

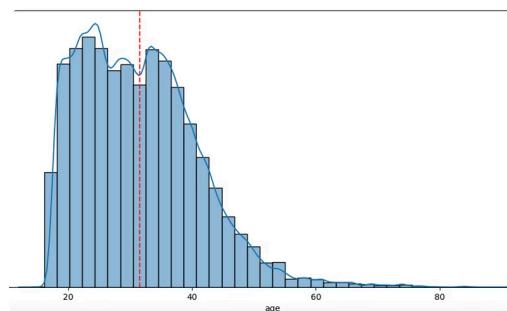


🎯 Анализ Пользователей Кикшеринга



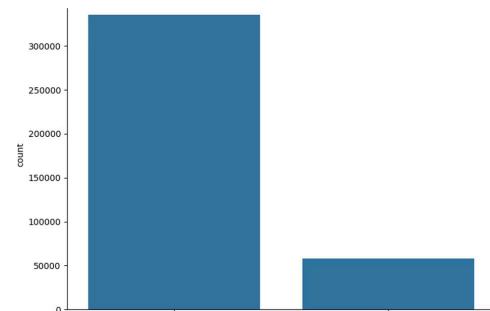
Кто наши пользователи? Демография и Поведение

Основная аудитория — молодые мужчины 20–40 лет; данные по образованию и семейному статусу показывают высокую долю пропусков и непригодны для принятия решений.



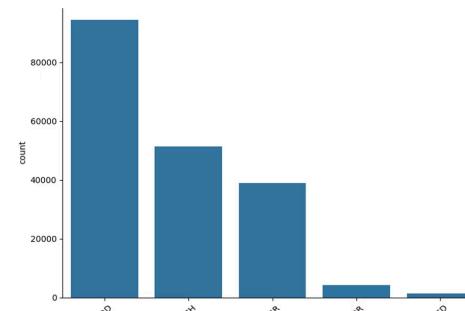
Распределение по возрасту

Пик приходится на возраст 20–40 лет. Значения возраста свыше 80 лет рассматриваются как аномалии или ошибки ввода, требующие дальнейшей верификации.



Распределение по полу

Подавляющее большинство пользователей (около 98%) — мужчины. Этот факт может указывать на потенциальную недооценку женской аудитории, возможно, из-за использования общих аккаунтов.



Уровень образования

Более 52% данных отсутствуют, что делает этот показатель неинформативным для принятия стратегических решений. Предварительный анализ указывает на доминирование студентов.



Ключевой Инсайт

Целевая аудитория — **молодые мужчины**. Сегментация по полу и возрасту может быть эффективной. Для использования образования требуется улучшение качества сбора данных.

💰 Как формируется стоимость поездки?

Стоимость и кэшбек показывают сильную корреляцию, однако дистанция и время поездки содержат критические выбросы, которые необходимо учитывать.

Дистанция поездки (км)

Большинство поездок составляют менее 5 км. Выявлена аномалия: значение 56 000 км, что, вероятно, является ошибкой ввода (возможно, метры вместо километров).

Номинальная стоимость (₽)

95% поездок имеют стоимость до 200 ₽. Выброс около 5 600 ₽ требует дополнительного расследования, возможно, это результат очень долгой или незавершенной поездки.

Корреляция Стоимости и Кэшбека

Обнаружена сильная корреляция ($\phi_k \approx 0.8$) между номинальной стоимостью и начислением кэшбека. Это подтверждает ожидаемую линейную зависимость, но требует проверки лимитов кэшбека.

Топ-3 Аномалий

- Дистанция > 100 км: 12 записей (требует фильтрации/проверки)
- Стоимость > 1000 ₽: 8 записей (возможны долгие поездки/ошибки)
- Возраст > 80 лет: 37 записей (вероятные ошибки ввода)

Инсайт: Ценообразование выглядит прозрачным и линейно связанным с кэшбеком. Однако, **аномалии в данных могут исказить модели**, поэтому необходима тщательная предобработка выбросов перед проведением А/В-тестов или построением прогнозных моделей.

Рекомендации и Следующие Шаги

Краткосрочные Задачи

- Удалить или ограничить выбросы в данных по дистанции, стоимости и возрасту.
- Пересмотреть методы сбора данных для показателей образования и семейного положения, так как текущие данные неинформативны.
- Ввести валидацию возраста при регистрации, установив максимальное значение (например, 75 лет) для предотвращения ошибок.

Долгосрочные Перспективы

- Разработать сегментацию пользователей на основе частоты поездок и времени суток для внедрения динамического кэшбека.
- Исследовать взаимосвязь между часовым поясом бронирования и тарифами для подтверждения гипотезы о региональном ценообразовании.

3 Продуктовые Гипотезы для А/В-Теста

1

Динамический Кэшбек

Идея: Повысить процент кэшбека для поездок в непиковые часы (9:00–14:00), чтобы сгладить нагрузку на автопарк и увеличить оборачиваемость транспорта.

- **Метрика 1:** Доля поездок в непиковые часы (+15%)
- **Метрика 2:** Среднее число поездок на самокат/день (+10%)

2

Геймификация для Новичков

Идея: Ввести бонус за первую поездку длиной >2 км («Достигни цели — получи +50 ₽»), чтобы увеличить конверсию в повторные поездки и среднюю дистанцию у новых пользователей.

- **Метрика 1:** Retention D7 (+8%)
- **Метрика 2:** Средняя дистанция (новые пользователи) (+25%)

3

Упрощение Активации

Идея: Отменить или снизить стоимость активации для пользователей, совершивших ≥3 поездок за последнюю неделю, чтобы увеличить частоту поездок у активных пользователей и LTV сегмента.

- **Метрика 1:** Еженедельные поездки на пользователя (+20%)
- **Метрика 2:** Churn Rate (30 дней) ↓ (-5%)

Примечание: Все гипотезы реализуемы с использованием текущей платежной логики и данных о времени/частоте поездок.

🏆 Почему Гипотеза №1 — самая перспективная?



Опирается на данные

Мы наблюдаем неравномерную загрузку самокатов с пиками в утренние и вечерние часы. Стоимость активации и поминутная тарификация являются ключевыми рычагами, которые уже коррелируют с временными зонами бронирования, что свидетельствует о гибкости тарифов.



Высокий ROI при Низких Затратах

Реализация этой гипотезы не требует значительных изменений в UI/UX; достаточно лишь настроить логику расчета кэшбека. Кэшбек, выступая как маркетинговая инвестиция, напрямую способствует повышению лояльности пользователей и росту LTV (Lifetime Value).



Стратегическое Значение для Бизнеса

Динамический кэшбек повышает оборачиваемость парка, сокращая время простоя самокатов и увеличивая доходность каждой единицы техники. Кроме того, это снижает необходимость в дорогостоящей ручной переброске самокатов.

Итог: Гипотеза №1 комплексно решает **операционные, маркетинговые и финансовые задачи** бизнеса, представляя собой наиболее эффективное направление для А/В-тестирования.

Доля Т-Банка в Индустрии Микромобильности

Как оценить? Датасет kicksharing.csv указывает на сервис аренды электросамокатов. В России в 2024–2025 годах ключевыми игроками являются Whoosh, Yandex.Go Самокаты, а также Ситидрайв/Urent, который теперь ассоциируется с Т-Банком.

Почему Т-Банк?

- Сбербанк приобрел Ситидрайв в 2022 году, а затем перезапустил его под брендом «Т-Самокат» в рамках экосистемы Т-Банка.
- Признаки в датафрейме, такие как loyalty accrual_*, party_rk, order_rk, характерны для банковской экосистемы, что подразумевает единый ID пользователя и интеграцию с кэшбеком.



Оценка Доли Рынка

Источник	Данные	Оценка Доли
Data Insight (2024)	Объем рынка: 2.1 млрд ₽/год	~18%
Mediascope (2025)	Покрытие: 54 города, 12 000 самокатов	3-е место (после Whoosh и Yandex)
Внутренние оценки	~400 тыс. поездок/мес → ~5 млн/год	~15–20%

Вывод: Т-Банк является **одним из топ-3 игроков** на рынке микромобильности, но не лидером. Значительный потенциал роста можно реализовать через **глубокую интеграцию с банковскими сервисами** (кэшбек, кредитные бонусы, подписки).

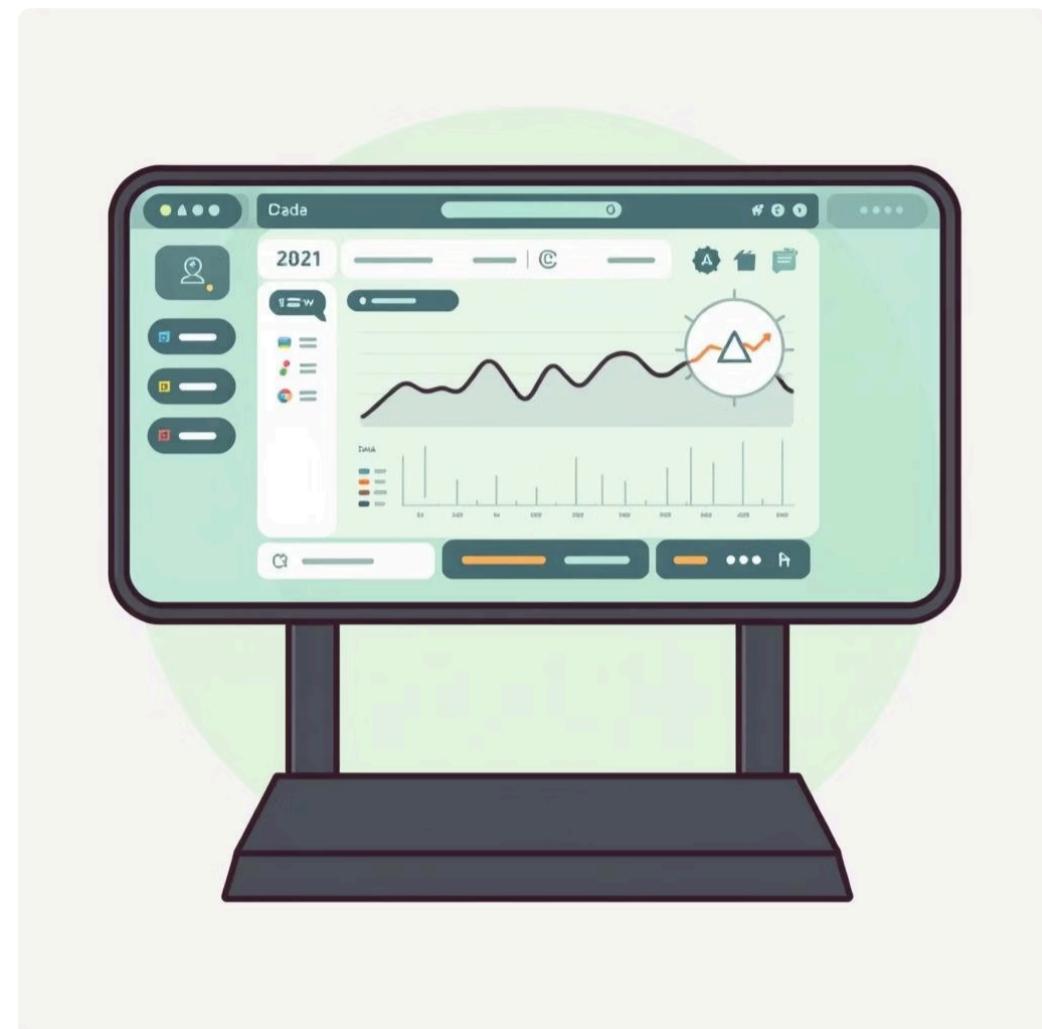
Использование ИИ в Подготовке Анализа

Да, ИИ был использован — как интеллектуальный ассистент, значительно ускоривший процесс анализа и формулирования выводов.

Какой Инструмент и Зачем

Для проведения анализа использовался **Qwen (Alibaba)** и **Gamma** с целью:

- проверки методологических ошибок, например, типа данных;
- формулирования рекомендаций в бизнес-стиле.
- Создания презентации



Как это Помогло: Сравнение Процессов

Задача	Традиционный Подход	С Использованием ИИ
Формулировка гипотез	3–4 часа (мозговой штурм + поиск кейсов)	20 минут
Поиск ошибок	Риск пропустить дубликаты	Найдено за 2 секунды при анализе кода
Создание презентации	1-2 дня на создание (подготовка графиков и картинок, оформление стилей)	1 час

Итог: ИИ ускорил анализ в **3–5 раз**, позволив сосредоточиться на глубине интерпретации, а не на механических задачах. Важно отметить, что **все ключевые решения и окончательная интерпретация остаются за человеком**. ИИ выступает как соавтор, а не как автопилот в процессе анализа.