



На колёсах к успеху: исследуем рынок самокатов

Огурцов Константин

EDA – отчистим данные от мусора



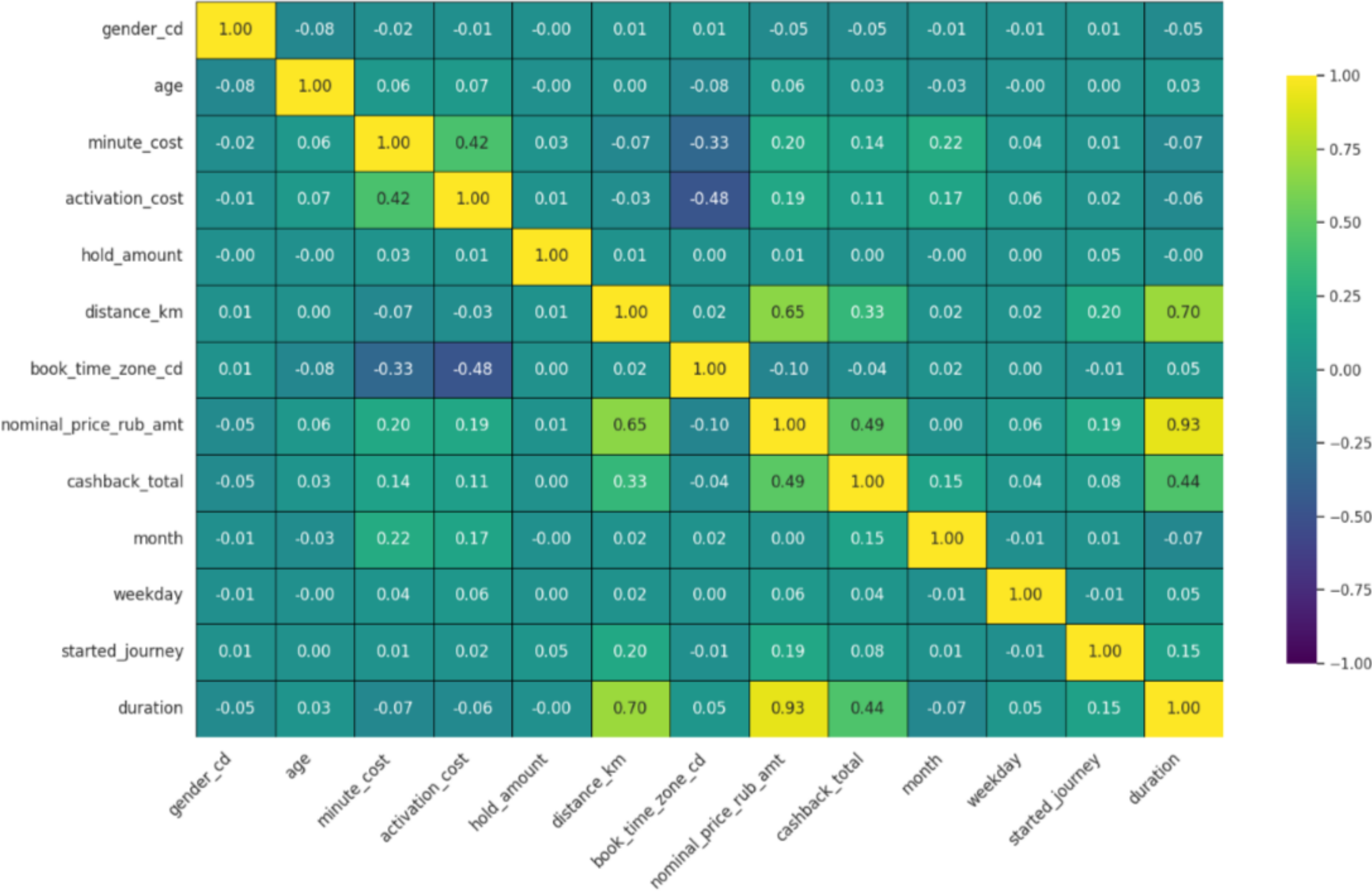
Избавились от данных:

- 1 Выбросы по пройденной дистанции (топ 1% верхних значений)
- 2 Выбросы по общему кэшбеку (топ 1% верхних значений)
- 3 Выбросы по ценам (топ 0,1% верхних значений)
- 4 Поездки без данных о гендере пользователя

Всего удалили
2,56% данных

Сделали тип гендера числовым параметром, добавили новые столбцы: продолжительность поездки, день недели и так далее

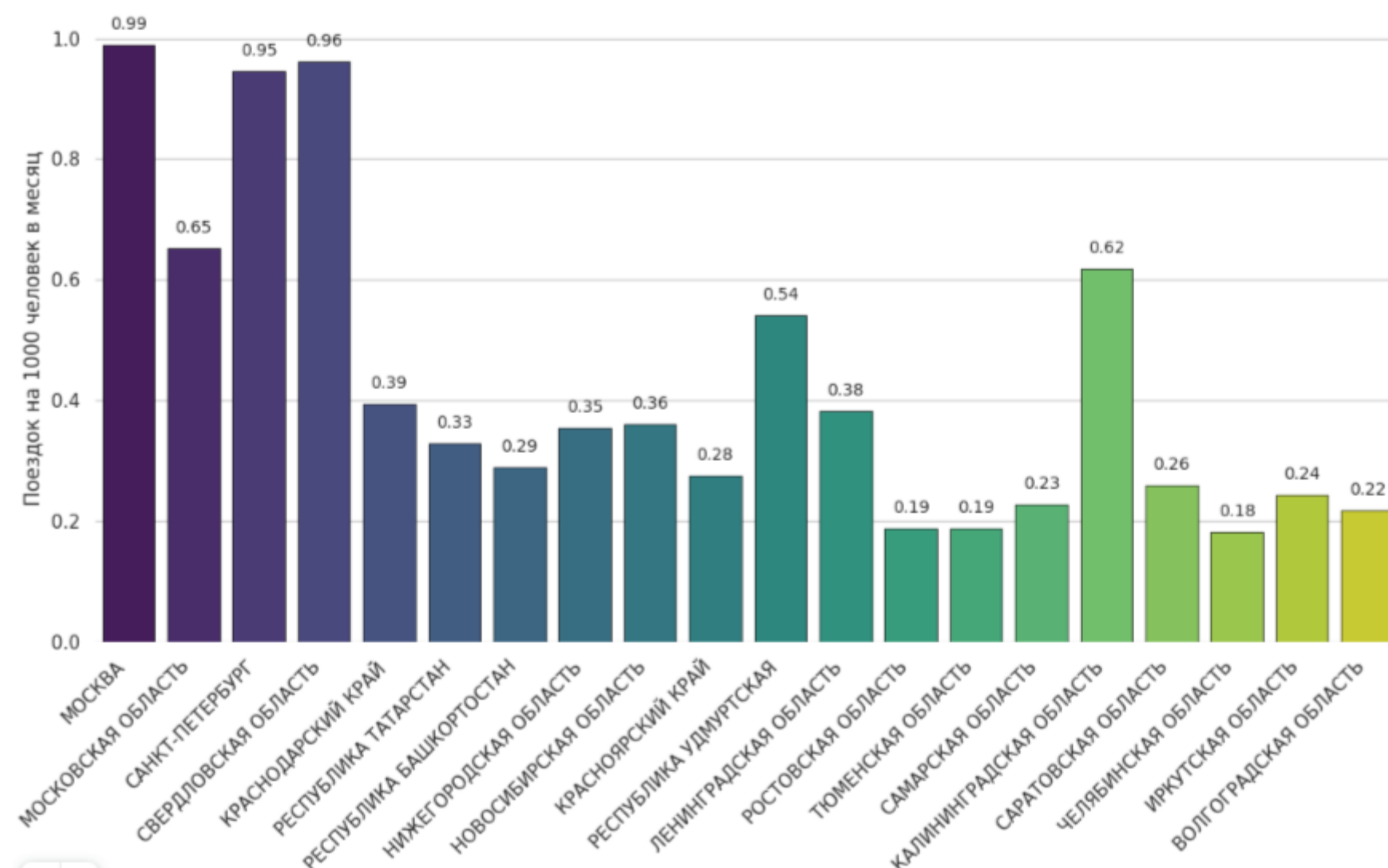
Корреляционная матрица признаков



Неочевидных корреляций признаков не выявлено

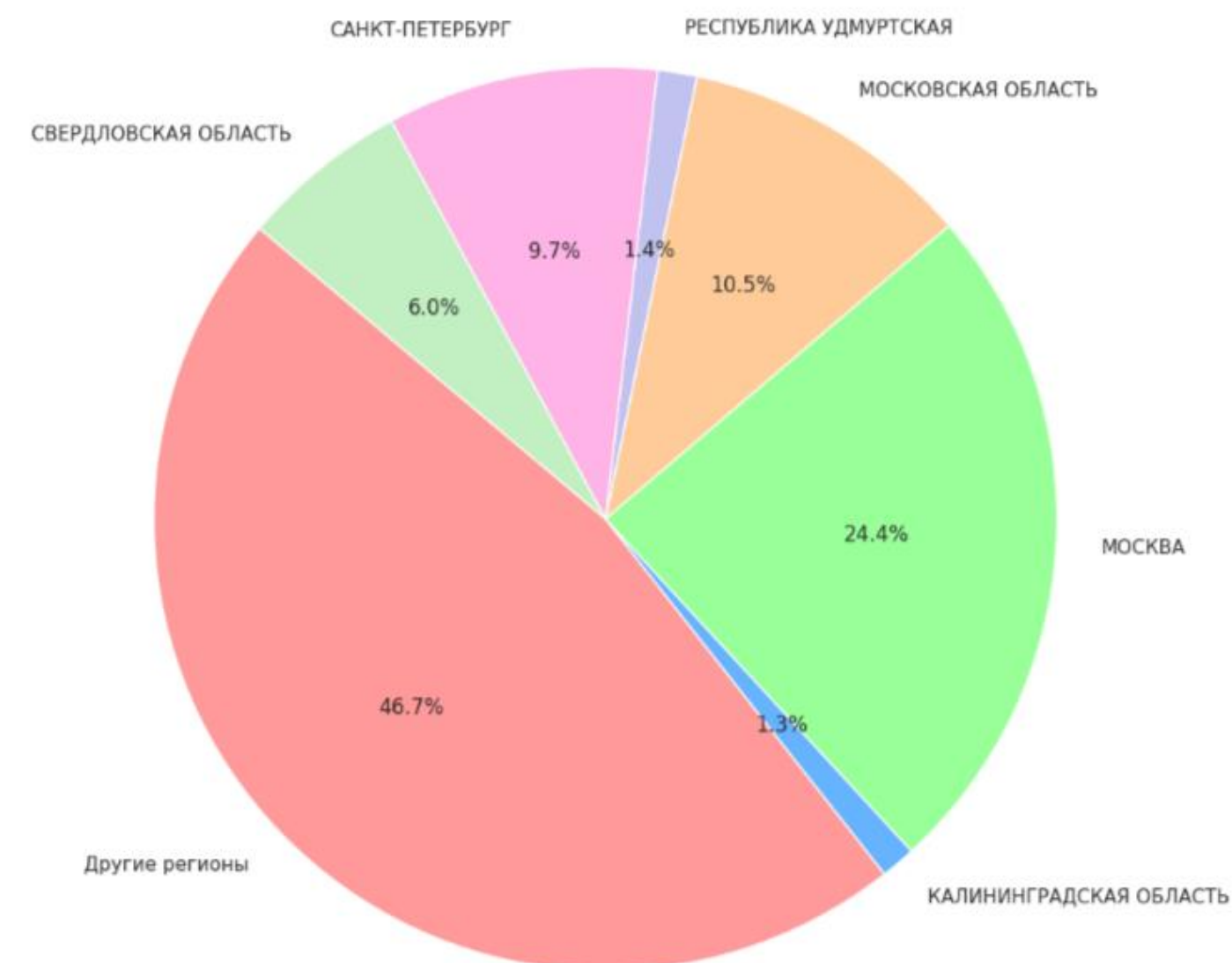
Ключевые графики

Среднее количество поездок в месяц на 1000 человек для основных регионов



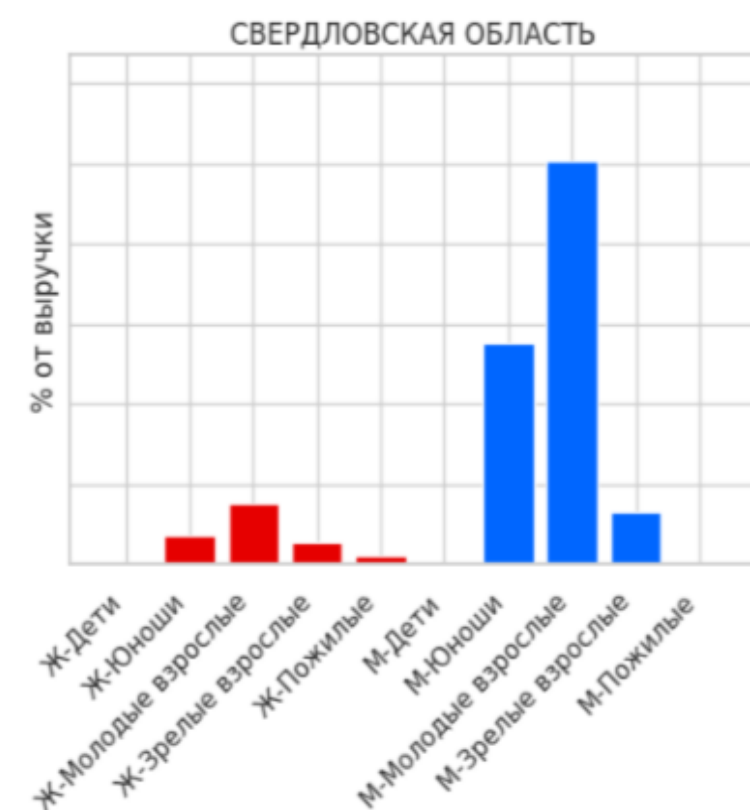
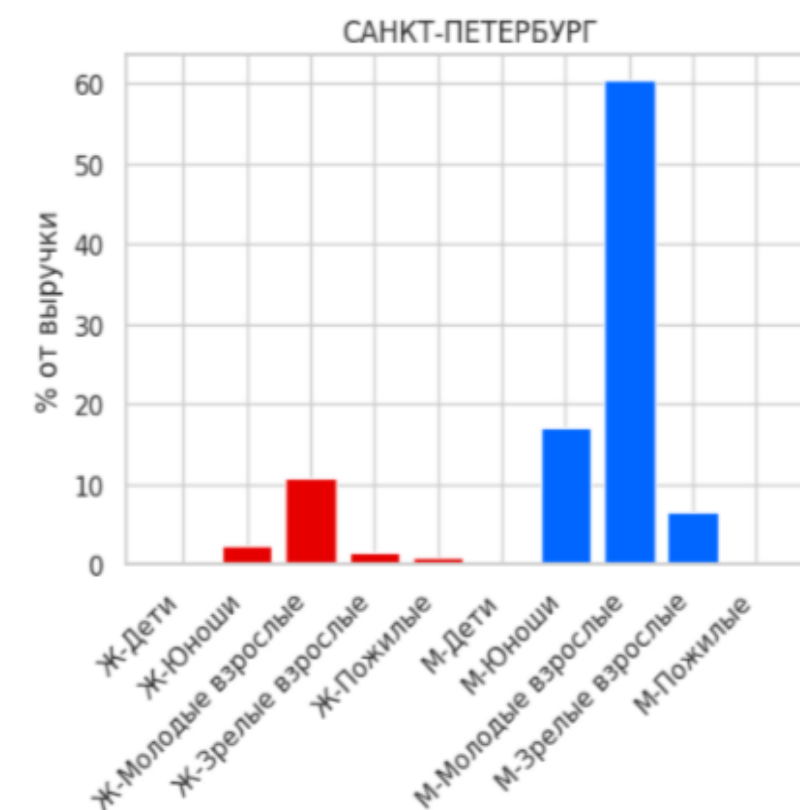
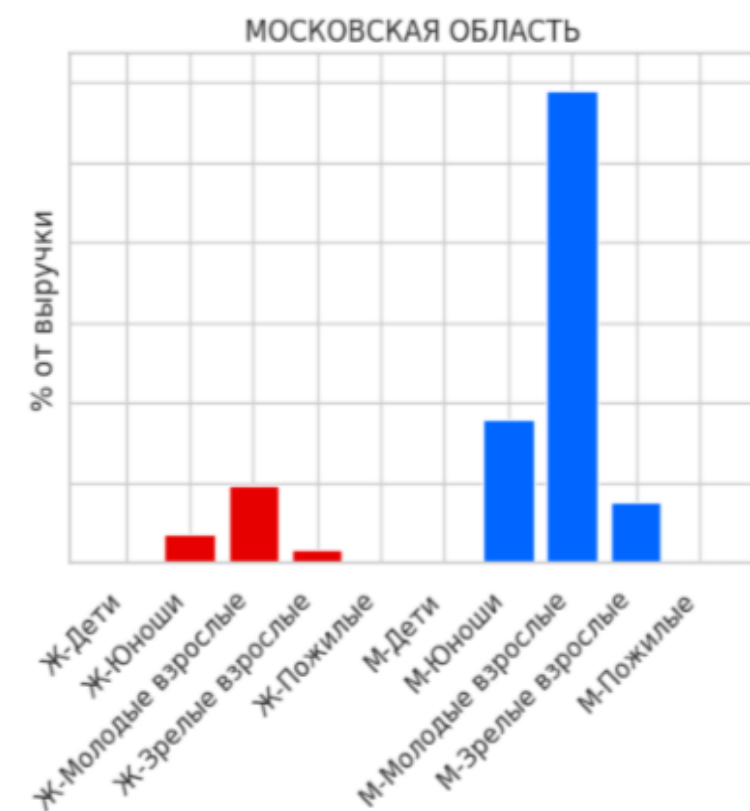
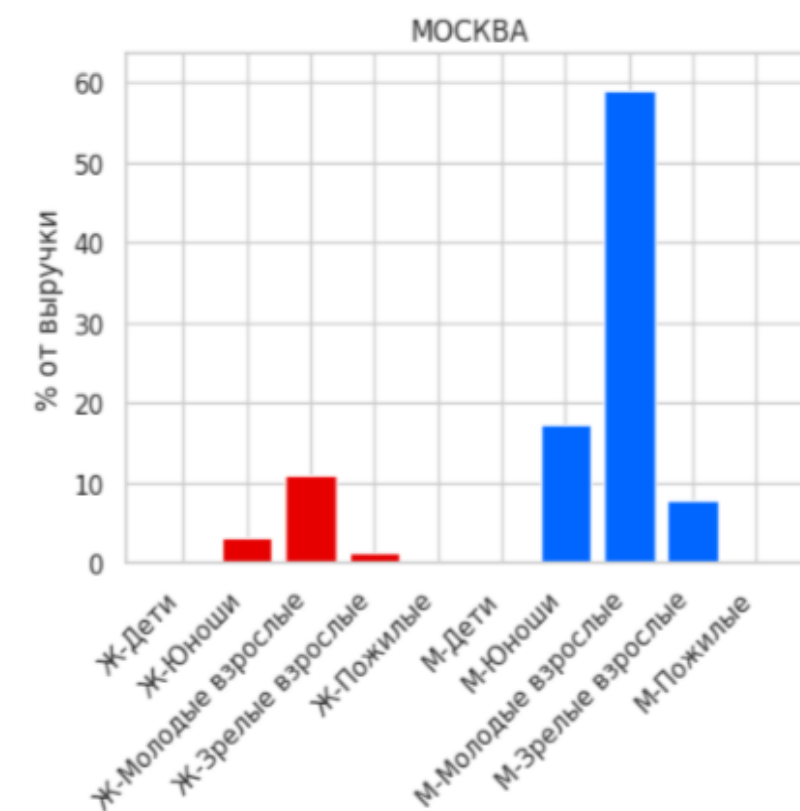
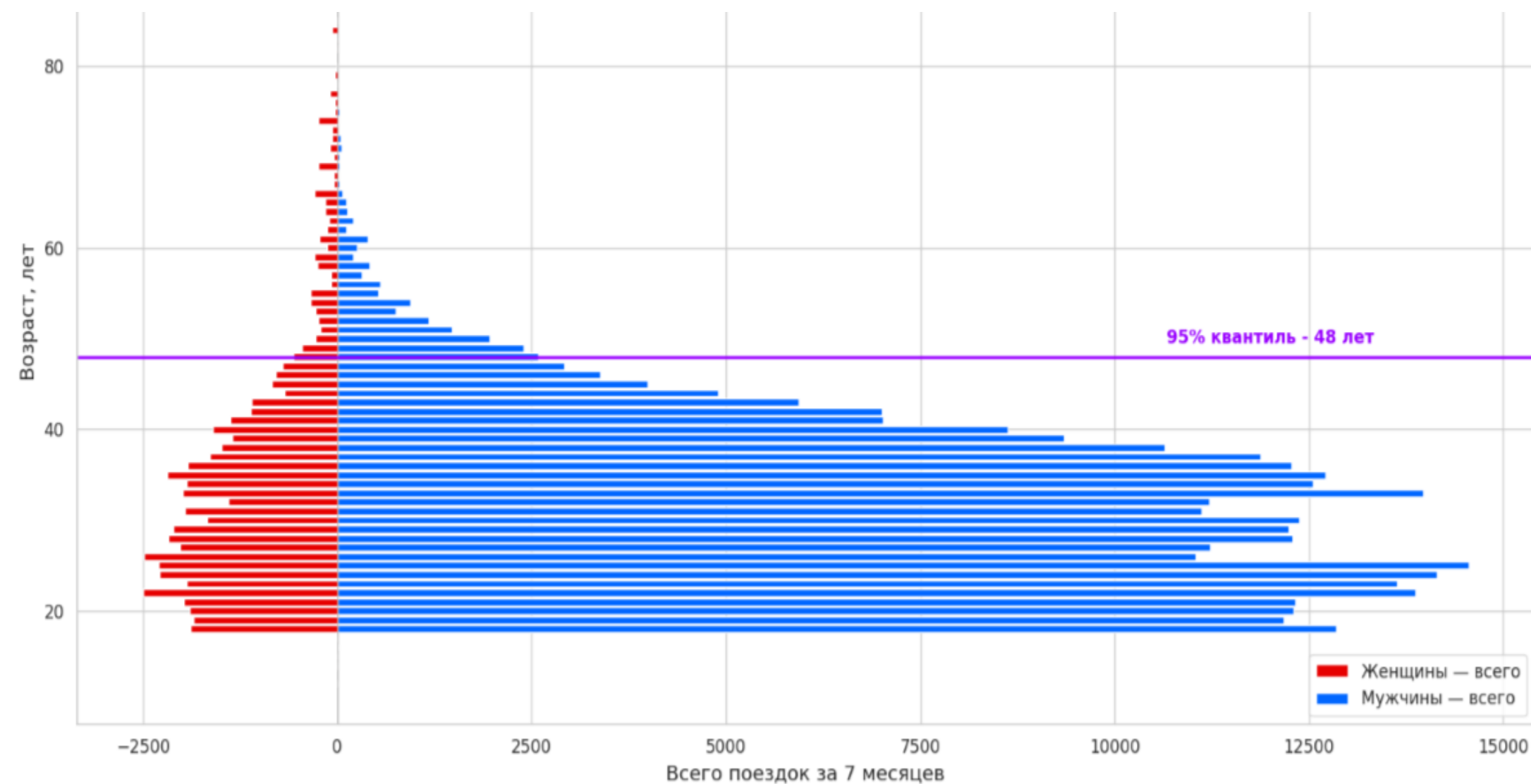
Самокаты как вид транспорта пользуется наибольшей популярностью в Москве, Московской области, Санкт-Петербурге, Свердловской области, Удмуртской республике и Калининградской области

Доля региона в общей выручке



Большую часть выручки приносят Москва, Московская область и Санкт-Петербург

Половозрастная структура



85% поездок совершают мужчины

95% поездок совершают люди возрастом до 48 лет

Продуктовые гипотезы

Первая гипотеза:

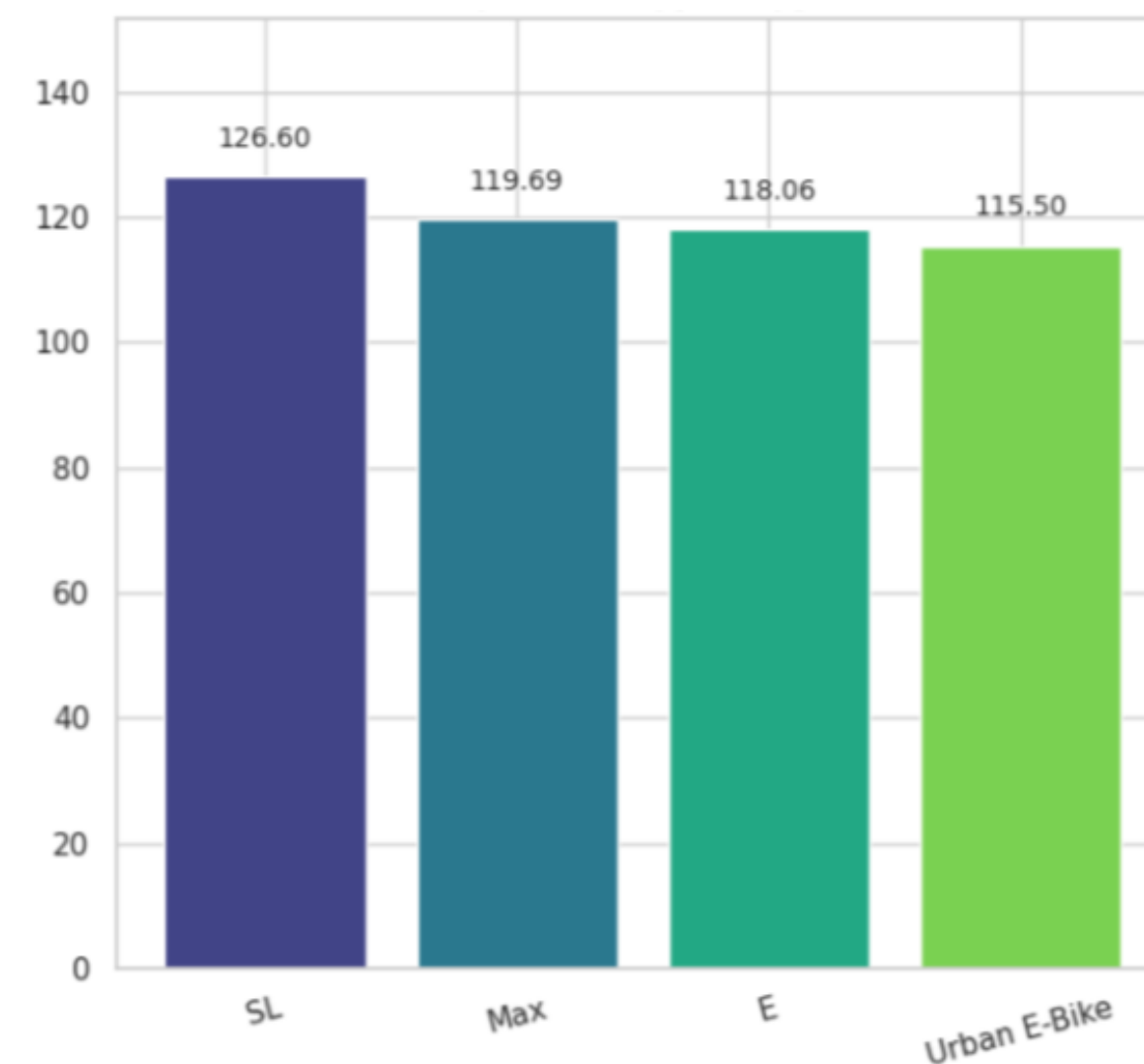
Если мы улучшим интерфейс приложения, добавив визуальные уведомления о низком заряде батареи и улучшим процесс активации самоката (например, через добавление инструкций или подсказок при проблемах с подключением), мы сможем снизить процент незавершенных поездок с 3% до 2,5%. Это приведет к увеличению количества завершенных поездок и повысит общую выручку на 0,25%.

3%
незавершенных
поездов

Вторая гипотезы:

Если самокаты модели Urban E-Bike заменить на модель SL в Москве, то средняя выручка от этих самокатов увеличится на 5,6%, а общая выручка по Москве за сезон вырастет на 0,15%.

Средняя выручка



Перспективная гипотеза

Я выбрал вторую гипотезу

Для уверенного подтверждения первой продуктовой гипотезы требуется много дополнительных данных о причинах о том, почему такое количество поездок остались незавершенными. В зависимости от причины нужно разрабатывать решение и дальше проводить А/Б тестирование или использовать другие способы проверки гипотезы. К тому же, велик шанс, что в некоторых случаях мы просто имеем некорректные данные.

Вторую гипотезу можно **подтвердить сразу** на имеющихся исторических данных, используя методы математической статистики (двухвыборочный Т критерий Стьюдента для проверки гипотезы о равенстве матожиданий). В результате, мы смогли отвергнуть нулевую гипотезу о том что матожидания выручки 2 моделей равны на уровне значимости 5%. Гипотеза подтверждена, значит, мы **сразу можем принять решение, повышающее прибыльность и эффективность бизнеса.**

Минимальный оцениваемый эффект
от решения, предлагающегося
второй гипотезой

8,96
рублей с каждой
поездкой

Изменения в выручке для самокатов
Urban E-Bike: **7.76%**

Изменения в выручке в Москве от
замены самокатов модели Urban E-
Bike на SL: **0.21%**

ИИ был использован для построения некоторых графиков, изменения стиля, а также для исправления ошибочного кода



That`s all

