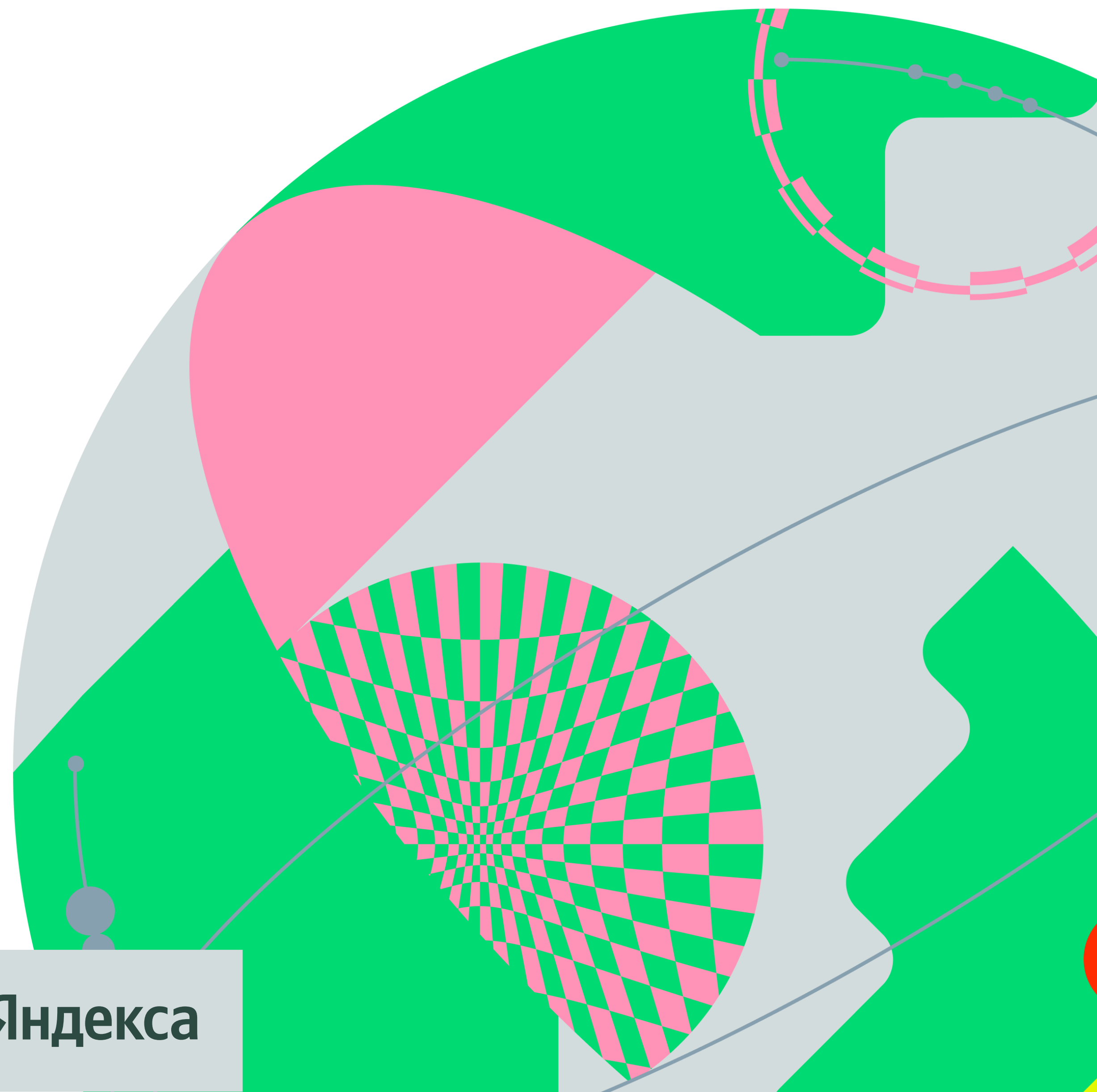


«Не изобретай велосипед!» - говорили они...

Продуктовая аналитика сервиса
по прокату велосипедов

Академия Яндекса



Это мы — «Колеса номер один»

2



Муляр Никита

Анализировал поведение пользователей, делал предсказания



Калугин Константин

Капитан команды, анализировал велосипедные станции, формулировал бизнес-задачи



Зейгман Константин

Дизайнер, анализировал передвижения велосипедов, внешние данные

Сервис Divvy

Divvy - это сервис проката велосипедов (электрических и классических) в городе Чикаго, США и его окрестностях

Функционирует с 2013 года. В 2017 и 2019 годах произошли самые массовые расширения сети.

🚲 ~ 5500 (2022)

👥 ~ 4M+ (2021)

📍 ~ 700 (2022)

** Информация из открытых источников*



Бизнес – дело непростое

1. Бизнес-цель	Чего хотим достичь в смысле бизнеса?	Выявить недостатки в управлении велосипедной сетью и придумать новые решения для компании
2. Бизнес-задача	Что нужно сделать для достижения поставленной цели?	Исследовательский проект
3. Бизнес-решения	Какие решения нужно предпринять для выполнения задачи?	<i>Это и предстоит узнать!</i>

Пайплайн исследования

01



02



03



04



Разведочный анализ данных

- Исследование структуры данных
- Нормализация данных
- Изучение зависимостей

Глубокий анализ сегментов сервиса

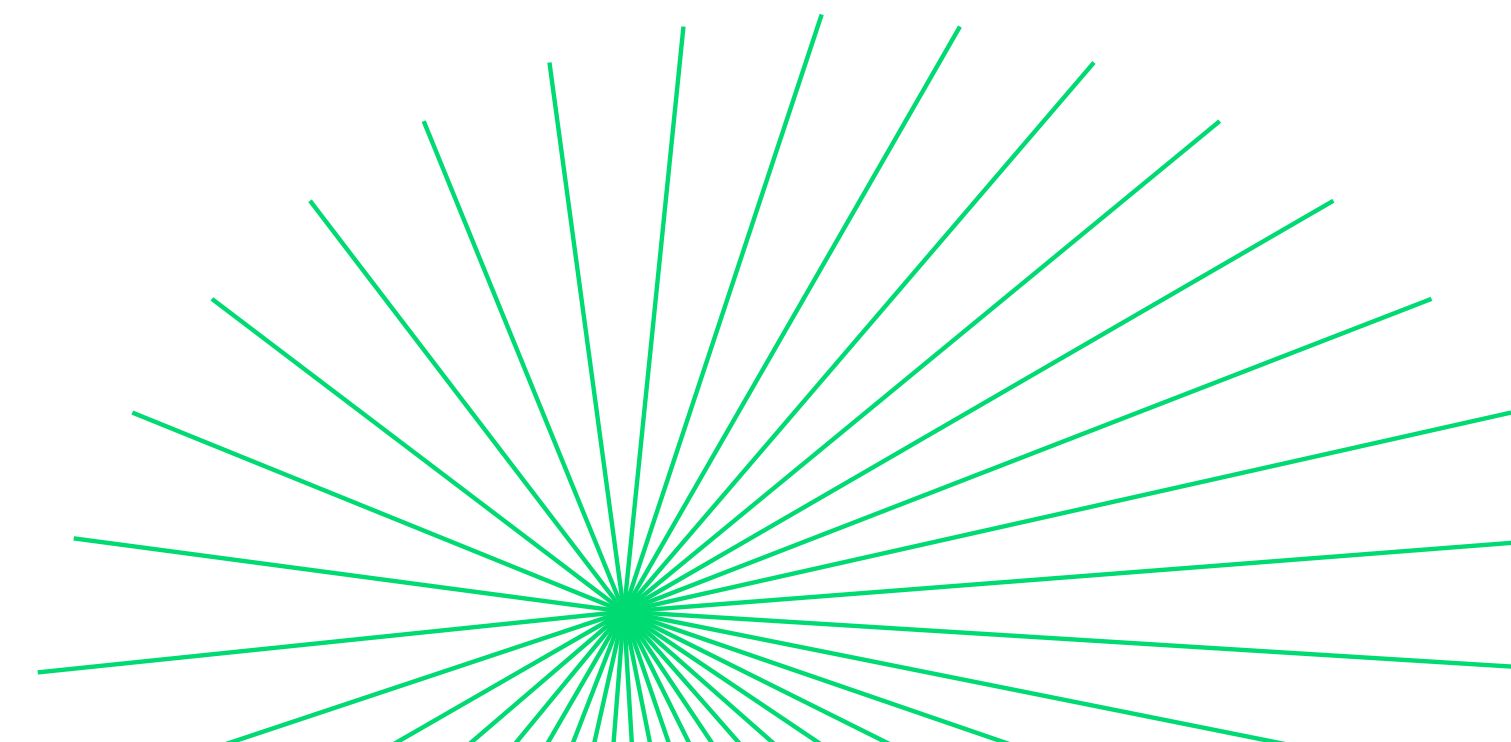
- Определение портрета целевой аудитории
- Анализ типов велосипедов

Прогнозы и анализ внешних данных

- Расчет экономики на 2023 год
- Прогнозирование данных
- Анализ внешних данных

Выводы и предложения

- Подведение итогов
- Предложения



Милльон терзаний от датасета

Name

202004-divvy-tripdata.zip

202005-divvy-tripdata.zip

.....

Divvy_Trips_2017_Q1Q2.zip

Divvy_Trips_2017_Q3Q4.zip

Divvy_Stations_2013.csv

Divvy_Stations_2013.shp.zip

Divvy_Trips_2013.csv

README.txt

01 Стандарты хранения данных изменились в 2020 году

02 Информация об остановках может храниться отдельно от поездок

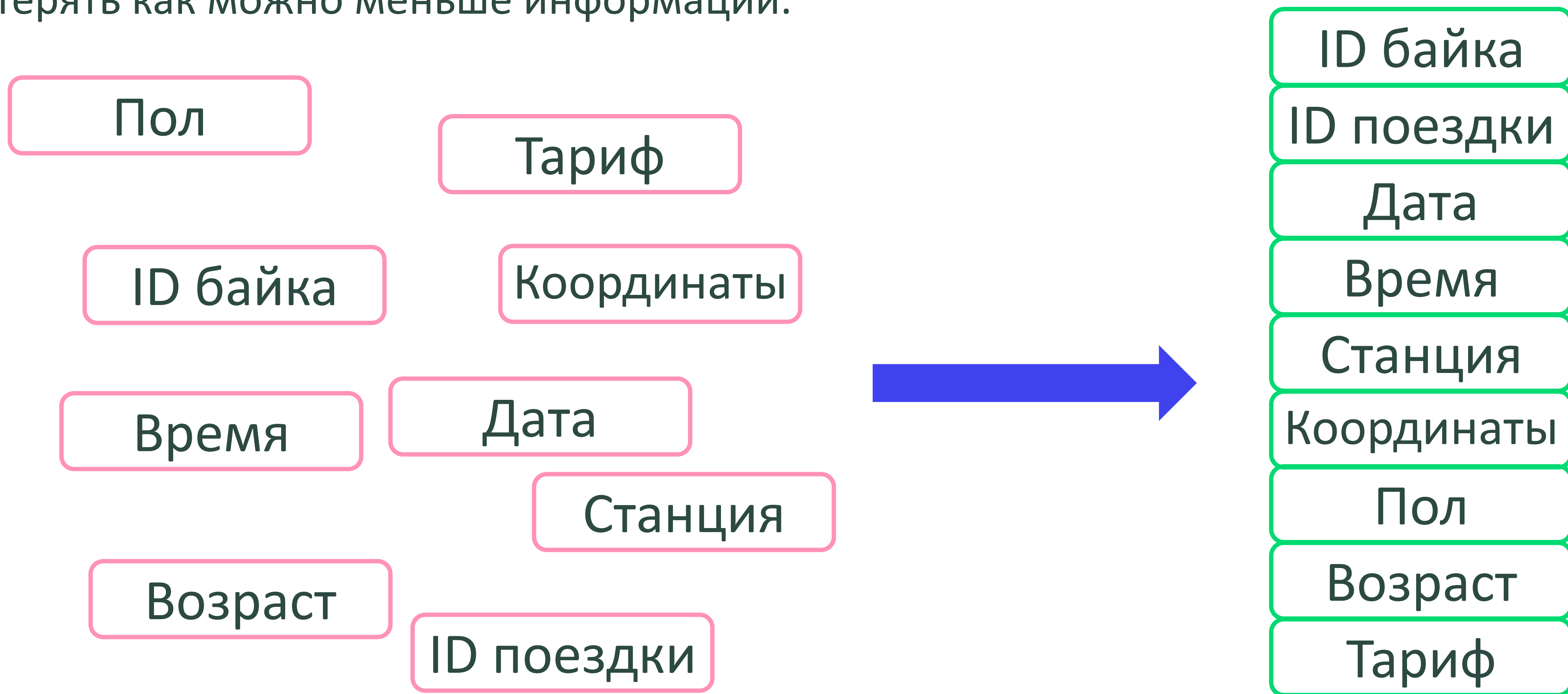
03 Разные количество и порядок колонок

04 Много пропусков

Как нормировать данные, чтобы с легкостью обрабатывать их?

Бигдательский Minecraft

Цель: собрать витрины данных по годам с файлов, которые имеют разный формат и потерять как можно меньше информации.

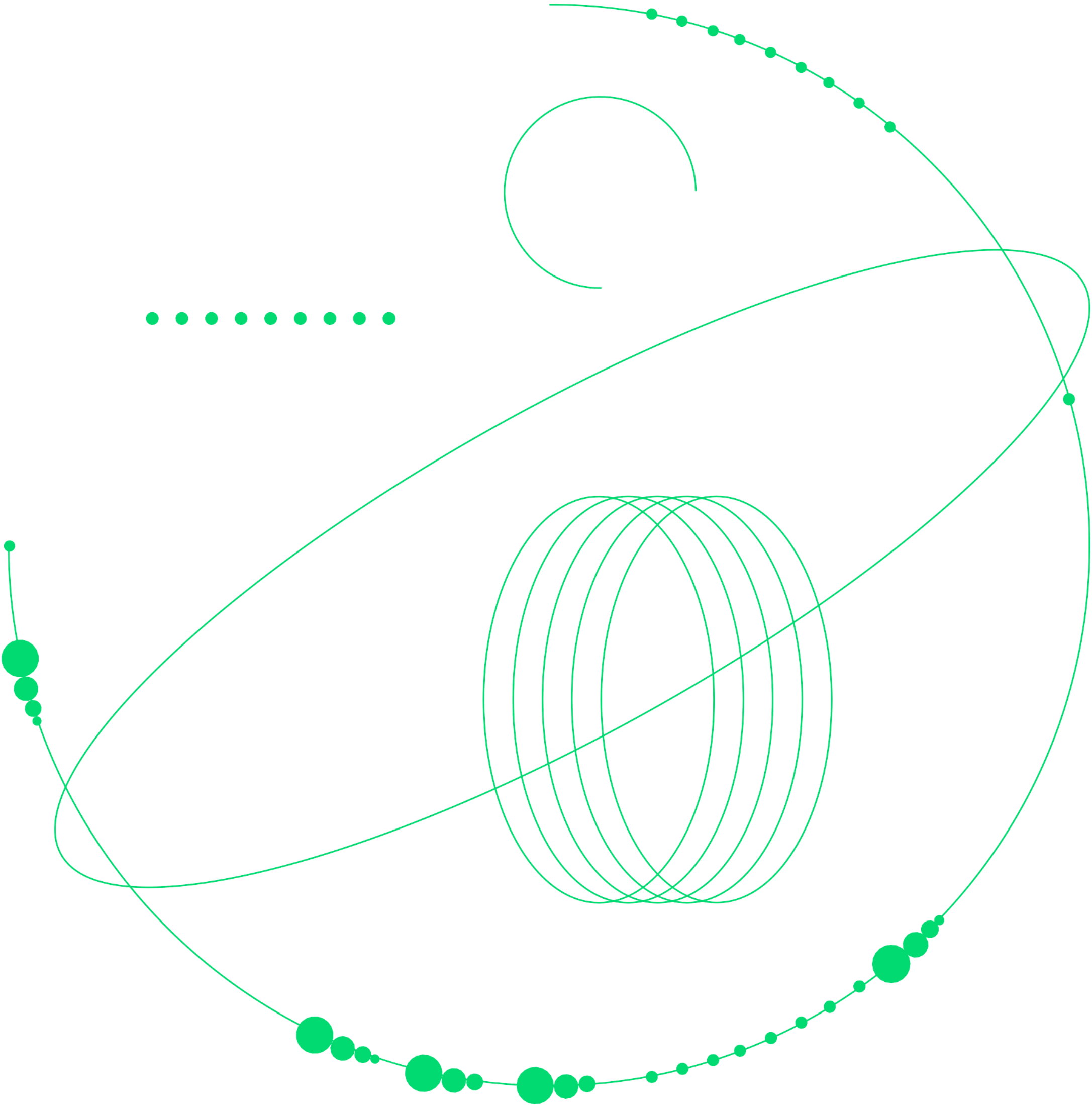


Аналитики не сдаются

2013						
	id	biketype	starttime		endtime	\
0	8163	NaN	2013-06-29T13:22:00	2013-06-29T13:45:00		
1	8164	NaN	2013-06-29T13:22:00	2013-06-29T13:45:00		
2	8826	NaN	2013-06-29T16:40:00	2013-06-29T17:19:00		
3	8827	NaN	2013-06-29T16:41:00	2013-06-29T17:19:00		
4	8855	NaN	2013-06-29T16:52:00	2013-06-29T17:27:00		
	startstation		endstation	startlat	startlng	\
0	Lake Shore Dr & Monroe St		Michigan Ave & Oak St	41.880958	-87.616743	
1	Lake Shore Dr & Monroe St		Michigan Ave & Oak St	41.880958	-87.616743	
2	Lake Shore Dr & Monroe St		Michigan Ave & Oak St	41.880958	-87.616743	
3	Lake Shore Dr & Monroe St		Michigan Ave & Oak St	41.880958	-87.616743	
4	Lake Shore Dr & Monroe St		Michigan Ave & Oak St	41.880958	-87.616743	
	endlat	endlng	member_casual	gender	age	bikeid
0	41.90096	-87.623777	Customer	NaN	NaN	153
1	41.90096	-87.623777	Customer	NaN	NaN	485
2	41.90096	-87.623777	Customer	NaN	NaN	417
3	41.90096	-87.623777	Customer	NaN	NaN	567
4	41.90096	-87.623777	Customer	NaN	NaN	49

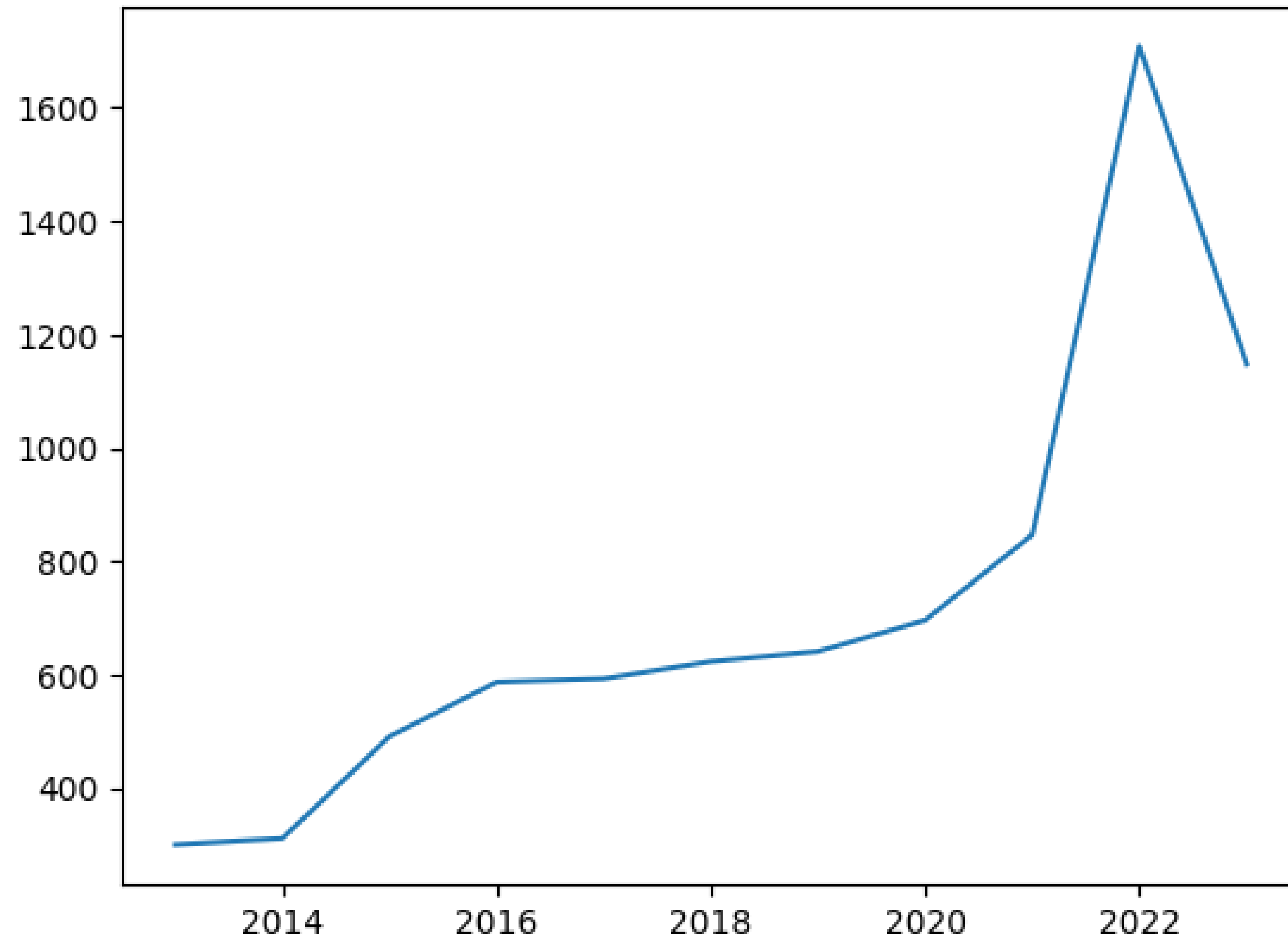
Структура базы за 2013 год

- 1. Подобрать алгоритм обработки для каждого года
- 2. Найти пересекающиеся колонки
- 3. Объединить данные в один файл (по годам)
- 4. Присоединить станции, где необходимо

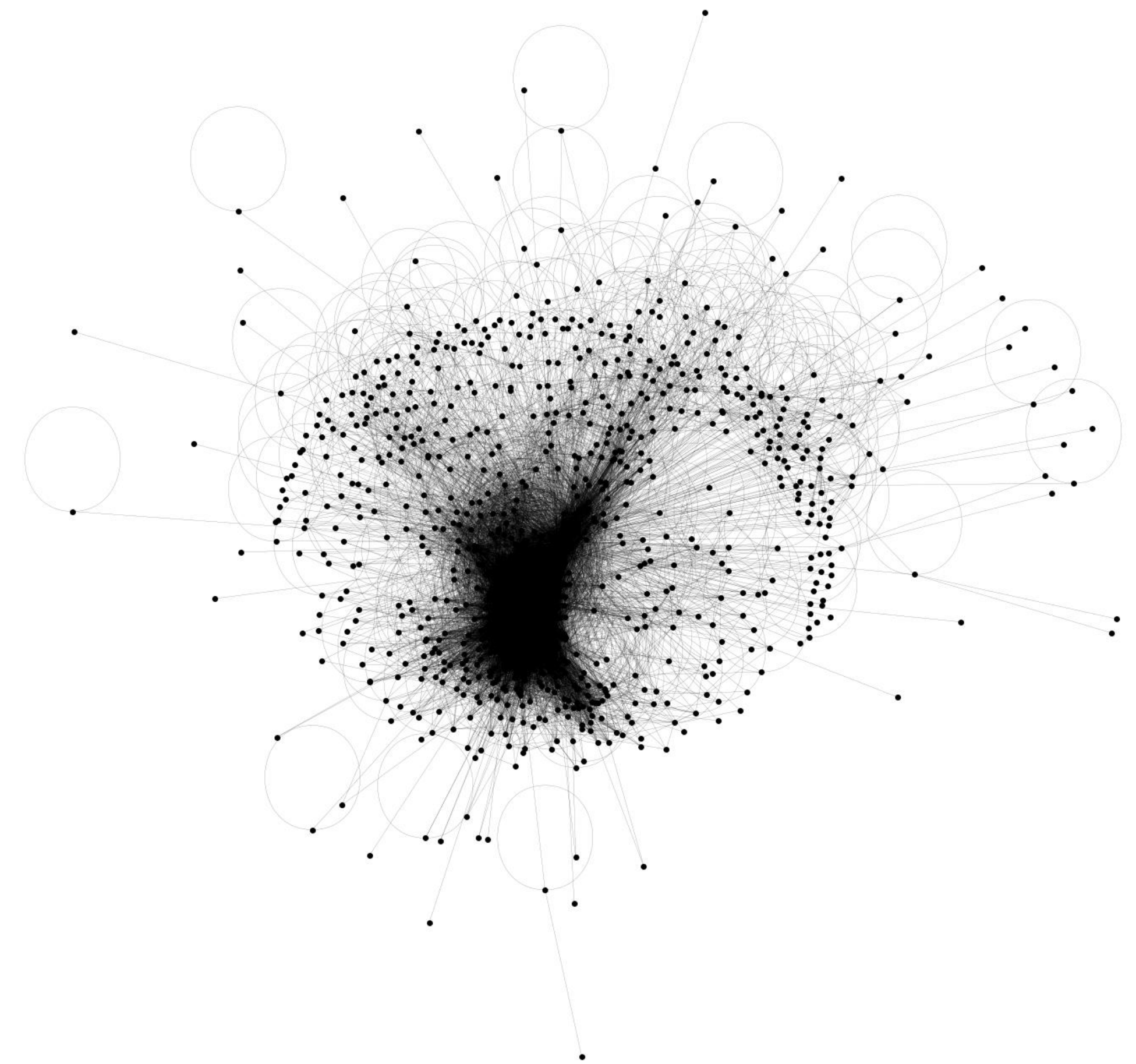


Разведка докладывает...

Количество станций по годам

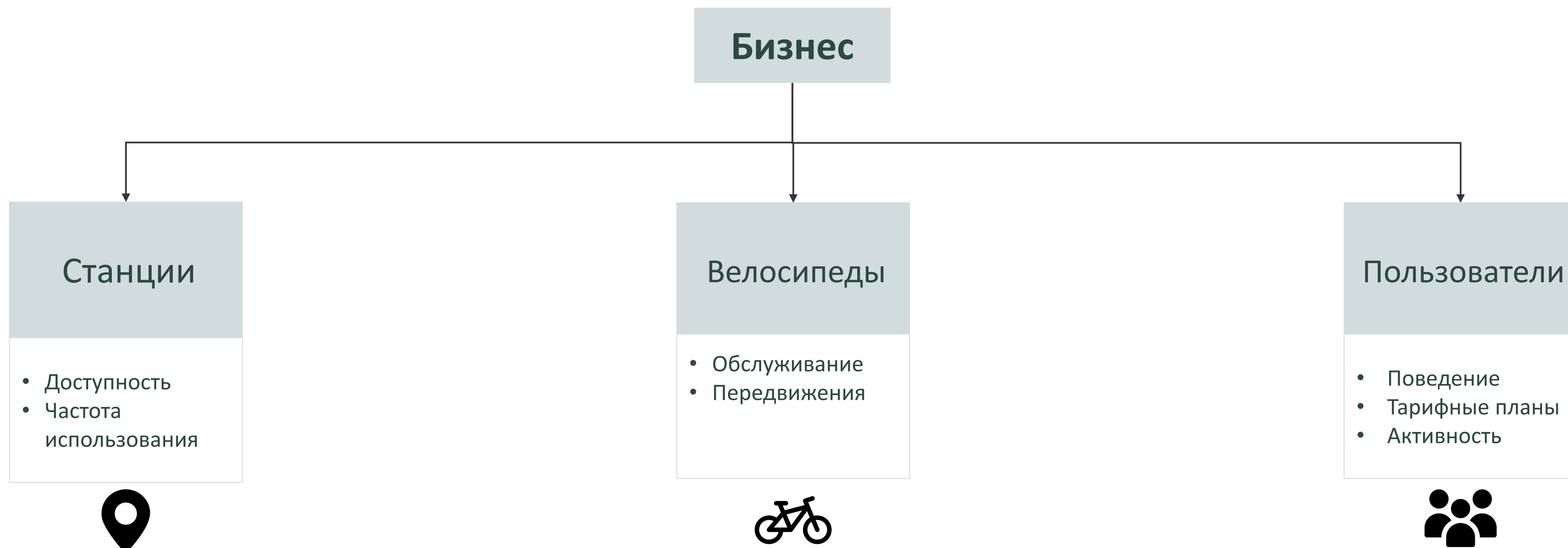


Граф передвижений за 2023 год



Анализ данных 2.0

Цель: исходя из результатов первичного анализа, изучить более детально каждый сегмент бизнеса по отдельности.



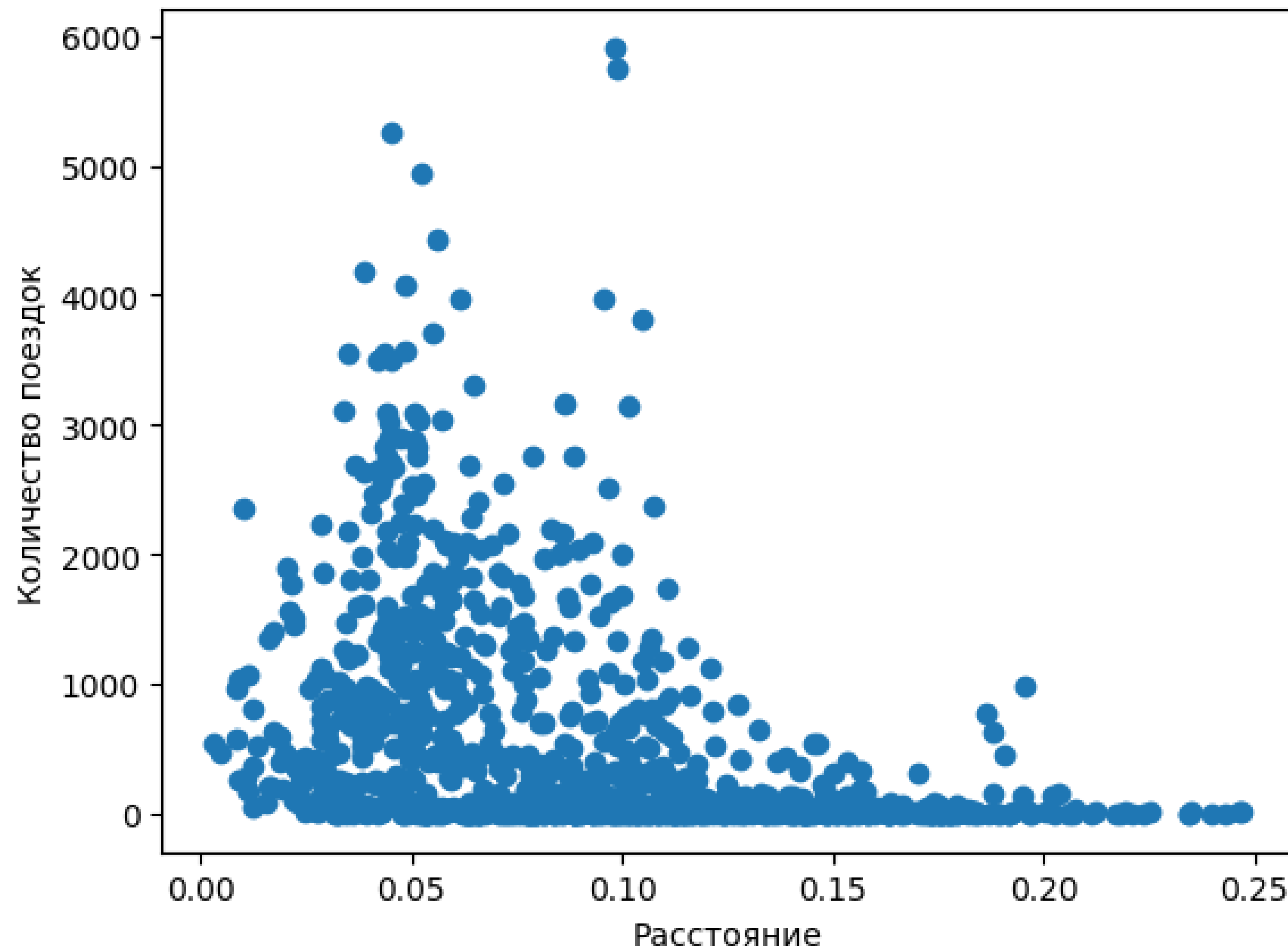
Из точки А в точку В

11

	Станция	Кол-во	Тип
	University Ave & 57th St	5908	classic_bike
	Ellis Ave & 60th St	5744	classic_bike
	Clinton St & Washington Blvd	5250	electric_bike
	Kingsbury St & Kinzie St	4932	classic_bike
💰	Clark St & Elm St	4423	classic_bike
	Canal St & Adams St	4184	classic_bike
	Clinton St & Madison St	4076	electric_bike
	Ellis Ave & 55th St	3981	classic_bike
	State St & Chicago Ave	3979	classic_bike
💰	Wells St & Concord Ln	3815	classic_bike



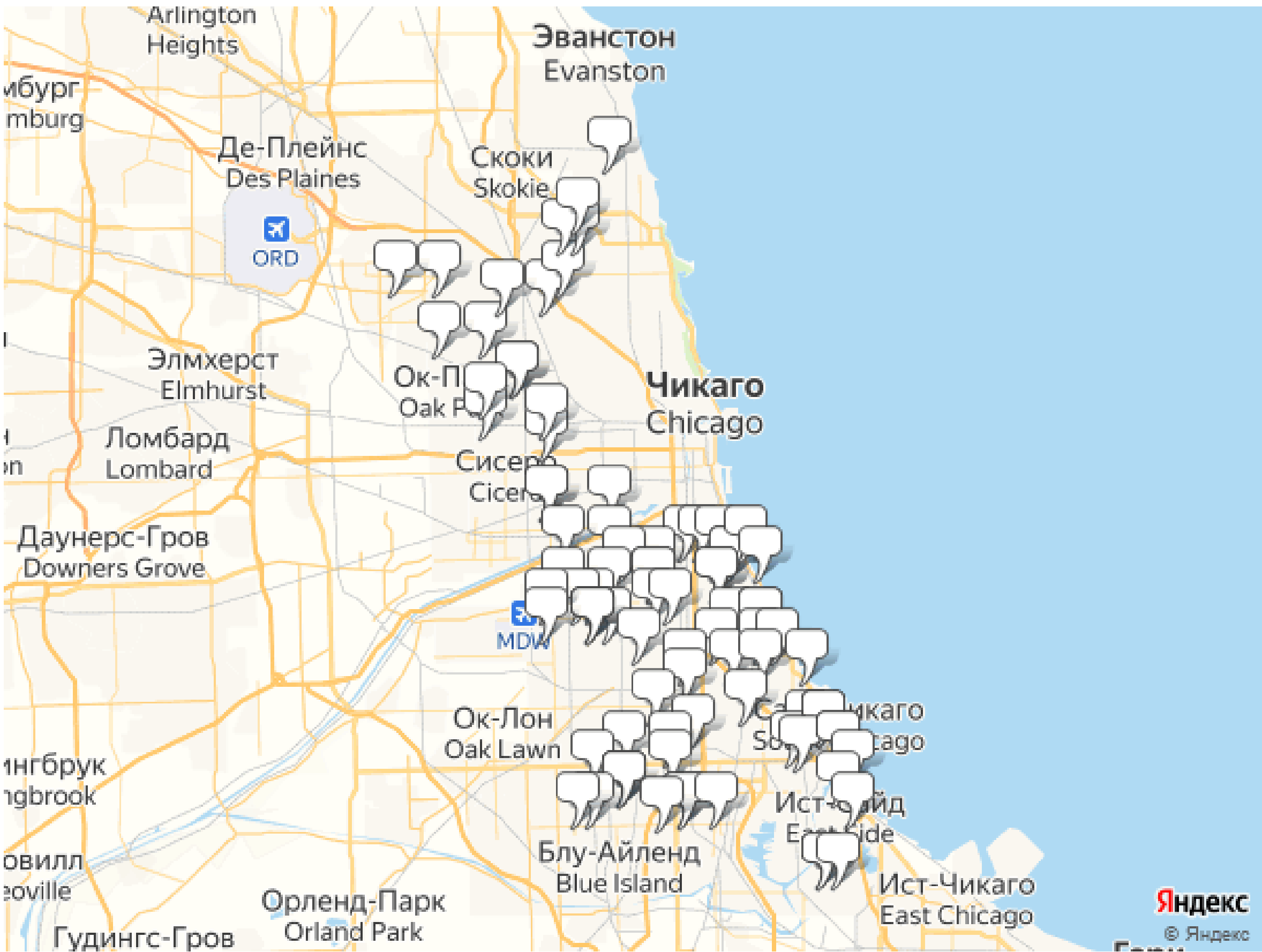
От типа велосипеда **не** зависит
популярность станции
(количество поездок за 2023)



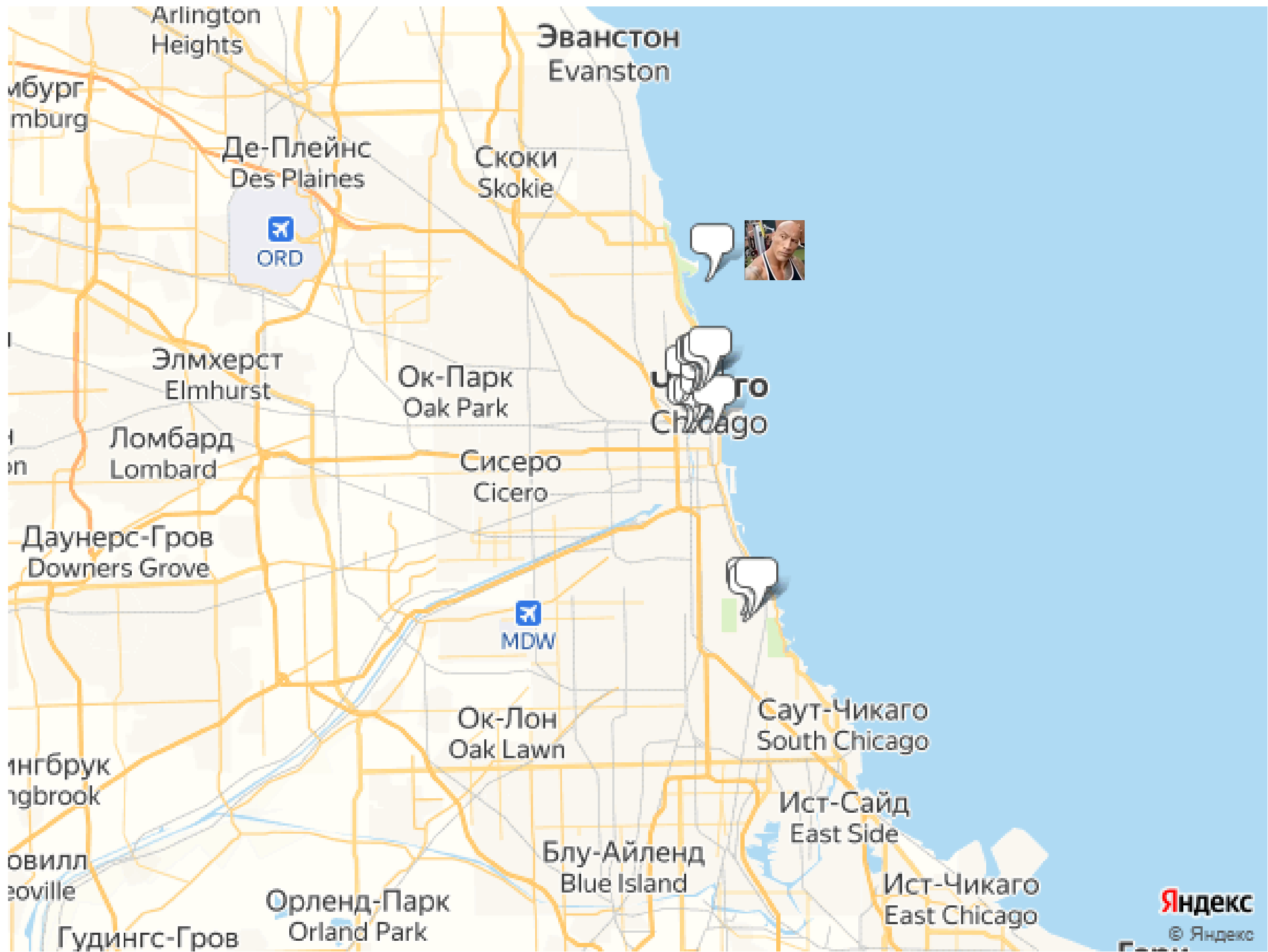
Расстояние между станцией и средним положением
всех станций имеют **обратную связь**

Водичка

Самые непопулярные станции



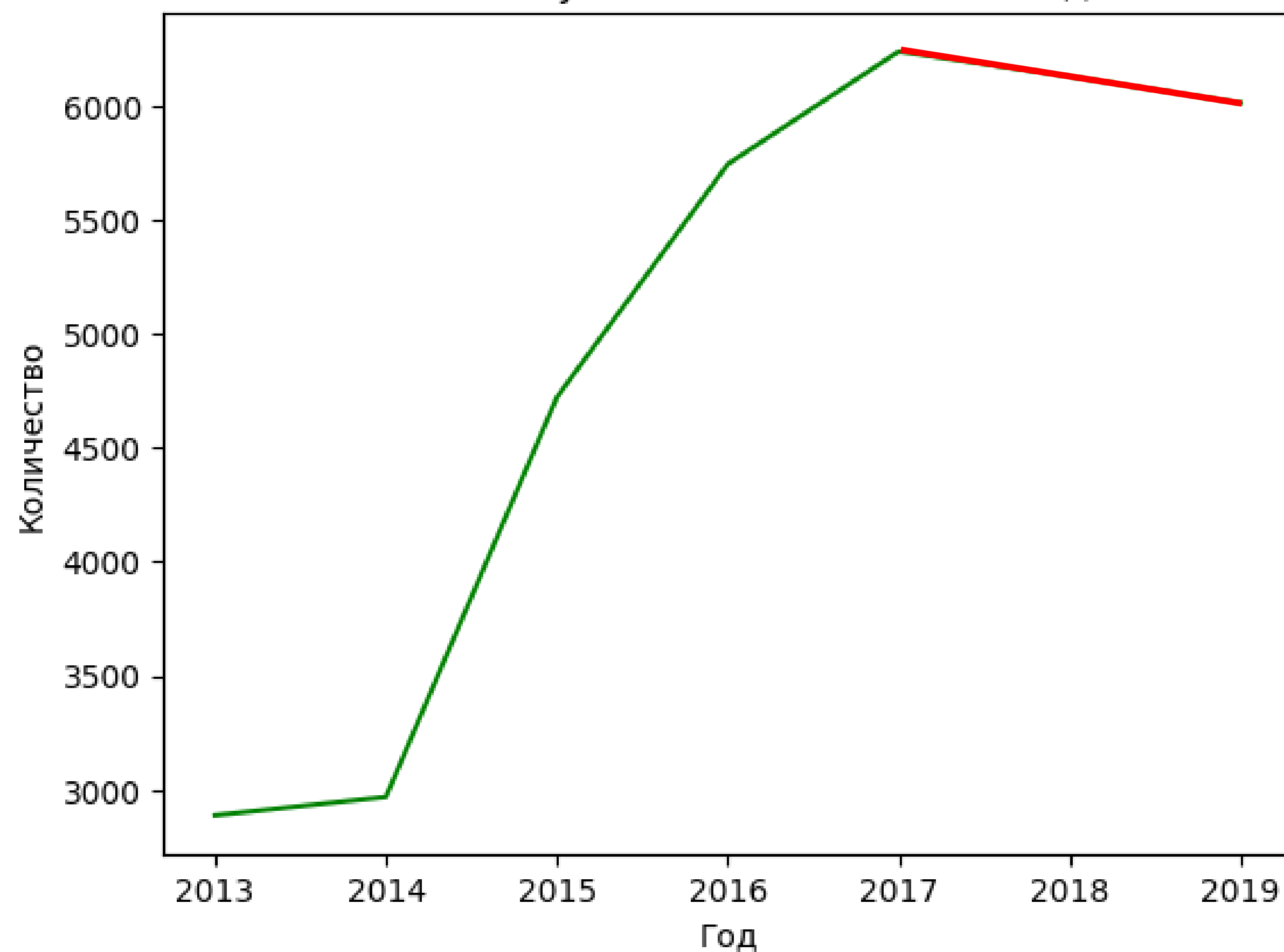
Самые популярные станции



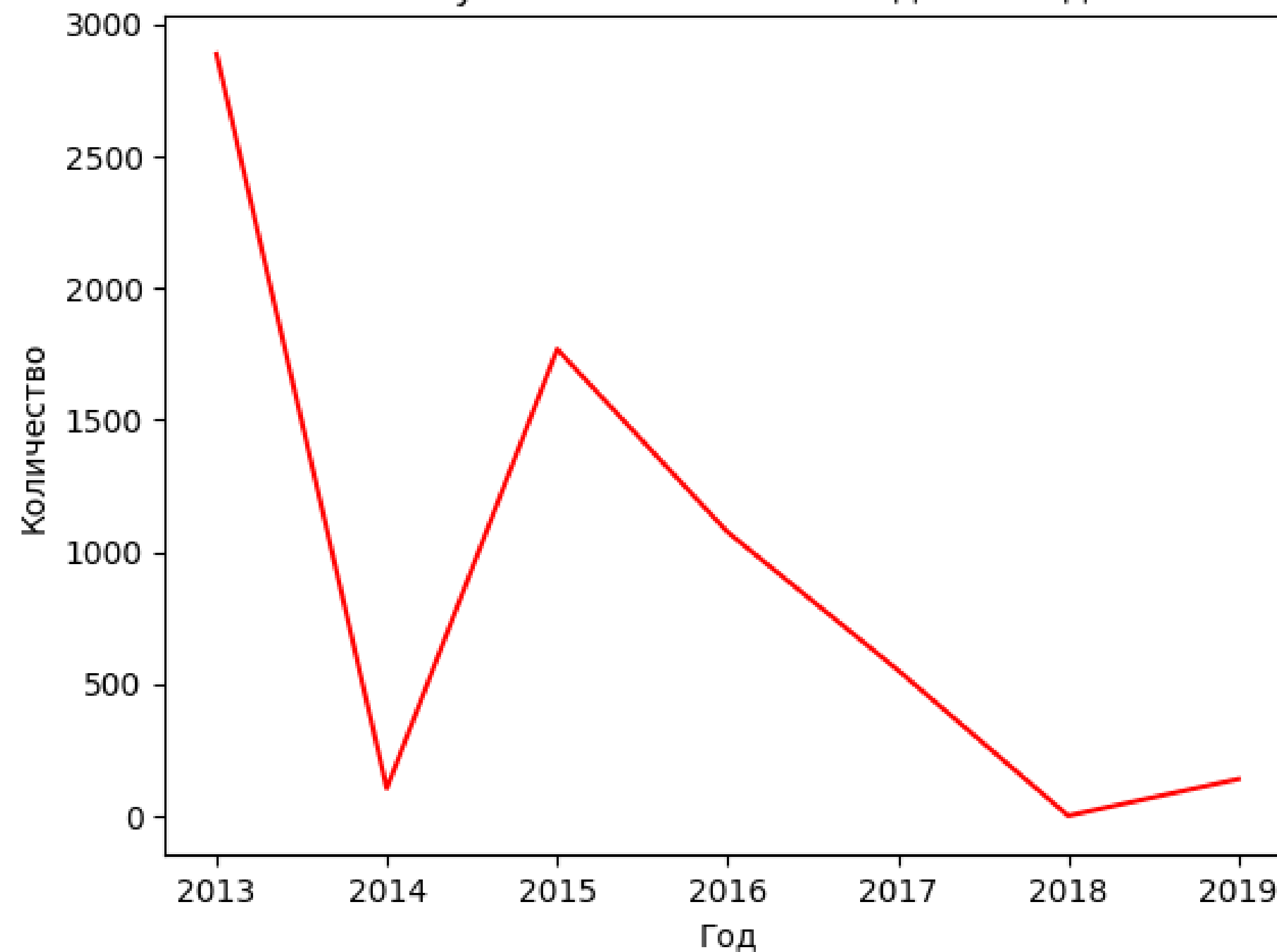
*Наиболее популярные станции расположены в **центре города**, а также в **парковых зонах**
Приток на некоторых станциях **превышает** допустимые значения*

Количество велосипедов

Количество уникальных байков по годам



Новые уникальные велосипеды по годам

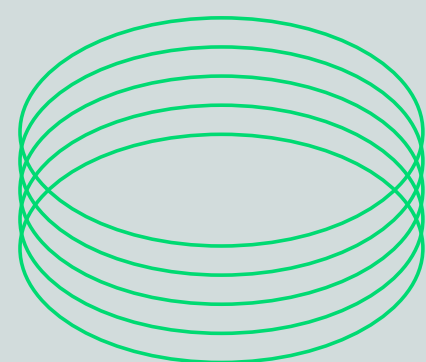


Как измерять пробег?

Время поездки

- + Самый точный критерий
- + Требуется только время
- + Обработывает прогулочные поездки
- Необходимость подбирать среднюю скорость

Среднюю скорость примем равной
12 км/ч



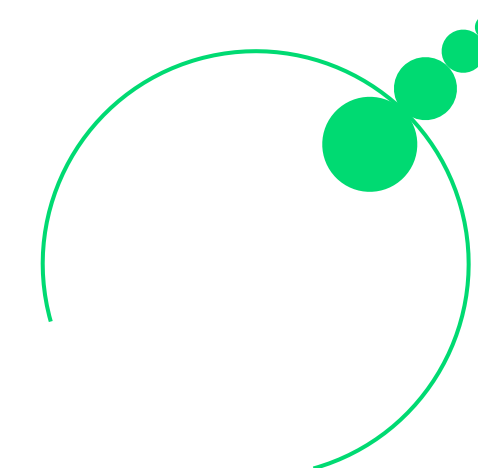
Расстояние между станциями по координатам

- + Точные измерения без подбора констант
- + Легко считать
- Прогулочные поездки (начинаются и заканчиваются в близлежащих точках)

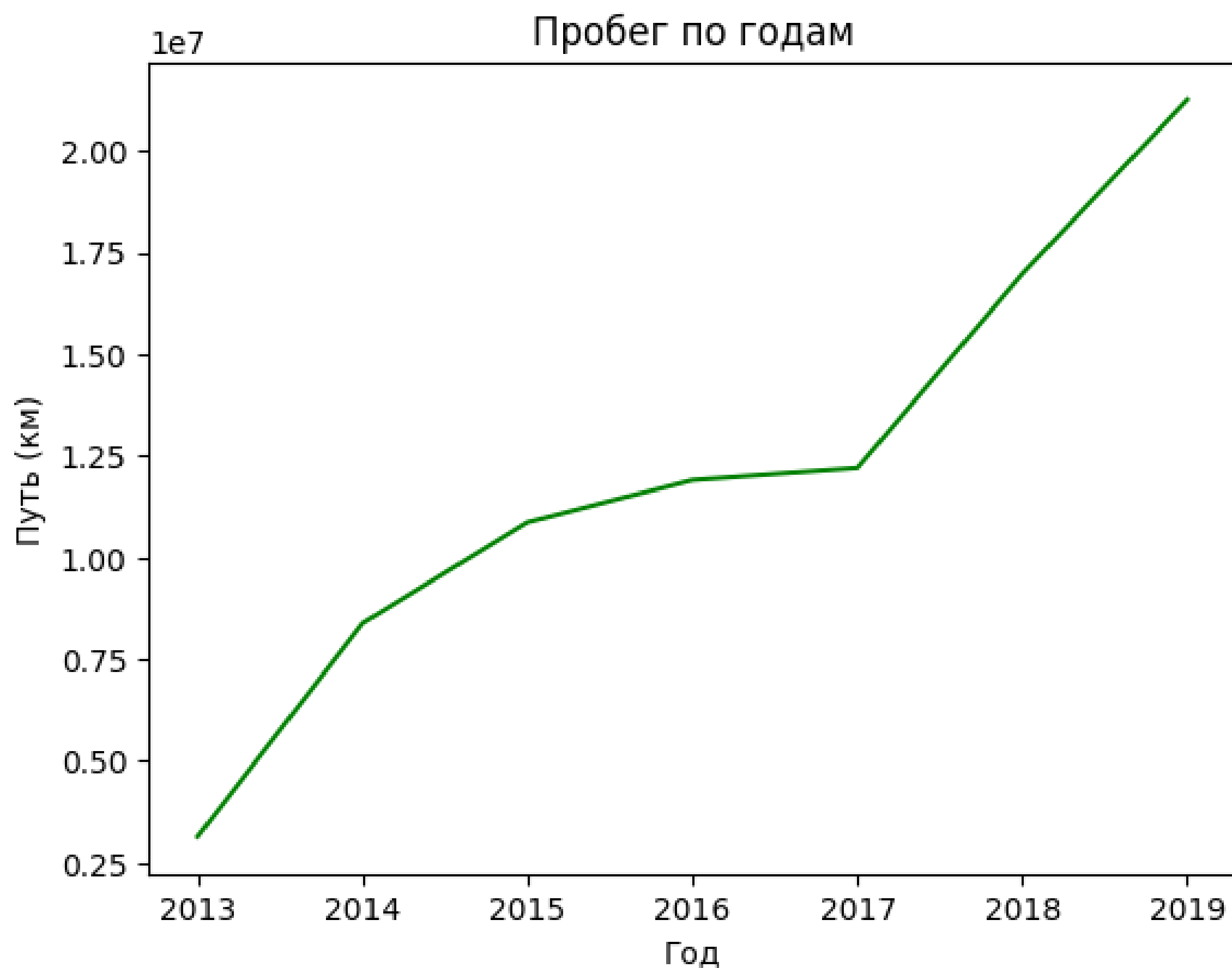


Фиксированная средняя длина поездки

- + Наименьшее количество данных, требуемых для расчета
- Трудно подбирать
- Неточная



Четырёхколёсные друзья



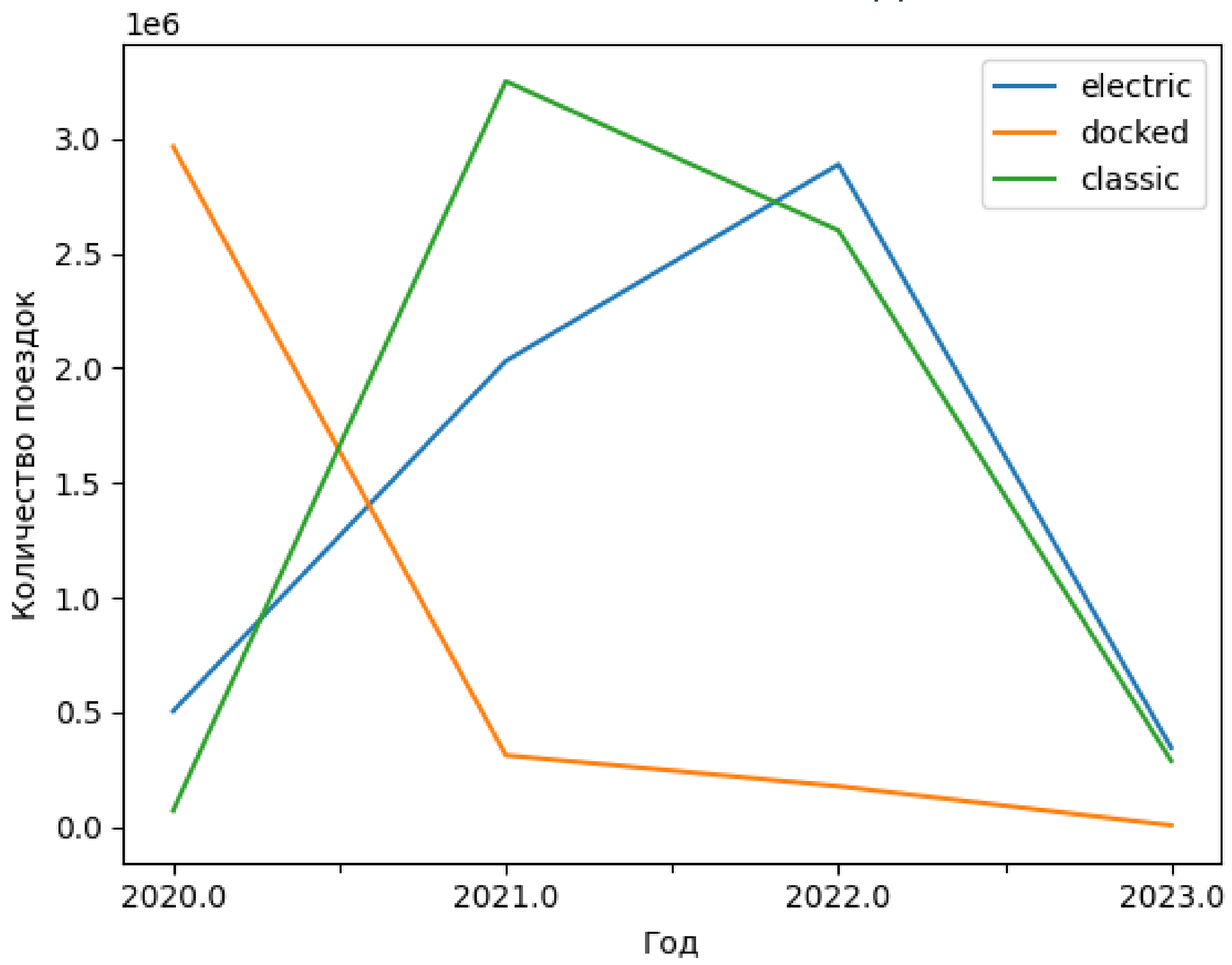
Информация только за 2013 – 2019 гг.



Примерно каждый **10000 км** пробега велосипед нуждается в ремонте

Симпл димпл или поп ит?

Соотношение количества поездок по типам



Информация только за 2020 – 2023 гг.



classic/docked
(Классический)



electric
(Электрический)

Одна поездка

- Базовый тариф
- Без дополнительных ограничений – берешь и едешь
- Все цены за минуту округлены до ближайшей минуты

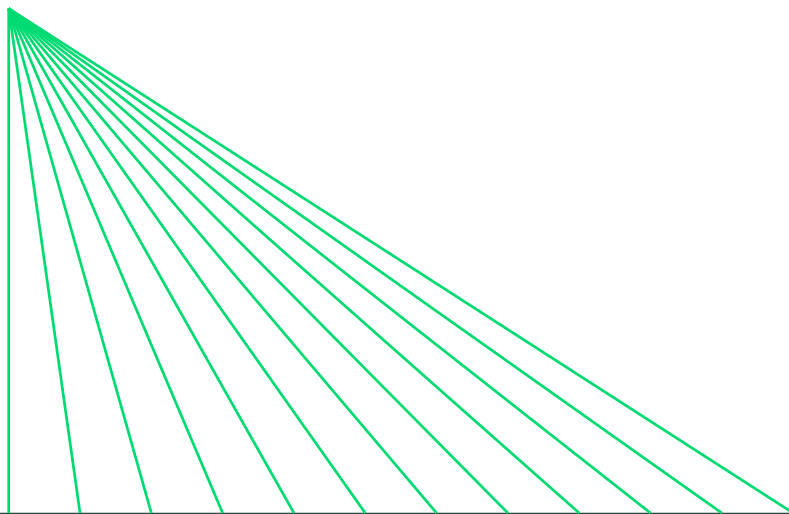
$1\$ + 0.17\$ \times \text{минута}$



Дневной абонемент

- Единоразовый абонемент на один день
- Поездка не более 3 часов
- 0.17\$ за превышение лимита

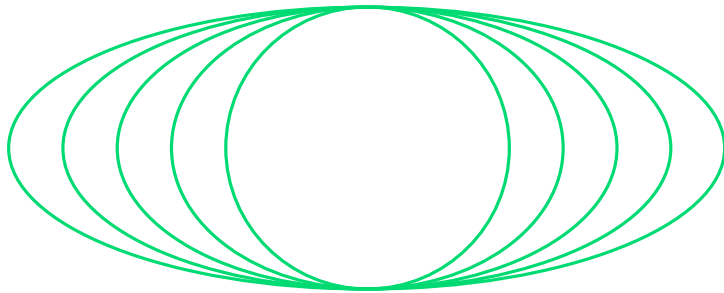
$16.50\$ \times \text{день}$



Годовое членство

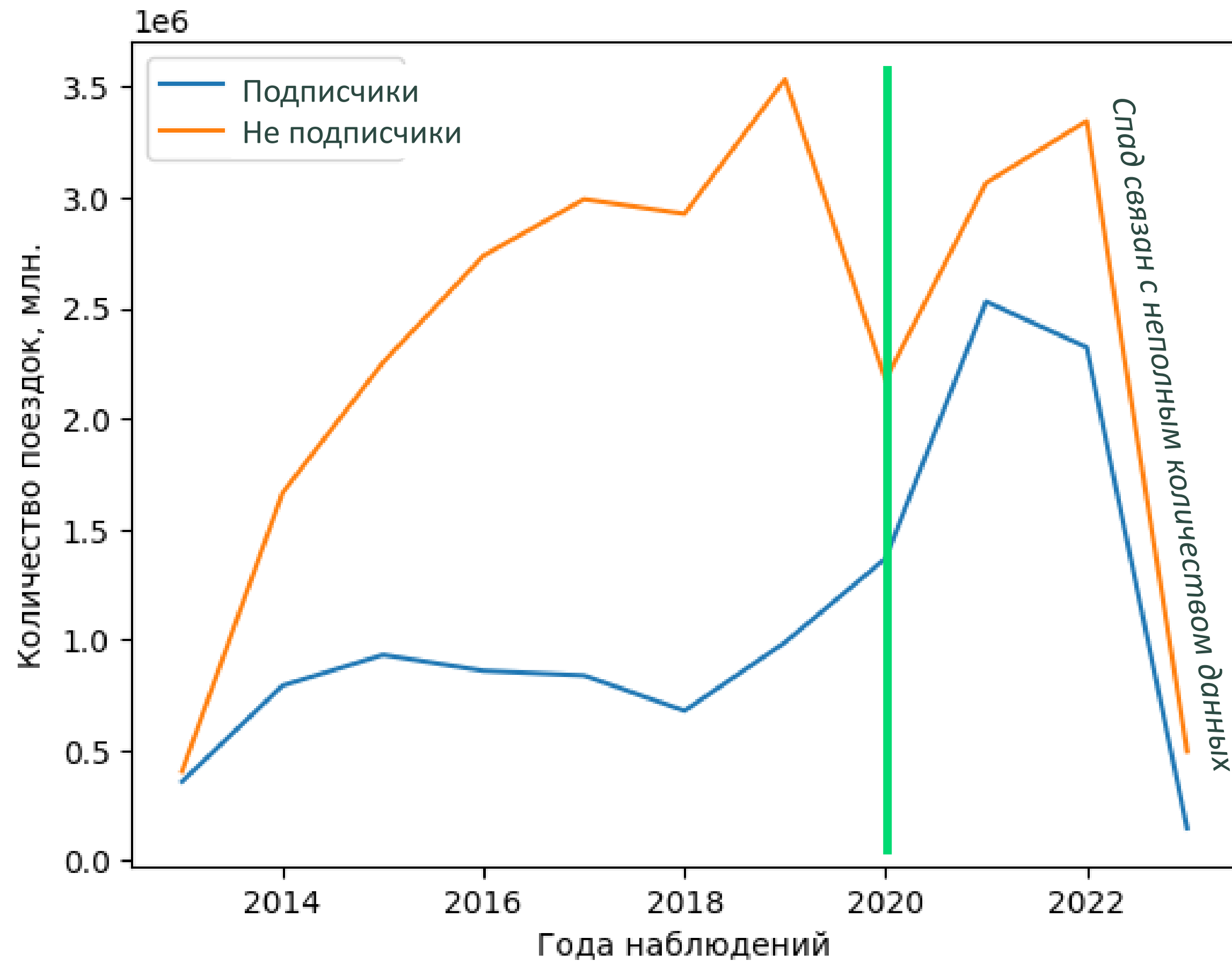
- Неограниченные поездки
- Поездка не более 45 минут
- 130.90\$ оплачиваются авансом за год

$11\$ \times \text{месяц}$



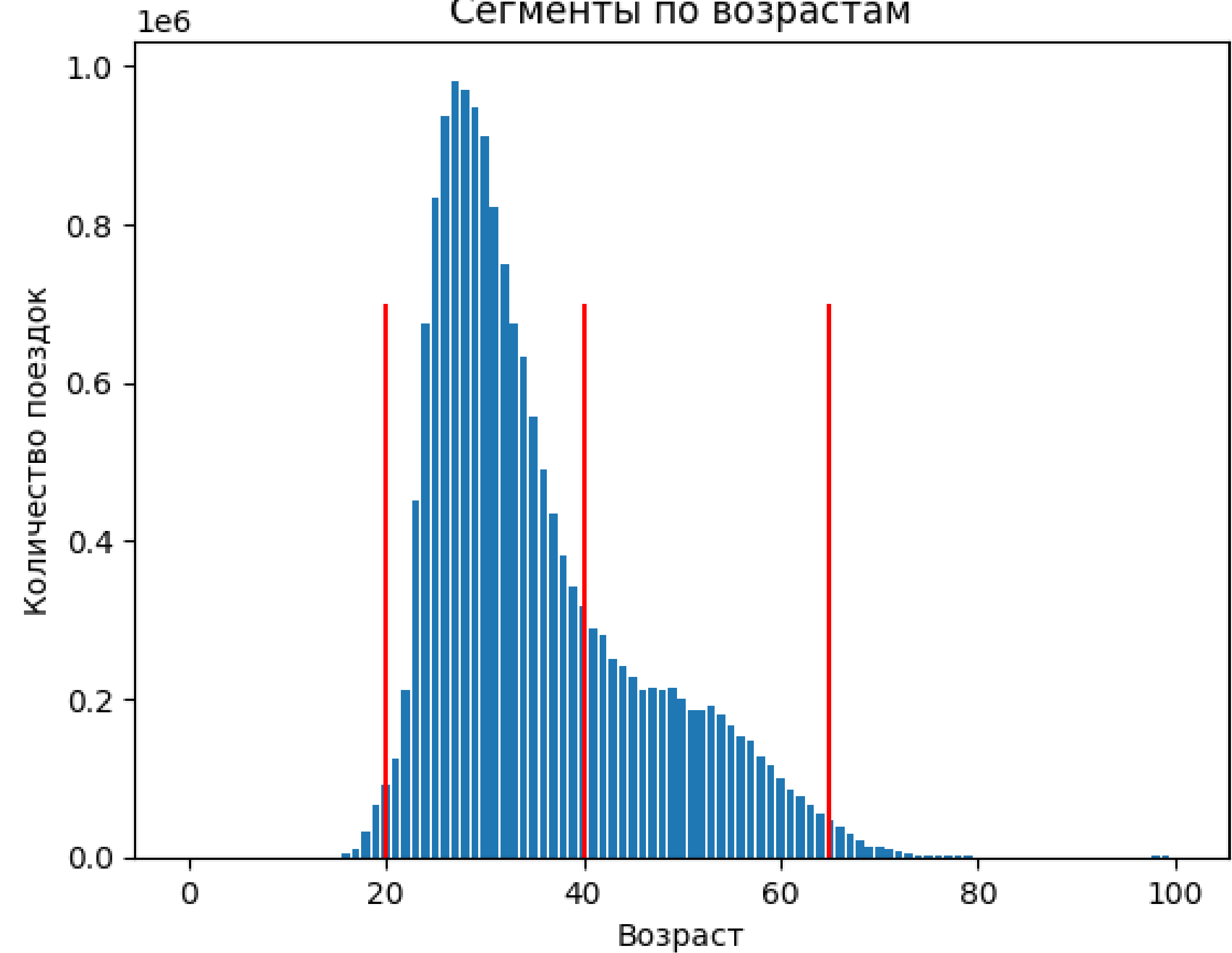
А кто это у нас тут катается?

Соотношение количества поездок по типам



Количество платных пользователей всегда **превышало** количество бесплатных

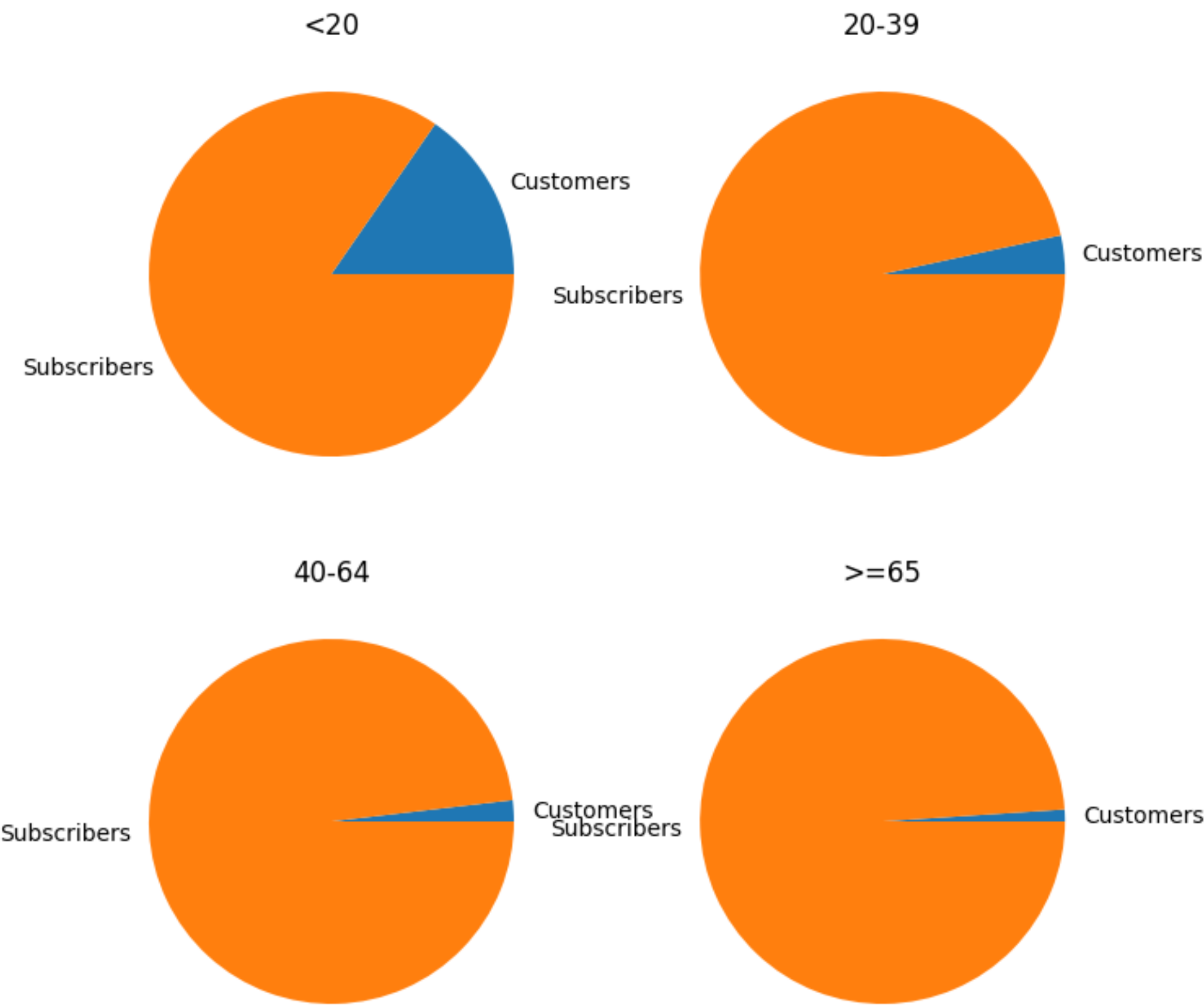
Сегменты по возрастам



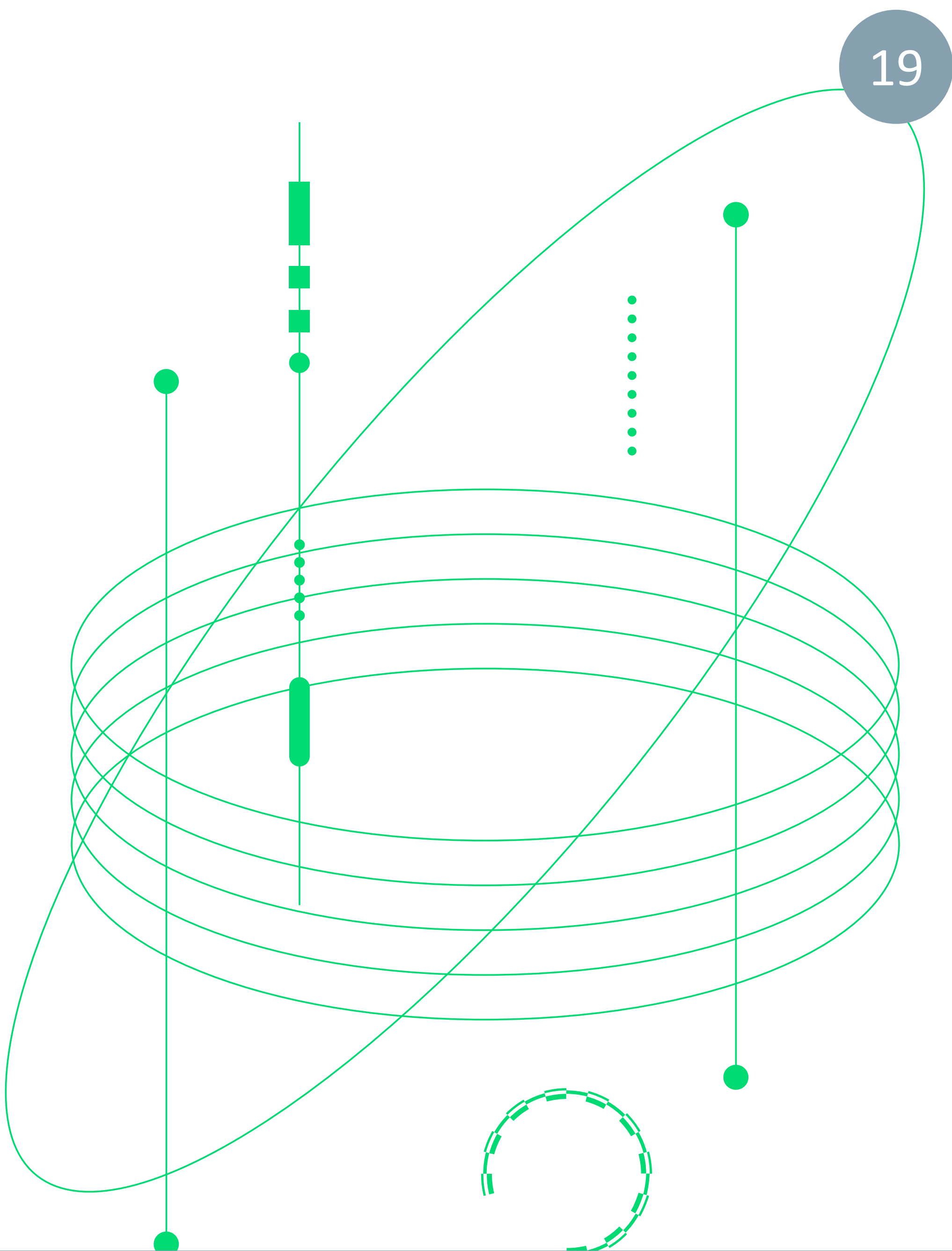
Пользуются очень часто – от **20 до 45 лет**

Аудитория

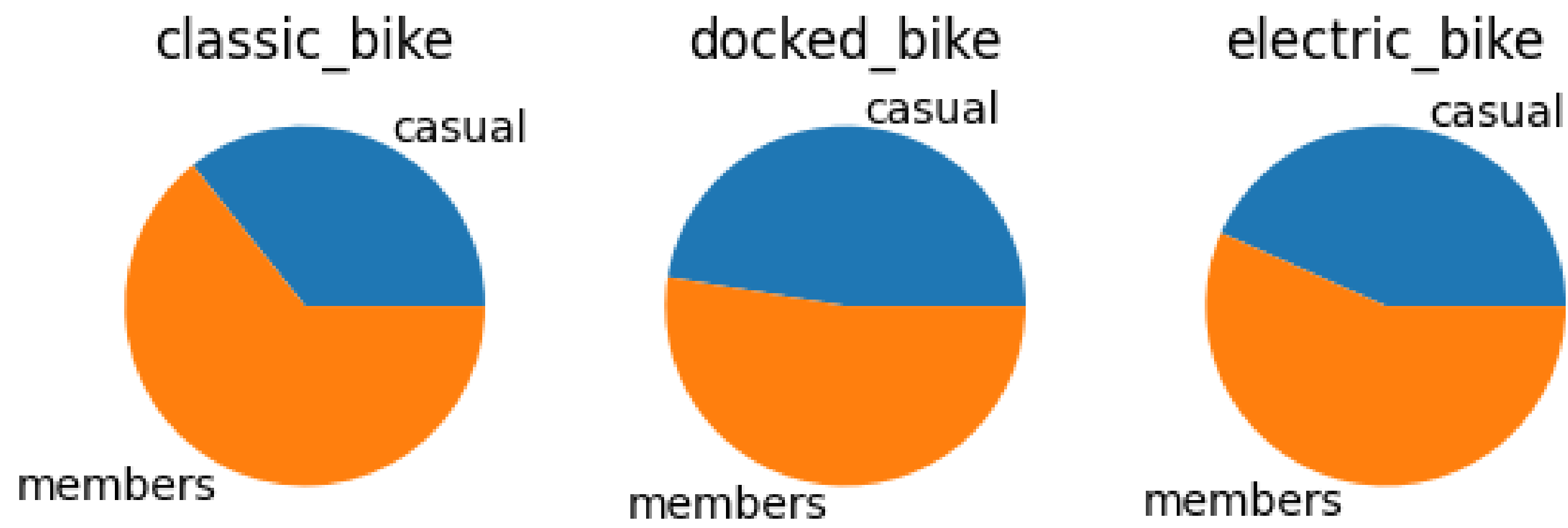
Доля подписчиков по возрастам



Соотношение меняется в пользу **подписчиков** с возрастом

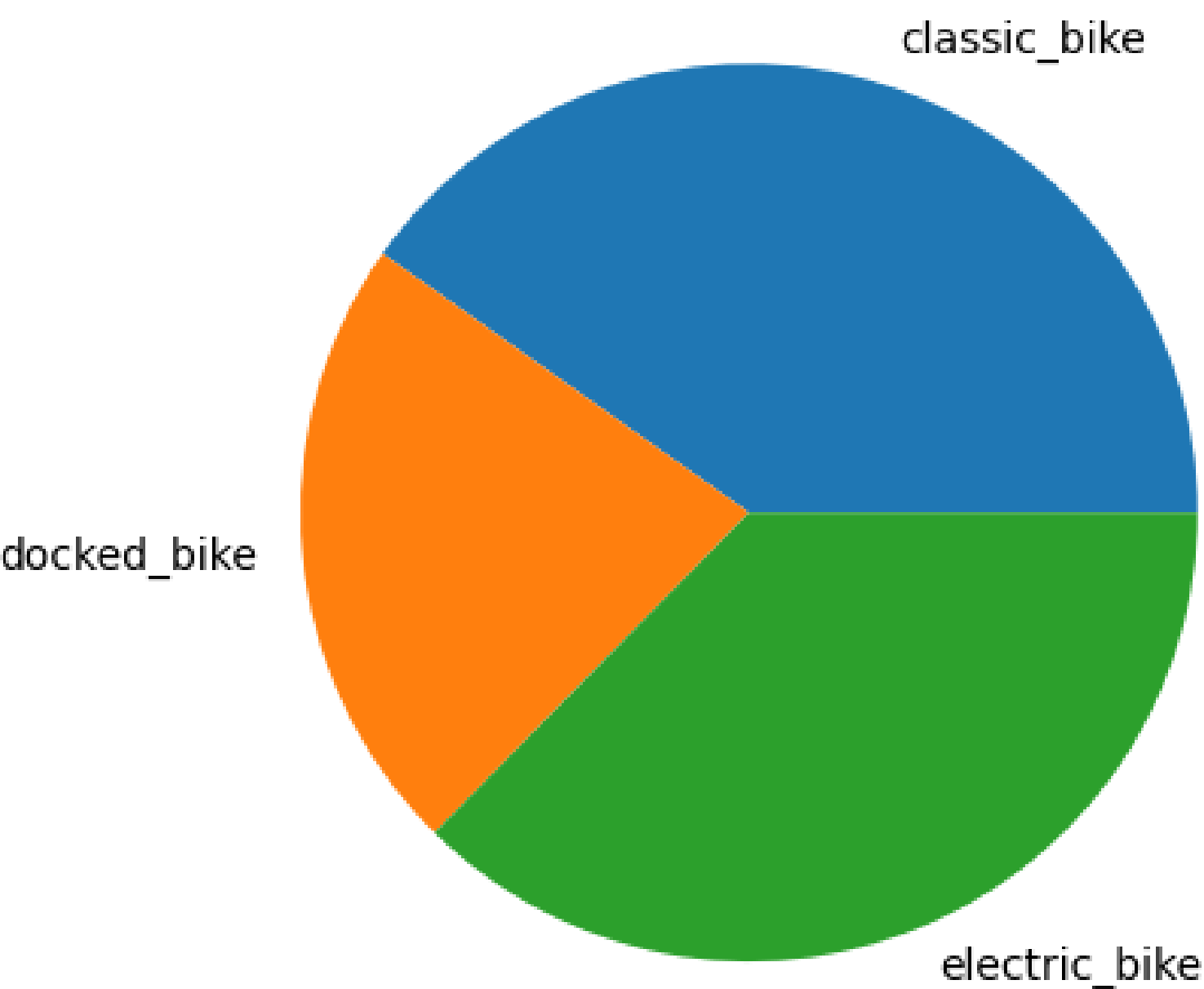


Соотношение типов пользователей к типам велосипеда



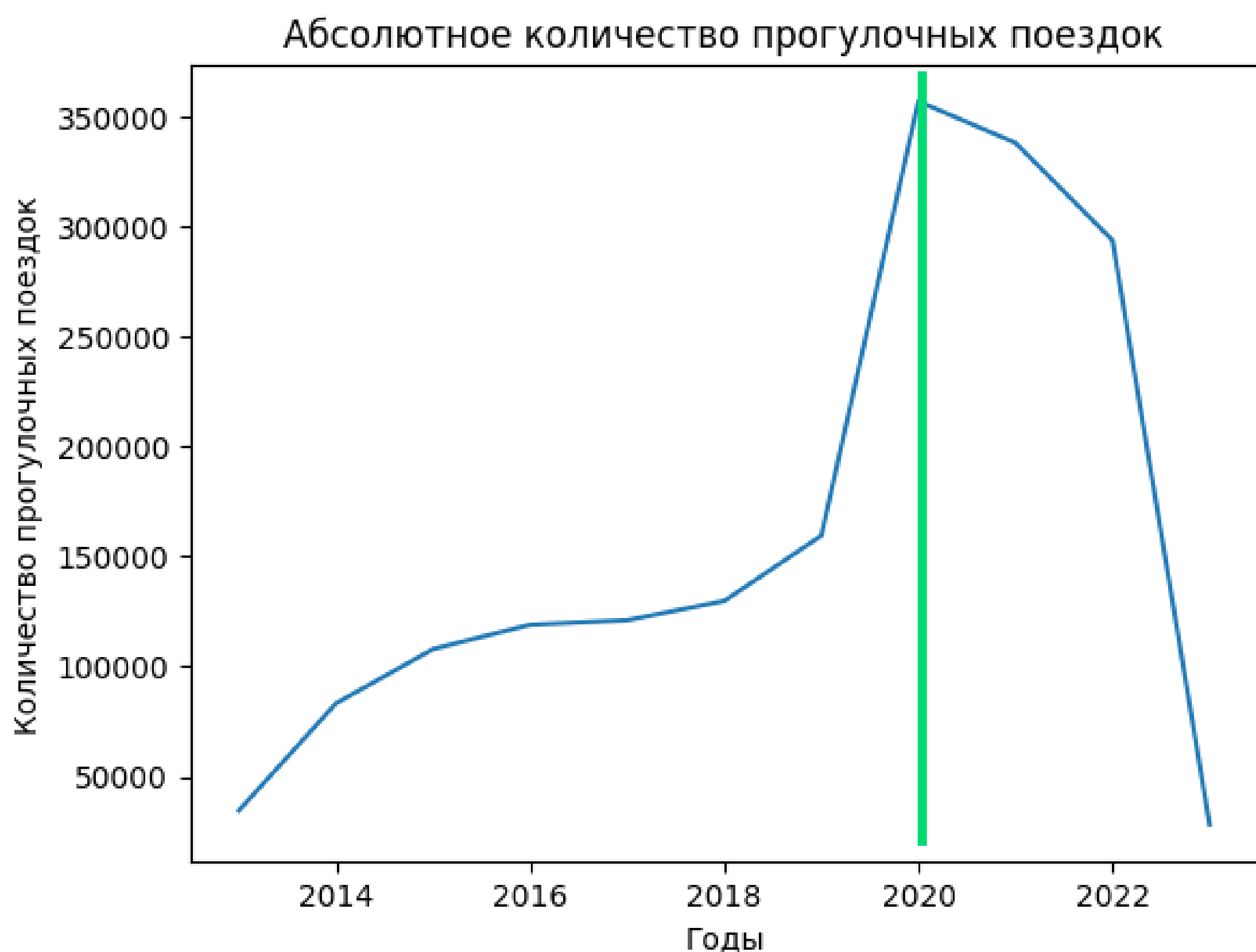
Нельзя сделать вывод о том, какие типы велосипедов популярны у платных пользователей и обычных

Среднее соотношение популярности байков



Классический велосипед остается в приоритете у пользователей

Аудитория

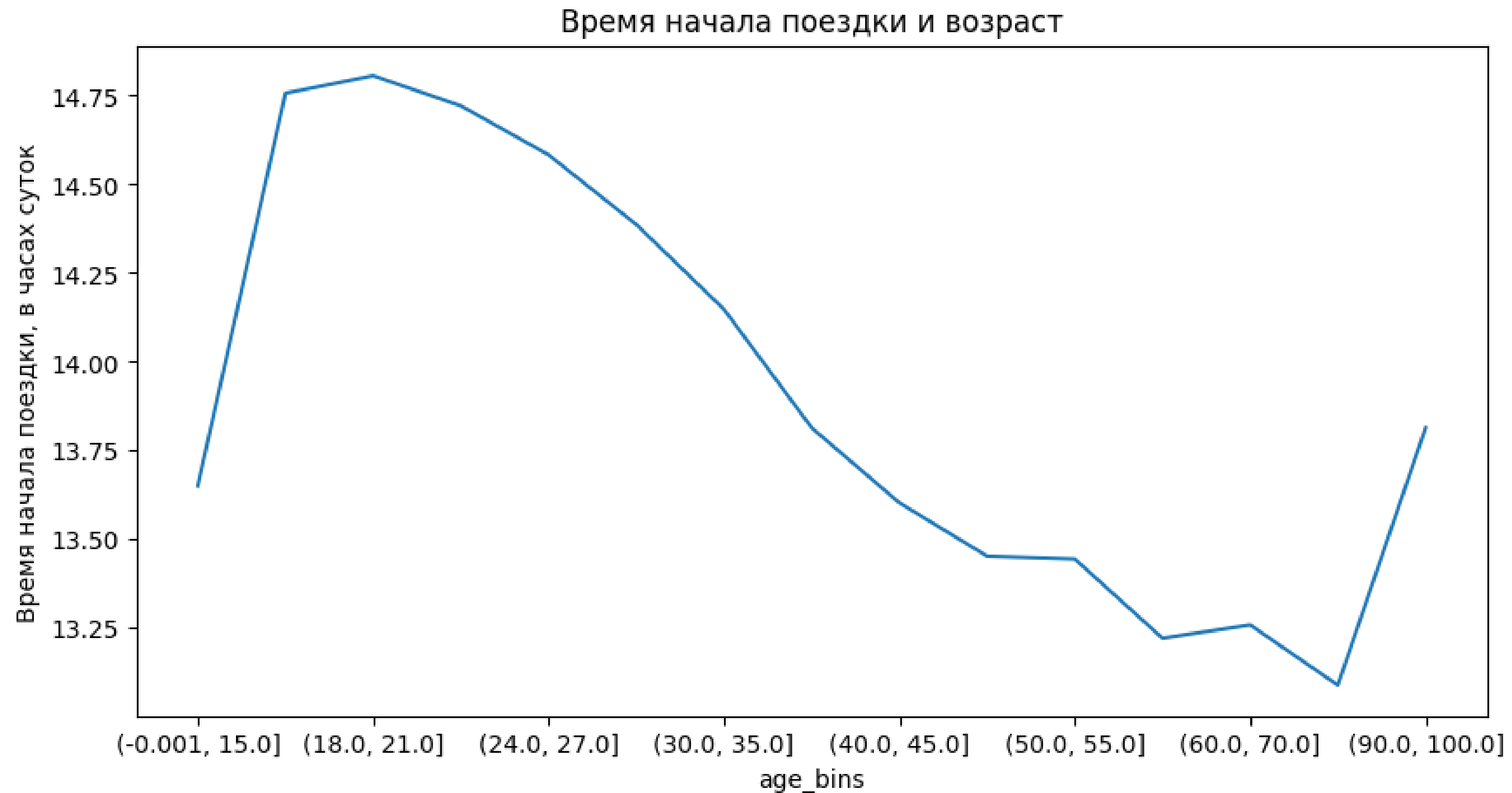


В 2020 в связи с пандемией доля прогулочных поездок сильно **возросла**

Аудитория

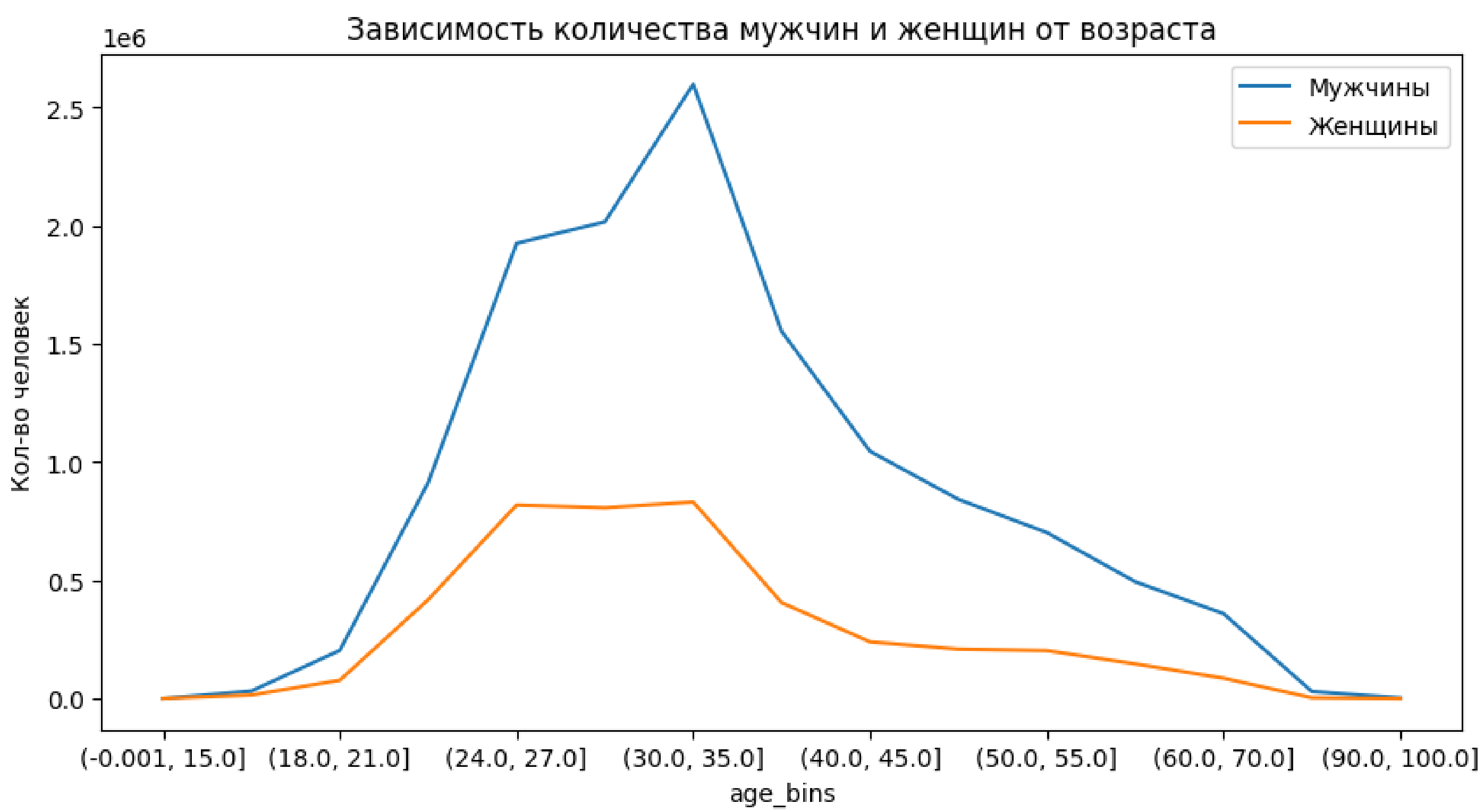


*У молодежи и пенсионеров средняя продолжительность поездки в среднем **больше***



Время начала поездки для каждой возрастной категории **разное** по нескольким причинам

Ох уж эти тренды...



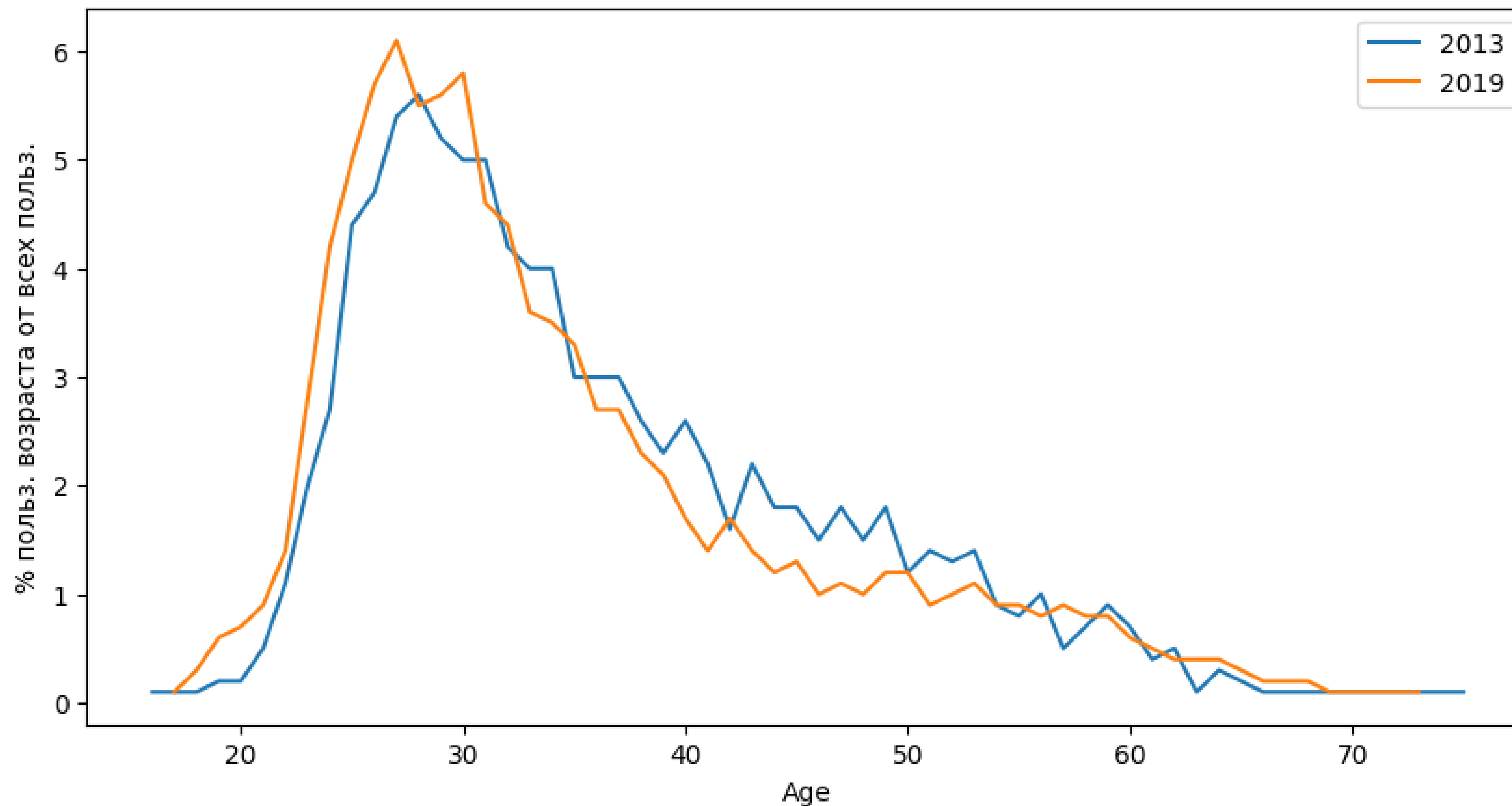
Молодые парни будут чаще «играться» с велопрокатом



Средний возраст клиентов постепенно уменьшался из-за стремления к трендам





Ох уж эти тренды...

Отношения возрастов к аудитории



Популярность среди молодежи **выросла**, а среди людей среднего возраста **упала**

Портрет пользователей

<div>Возраст</div> <div>От 20 до 40 лет</div>	<div></div>	<div>Подписка</div> <div>Платный пользователь</div>	<div></div>	<div>Длительность поездки</div> <div>~ 15 минут</div>	<div></div>
<div>Пол</div> <div>Мужской</div>	<div></div>	<div>Предпочитаемый велосипед</div> <div>Классический</div>	<div></div>	<div>Вероятность прогулочной сессии</div> <div>~ 10%</div>	<div></div>

Фиксируем прибыль

1.7M\$

Средние расходы на
покупку новых
велосипедов в год

12.9k\$

Средние расходы на
ремонт велосипедов в
год

10k\$

Средние расходы в год
на одну станцию

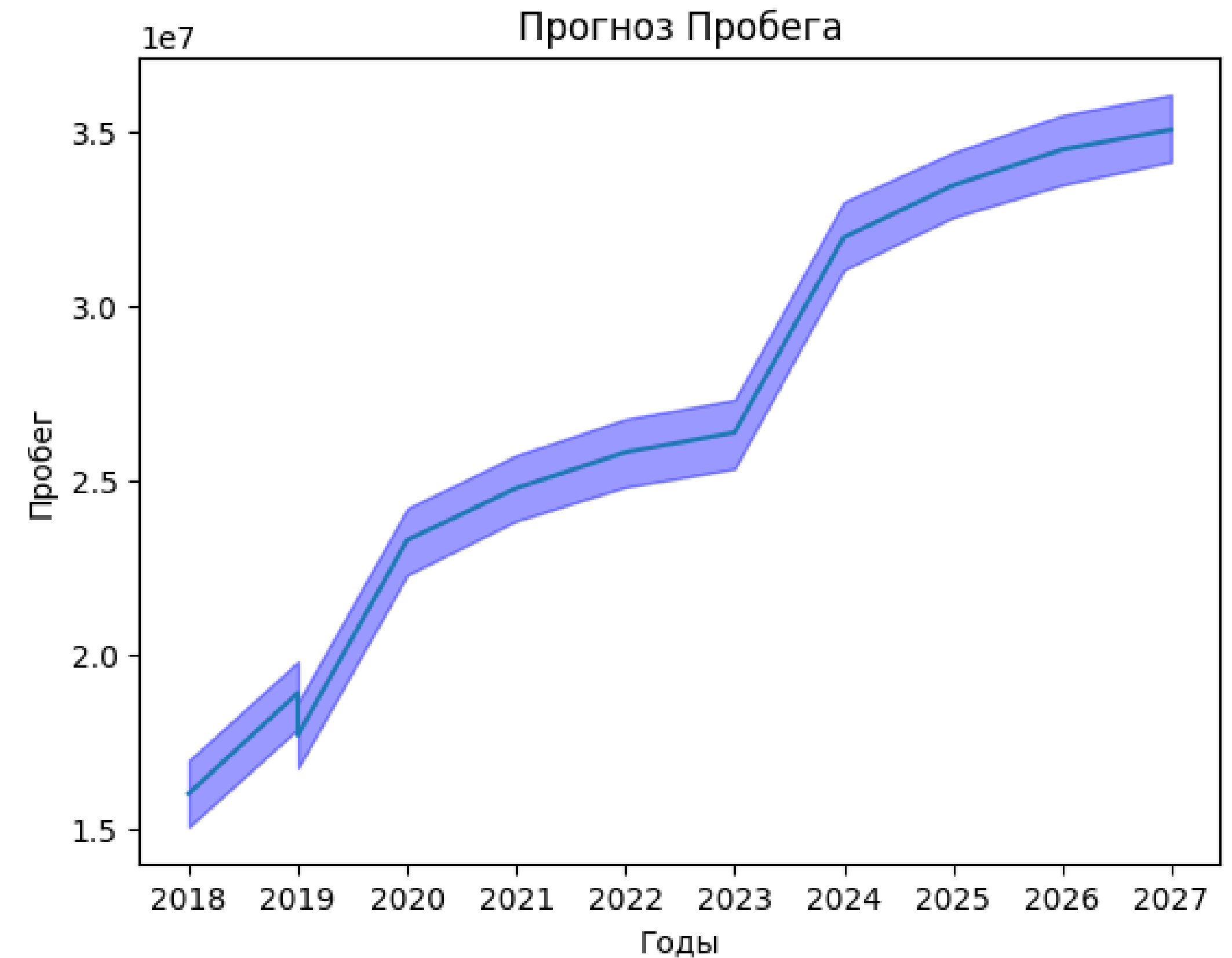
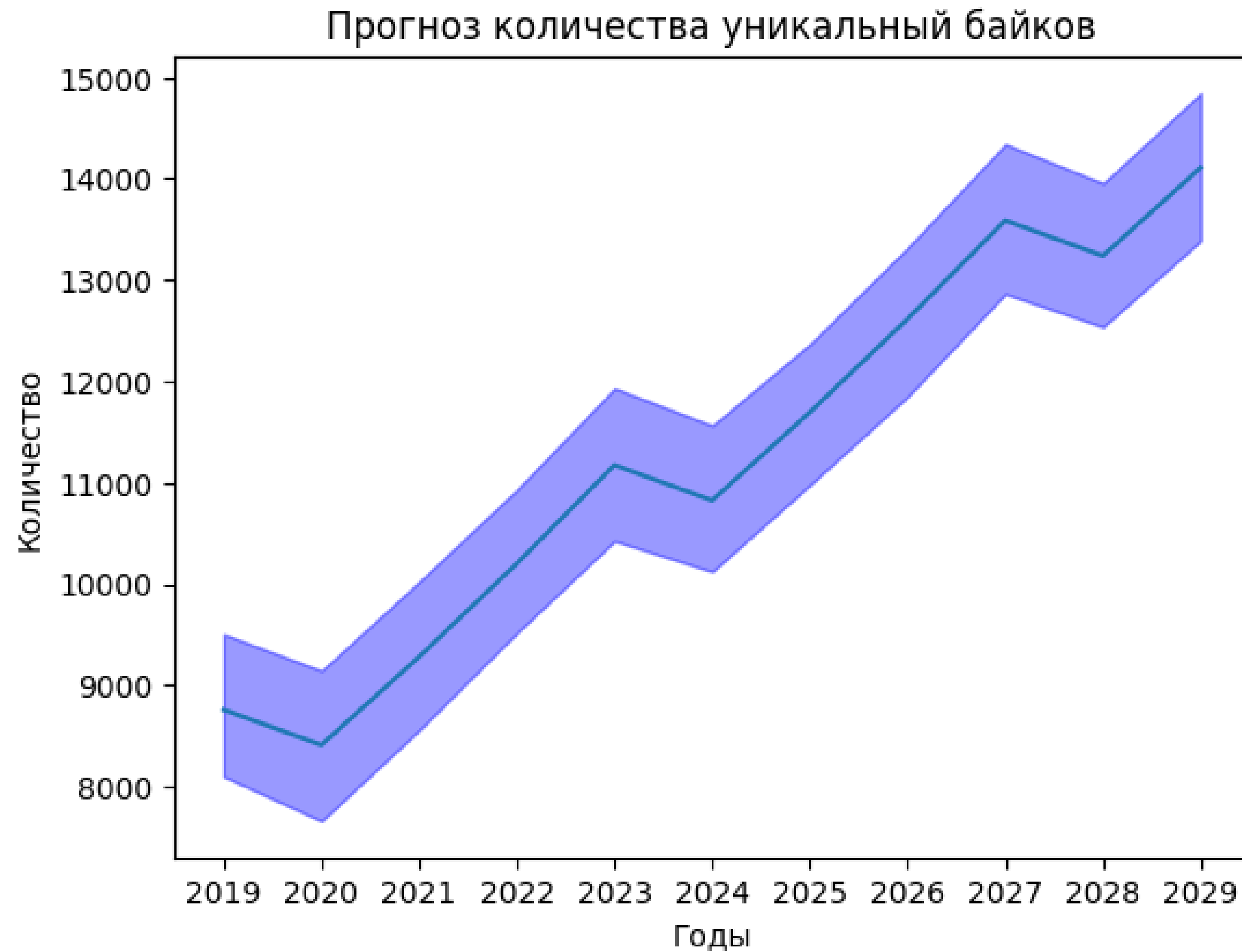
12.7M\$

Общие расходы за весь
2023 год



* Невозможно получить данные о платных пользователях

Prophet (не profit)



Согласно прогнозам, велосипедная сеть будет расширяться. Так ли все очевидно на самом деле?

Есть две новости...

Рейтинг Divvy на TripAdvisor



01 Неудовлетворенность клиентов сервисом

02 Проблемы с доступностью велосипедов

03 Участвовавшие случаи кражи

04 Плохая клиентская поддержка



свидетель 00d
Вулверхэмптон, UK • 1 публикация



Вводят в заблуждение и взымают то, что хотят
сент. 2022 г. • Для двоих

Многие пользователи жалуются на ошибочные списания, клиентскую поддержку



Эндрю Ф
1 публикация

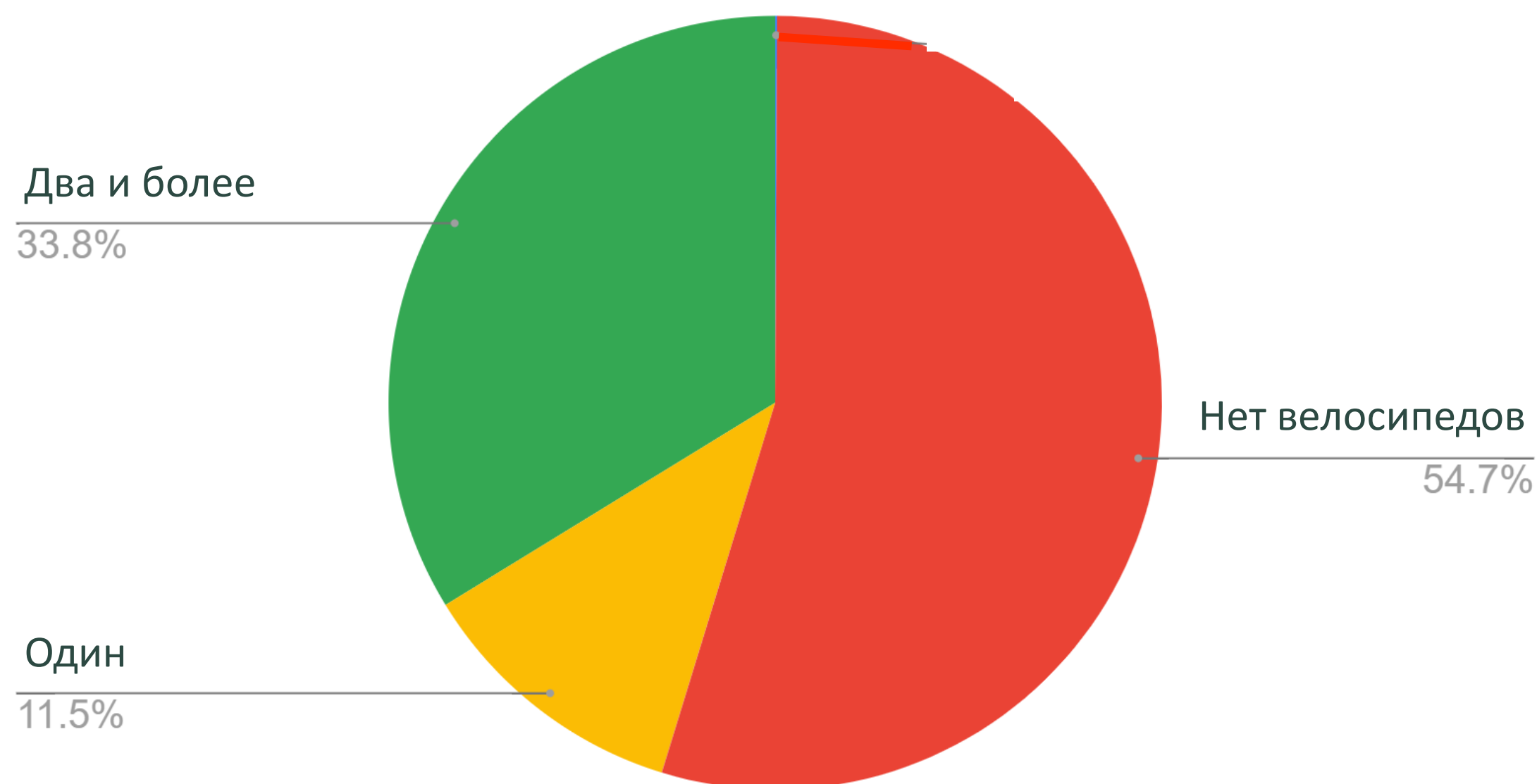


Ищу адвоката для судебного иска против Divvy
сент. 2022 г. • Для двоих

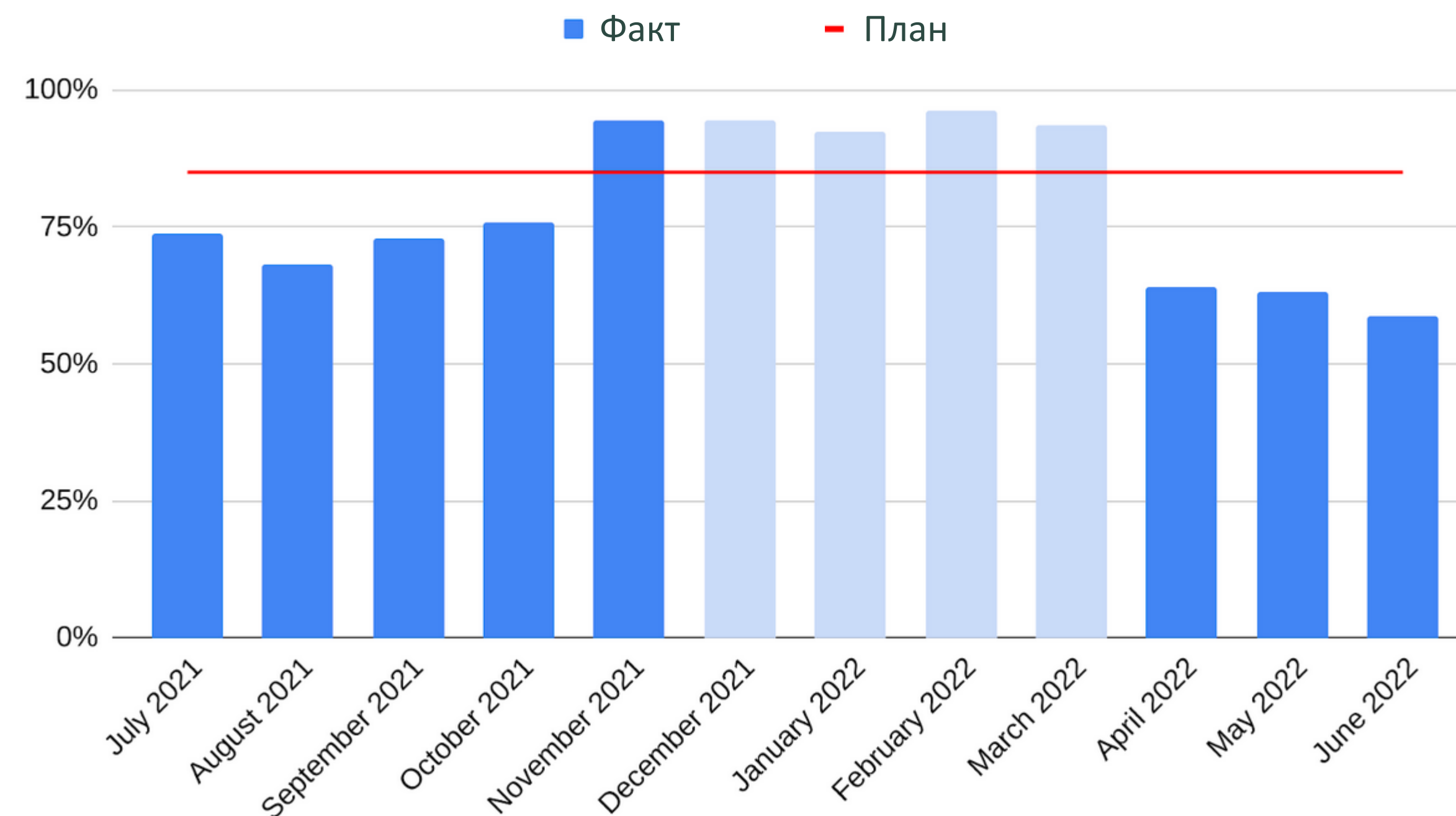
Кто-то идет на крайние меры

Велосипед временно недоступен

Количество классических велосипедов на одной из центральных станций
(лето 2022 года)



Процент действующего велопарка



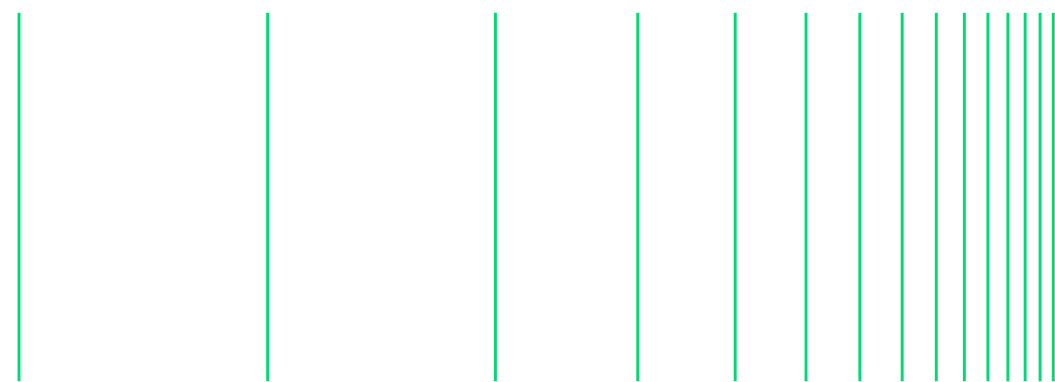
В 2022 Lyft и Divvy столкнулись с **серьезными проблемами** в управлении сети.
Прогноз на 2023 год невозможно из-за вероятных **изменений** в структуре компании

Пришло время подвести итоги

Исследование

За счет выбранных нами методов были:

- Спрогнозированы ключевые цифры бизнеса на 2023 год
- Сформирован портрет среднего пользователя, изучены базовые закономерности в поведении пользователей



Проблемы и решения

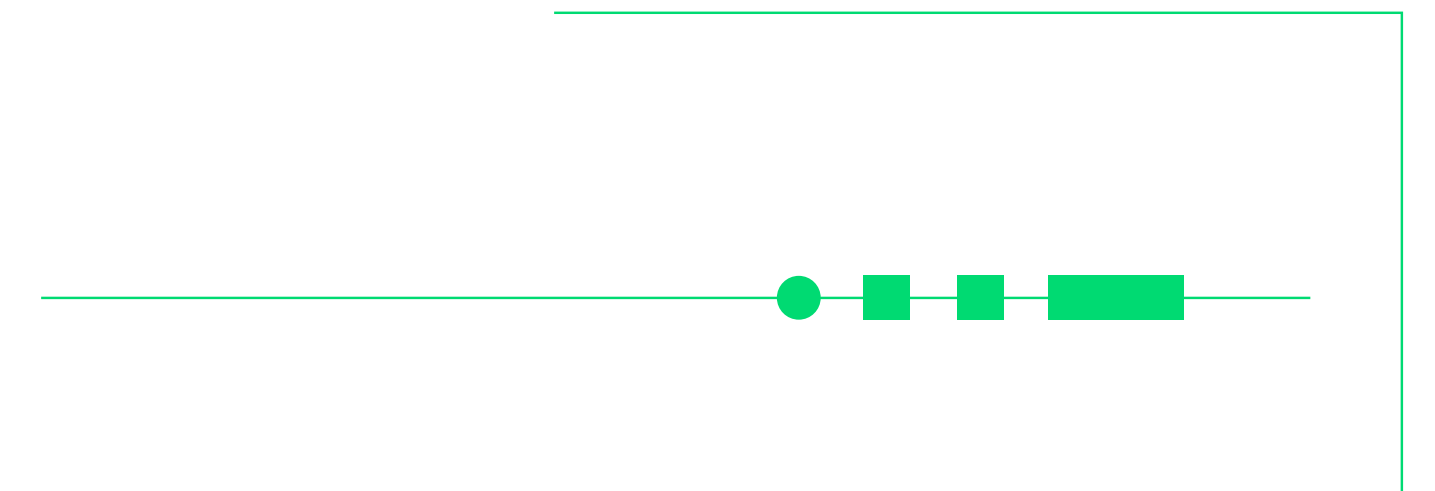
Исходный датасет, а также подтверждение из внешних данных, дают понять:

- На некоторых станция велопроката наблюдается аномальный приток
- Случаи кражи велосипедов участились.
- Клиенты не удовлетворены сервисом.
- Сервис Divvy Bikes испытывает не самые стабильные времена. Необходимо подождать новых решений компании.

Направления развития

Объективно, логов лишь о поездках было недостаточно для широкого анализа.

- Изучение логов из приложения сервиса проката
- Повышение точности прогнозов (машинное обучение)
- Опрос пользователей



Движение – жизнь!

Пришло время вопросов =)

