

# Кейс для продуктовых аналитиков (осень 2025)

---

Андрей

13 ноября 2025 г.

Т-Образование

# Знакомство с датасетом

## О датасете

Данный датасет содержит подробную информацию о поездках на электросамокатах, арендованных через приложение Т-Банка.

Каждый ряд таблицы описывает одну поездку и включает:

- параметры заказа (даты и время, стоимость, дальность, модель самоката);
- характеристики клиента (возраст, пол, образование, семейное положение);
- начисления и бонусы по лояльным программам.

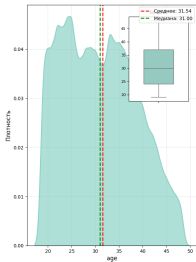
Датасет обезличен и предназначен для анализа пользовательского поведения, оптимизации тарифов и выявления закономерностей в использовании услуги аренды.

Источник: <https://dano.hse.ru/data2024>

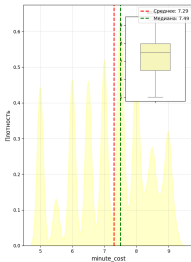
# Распределение величин

## Анализ распределений переменных в наборе данных

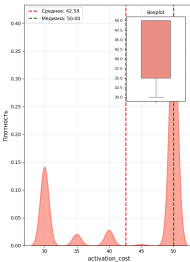
Распределение: age (5-95% перцентиль)  
удалено выбросов: 33184  
(n=396,487)



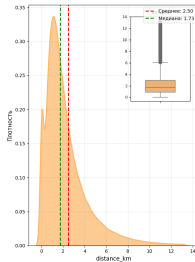
Распределение: minute\_cost (5-95% перцентиль)  
удалено выбросов: 19604  
(n=396,487)



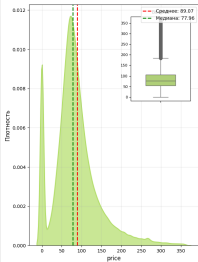
Распределение: activation\_cost (5-95% перцентиль)  
удалено выбросов: 15885  
(n=396,487)



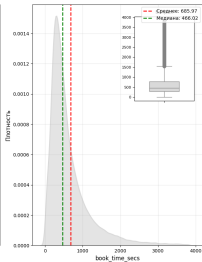
Распределение: distance\_km (99% перцентиль)  
удалено выбросов: 3965  
(n=396,487)



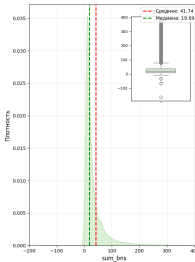
Распределение: price (99% перцентиль)  
удалено выбросов: 3965  
(n=396,487)



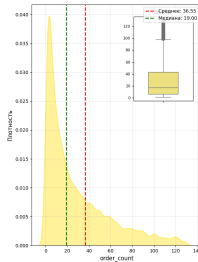
Распределение: book\_time\_secs (99% перцентиль)  
удалено выбросов: 3965  
(n=396,487)



Распределение: sum\_bns (99% перцентиль)  
удалено выбросов: 3965  
(n=396,487)



Распределение: order\_count (5-95% перцентиль)  
удалено выбросов: 19631  
(n=396,487)



# Основные выводы

- Объем данных: **396 749** наблюдений
- Средний возраст пользователей: **31.5** года
  - 75% пользователей младше 37 лет
- Типичная стоимость поездки:
  - Активация: **42.6** (медиана: 50.0)
  - Минута: **7.3** (медиана: 7.5)
  - Итог: **85.9** (медиана: 78.0)
- Стандартный залог: **300** (фиксированная величина для 75% случаев)
- Средняя дистанция: **3.7 км** (медиана 1.7 км)
  - Наличие экстремальных выбросов (макс: 56 012 км)
- Время бронирования: **11.4 минуты** (медиана 7.8 минут)
- Лояльность пользователей:
  - В среднем **36.5** поездок на пользователя
  - Медианное значение **19** поездок

## Аномалии и выбросы

- Отрицательные значения в стоимости поездок
- Экстремальные расстояния (max: 56 012 км)
- Сильное отличие медианы от среднего по ключевым метрикам

## Ключевые выводы

- Датчики расстояния электросамокатов моделей ES\*\*\* часто выходят из строя.
- Бонусы за использование кредитной карты рассчитываются не корректно или не очевидно.
- Надо бы все нормировать все при помощи **StandartScaler**, так мы не потеряем важные выбросы и сохраним основные отношения.

# Продуктовые гипотезы

- **Гипотеза 1: Зависят ли предпочтения пользователей по тарифам от времени суток?**

Гипотеза предполагает, что краткосрочные поездки популярнее утром в рабочие дни.

- **Гипотеза 2: Влияет ли семейное положение человека на вероятность поездок?**

Проверка: Одиноким людям более склонны к поездкам.

- **Гипотеза 3: Как влияет расстояние поездки на выбор тарифа?**

Исследование: склонны ли пользователи выбирать дорогой тариф ( $> 8.22 \text{руб/мин}$ ) при коротких поездках ( $< 286 \text{сек}$ ).

# Гипотеза: Одиноким людям более склонны к поездкам

## Статистическая значимость

- t-тест:  $p=1.17e-18$
- ANOVA:  $p=2.9e-17$
- Mann-Whitney:  
 $p=9.97e-23$
- Kruskal-Wallis:  $p=2.32e-21$

## Размер эффекта

- Cohen's d: 0.098 (малый)
- $\eta^2$ : 0.001 (0.1% дисперсии)

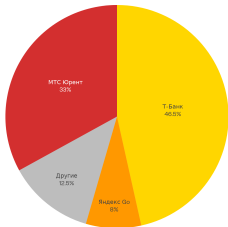
## Среднее число поездок (95% ДИ)

- Одиноким: 6.98 [6.75, 7.21]
- Семейные: 5.62 [5.42, 5.81]
- Прочие: 6.33 [5.62, 7.05]

## Практическая значимость

- Разница: +24.2% vs семейные
- Odds Ratio: 1.18

# Рынок кикшеринга в России: консолидация и рост



- Топ-3 оператора контролируют **>85%** рынка в денежном выражении

## Динамично растущий рынок городской мобильности

- Объем рынка в 2024 г.: **31,1 млрд руб. (+39.8% г/г)**
- Количество поездок: **281,6 млн**
- Аудитория: **25,3 млн пользователей**



# Whoosh (Т-Банк) — абсолютный лидер рынка

Рыночная доля: **47%** от всех поездок в России  
(Каждый второй самокат в стране — *Whoosh*)

## Ключевые показатели (2024)

- **149,7** млн поездок  
(+44% г/г)
- Парк: **245,3** тыс.  
самокатов
- География: **81** город  
(Россия и Латинская Америка)

## Сравнение с конкурентами

- *Whoosh*: **47%**
- *МТС Юрент*: **33%**
- *Яндекс Go*: **8%**
- *Прочие*: **12%**

Использовался Perplexity для

- Поиска информации
- Генерации громоздкого кода и верстки
- Определения доли рынка ТБанка

Так же благодарность Claude Sonet 4.5, DeepSeek, ChatGPT 5.