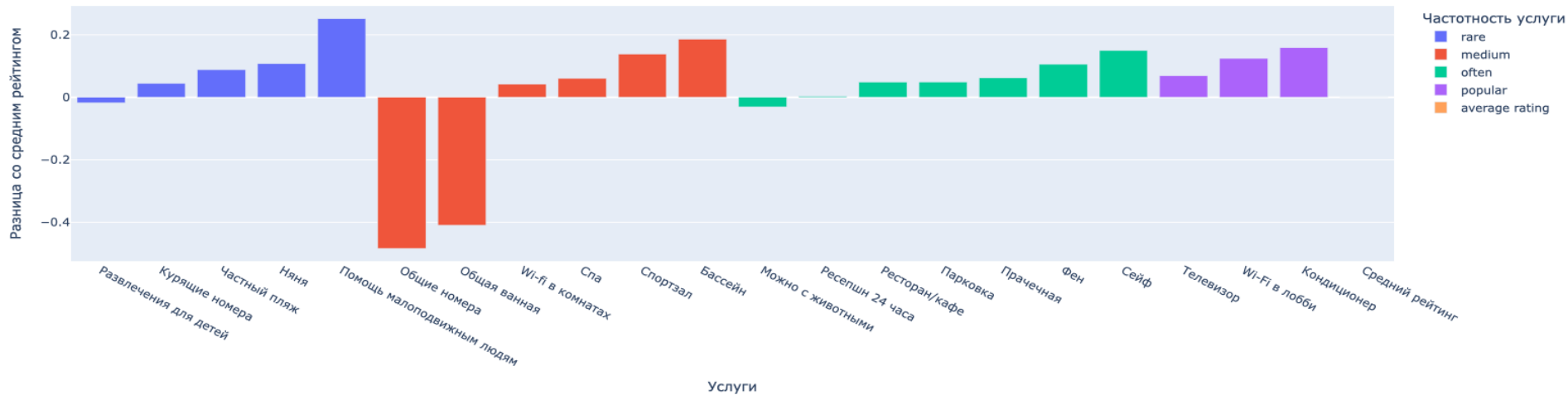
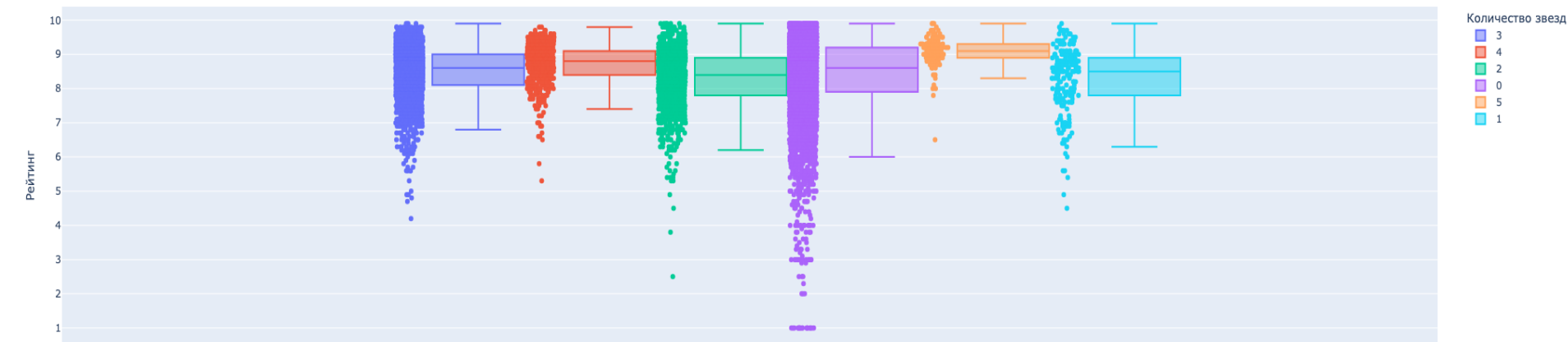
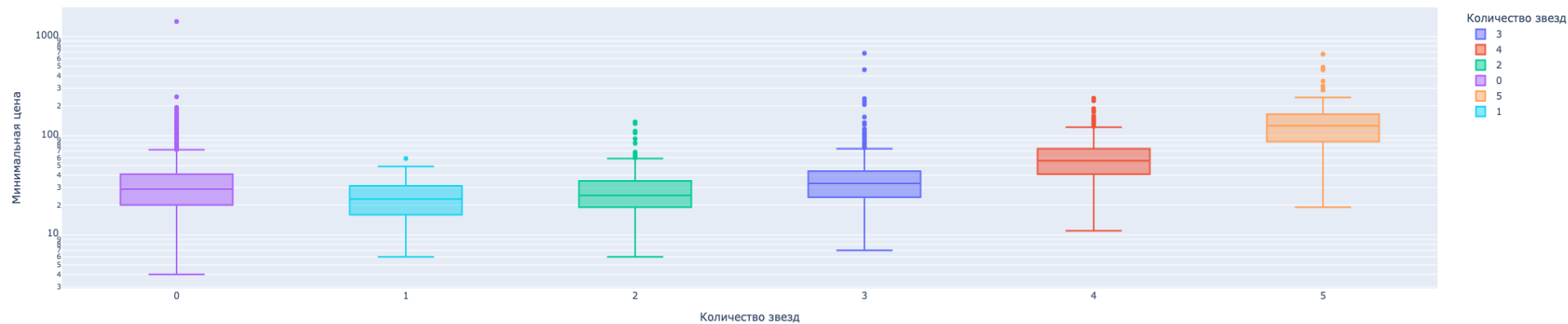


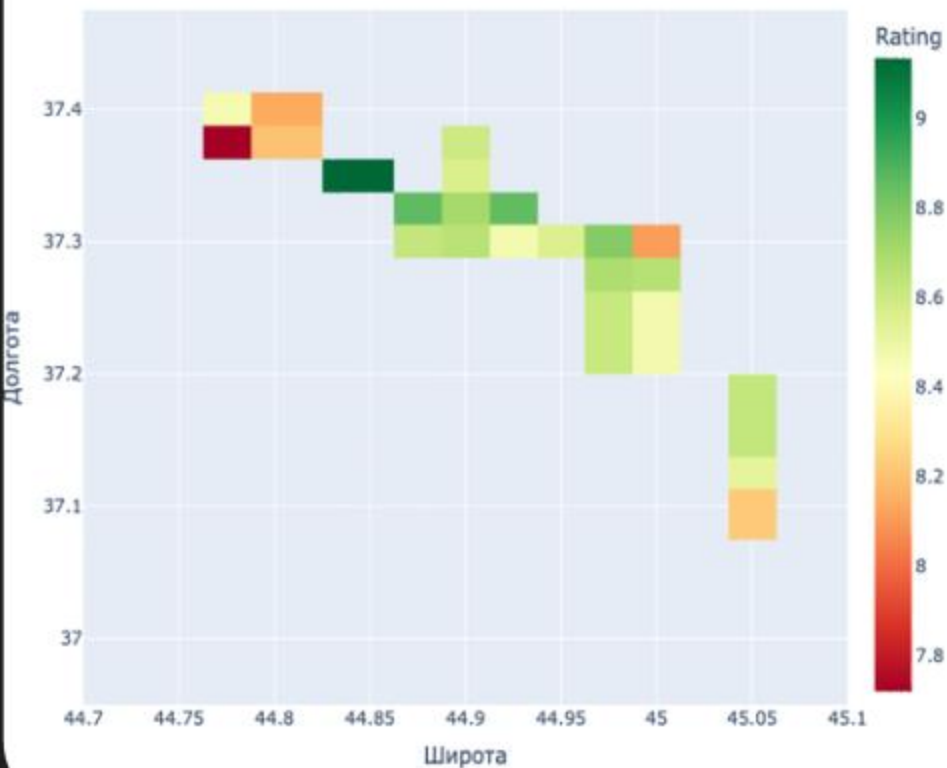
График разрыва между средним рейтингом и рейтингом по услугам с учетом частотности услуги



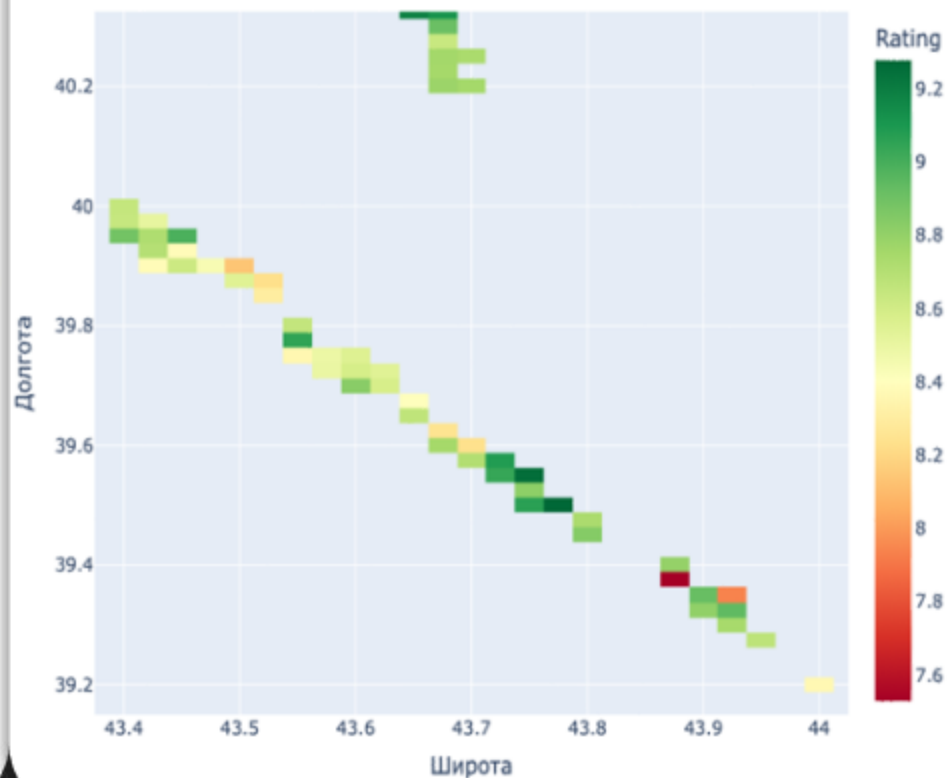
Распределение цен в зависимости от звездности отеля



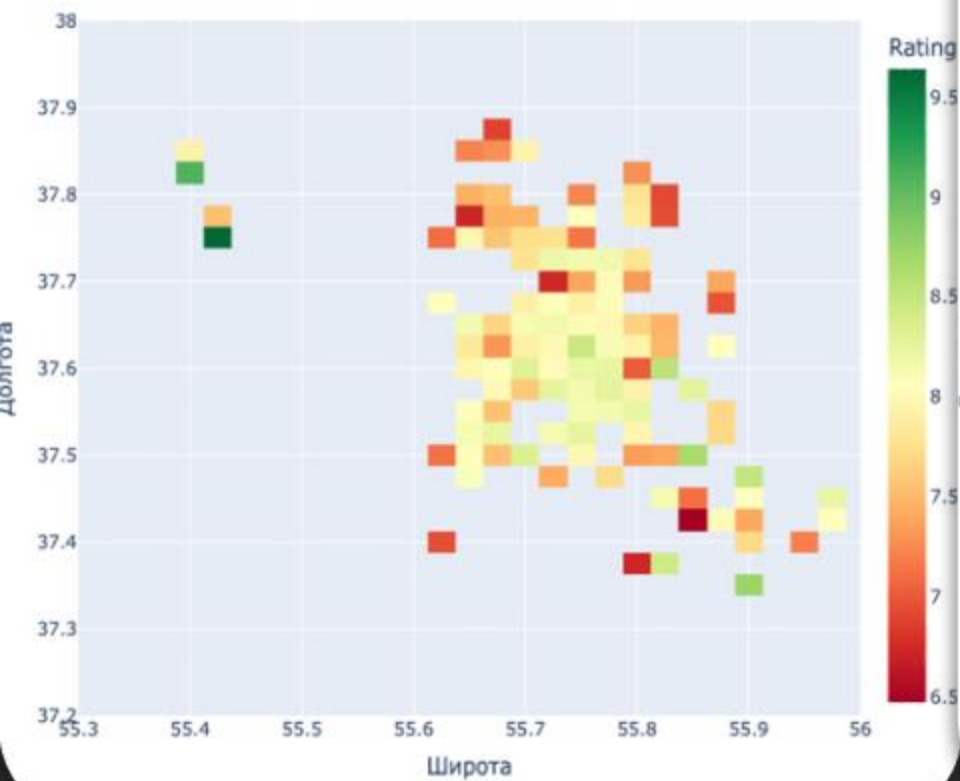
Тепловая карта рейтинга для Геленджика: средний рейтинг по районам



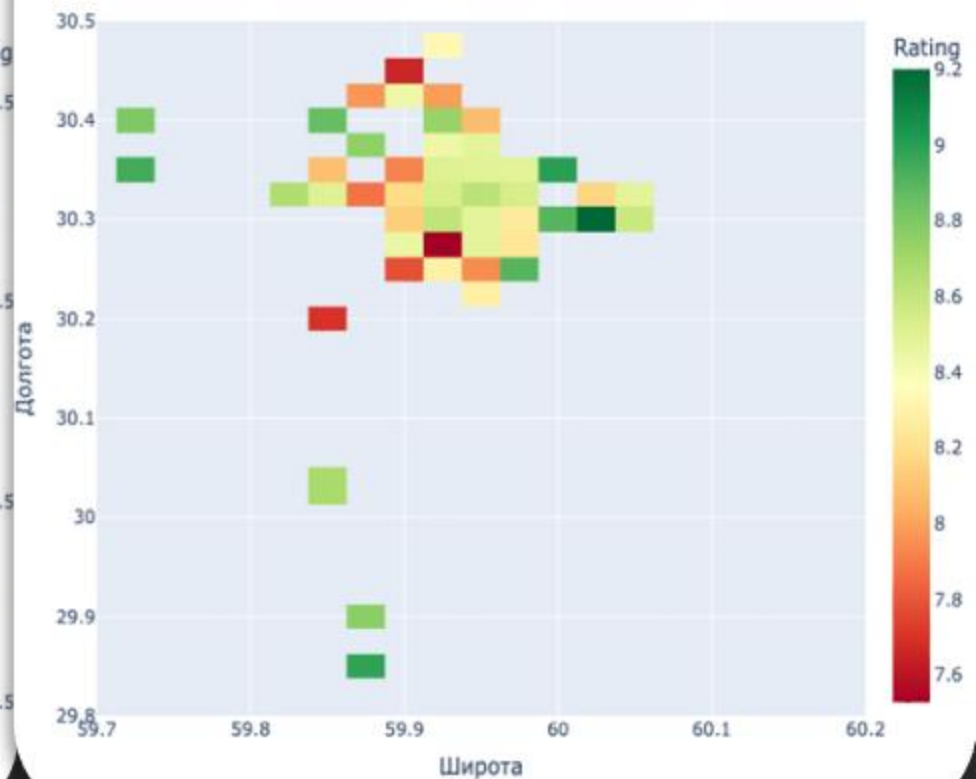
Тепловая карта рейтинга для Сочи: средний рейтинг по районам



Тепловая карта рейтинга для Москвы: средний рейтинг по районам



Тепловая карта рейтинга для Санкт-Петербурга: средний рейтинг по районам



# Выводы

## 1. Удобства и инфраструктура

- Некоторые услуги (общая ванная, общие номера) снижают рейтинг.
- Специализированные услуги (помощь маломобильным, бассейн) значительно улучшают оценку.
- Частота услуг не коррелирует с их влиянием на рейтинг.

## 2. Локация

- Отели в центральных или туристических (для приморских городов) районах получают более высокие оценки.
- Жилые районы ассоциируются с низкими рейтингами.
- На окраинах отели могут быть успешными за счёт дополнительных услуг или туристической привлекательности.

## 3. Рейтинг ≠ качество

- Рейтинг не всегда соответствует звёздности: 5-звездочные отели имеют разброс оценок из-за завышенных ожиданий.
- Высокая цена не гарантирует премиум-сервис: дорогие отели могут получать низкие оценки, а недорогие – высокие.

# Рекомендации

## Оптимизация набора услуг:

- Инвестировать в сервисы с выраженным положительным влиянием.
- Пересмотреть использование услуг с отрицательным влиянием и заменить их более качественными решениями.
- Сфокусироваться на качестве, а не на количестве предлагаемых удобств.

## Контроль за ожиданиями и ценообразованием:

- В 5-звездочных отелях уделять особое внимание управлению ожиданиями клиентов через сбор обратной связи и оперативное реагирование на отзывы.
- Обеспечение логики "высокая цена = реальное качество сервиса"

## Улучшение локационной привлекательности:

- Подчёркивать территориальные преимущества отелей, расположенных в центре, концентрировать маркетинг для них.
- Для объектов в жилых районах – разработать дополнительные услуги, чтобы компенсировать недостатки расположения.
- Для отелей на окраинах – использовать уникальные предложения для привлечения клиентов.