



Nestlé

Им бы поработать

Роттэ Матвей
Бубеков Константин
Пыльнов Владимир
Зуев Альберт

Наш датасет: «Тинькофф: сервис «Игры»

Описание игры

- Категория
- Цена
- Позиция в «ТОП 5000 популярных игр» в Steam

Описание клиента

- Пол
- Возраст
- Уровень образования
- Ежемесячный доход

Описание заказа

- Дата
- Количество товара
- Город

88699 записей

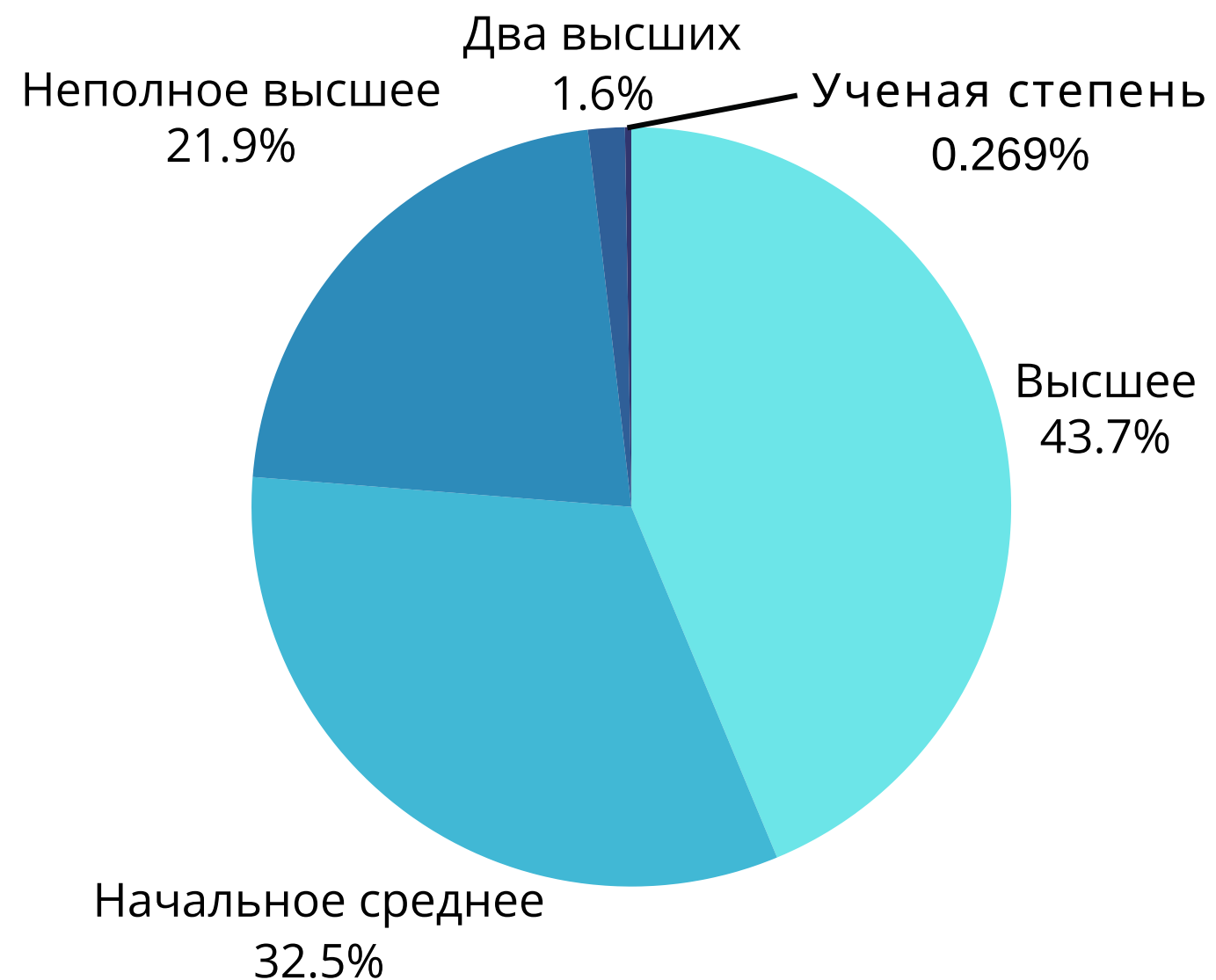
3237 уникальных игр

41298 клиентов

25 уникальных категорий

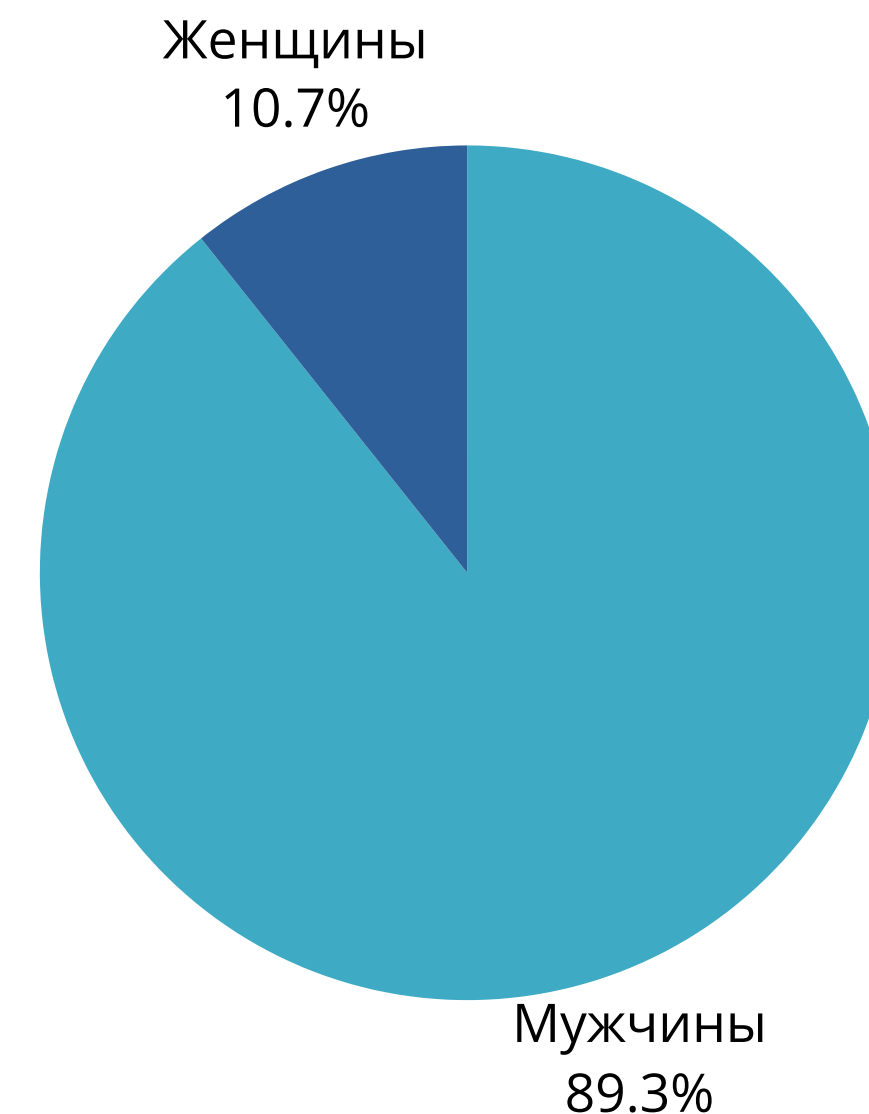
В датасете представлена информация по каждому заказу и играм, которые пользователь приобрёл в рамках этого заказа в период с 1 октября 2022 года по 29 октября 2023 года.

Распределение уровней образования



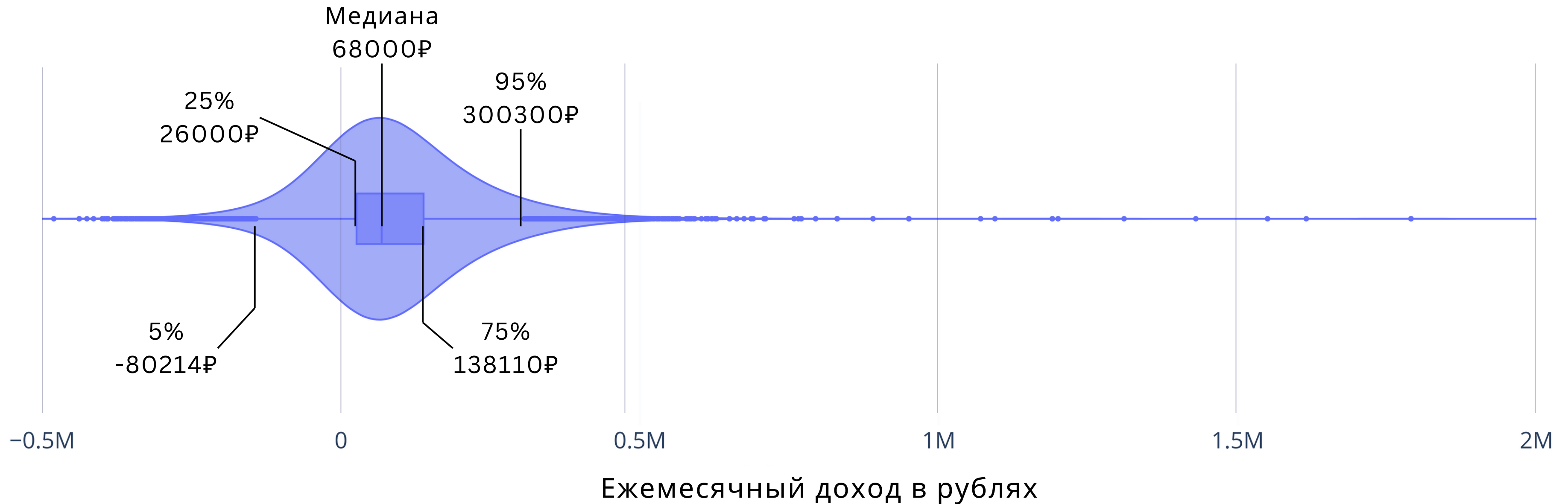
Преобладает высшее образование

Распределение мужчин и женщин



Среди пользователей сервиса «Тинькофф: Игры» мужчин в 9 раз больше, чем женщин

Распределение доходов



КОСВЕННЫЙ ИСТОЧНИК

НЕ ИСПОЛЬЗУЕТСЯ
В ИССЛЕДОВАНИИ

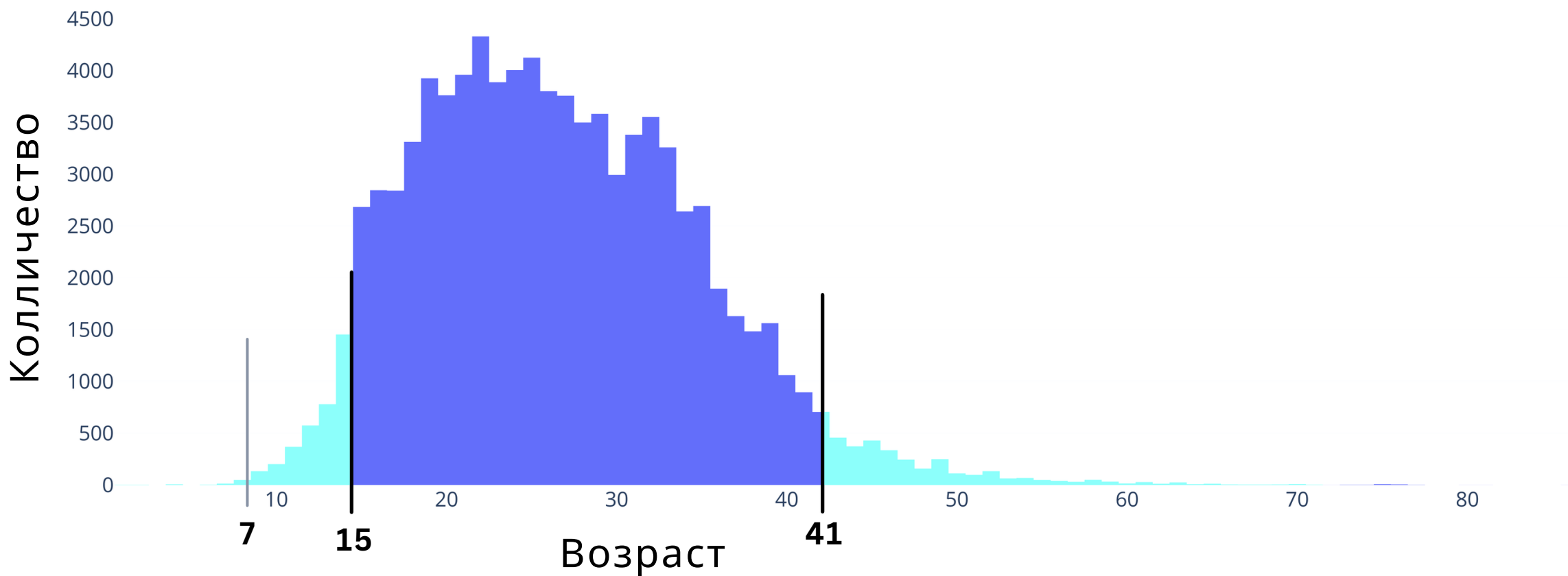
**МЫ НЕ ФИЛЬТРОВАЛИ
ДАННЫЕ ПО ДОХОДУ**

БОЛЬШЕ КОЛИЧЕСТВО
ОТРИЦАТЕЛЬНЫХ ДОХОДОВ

Матрица корреляции показывает отсутствие явных зависимостей

	good_price	age	monthly_income_amt	steam_popularity_score	order_timestamp	education_level
good_price	1	0.063	0.064	-0.137	0.07	0.005
age	0.063	1	0.098	-0.018	-0.113	0.097
monthly_income_amt	0.064	0.098	1	-0.002	-0.035	0.046
steam_popularity_score	-0.137	-0.018	-0.002	1	-0.009	-0.03
order_timestamp	0.07	-0.113	-0.035	-0.009	1	-0.015
education_level	0.005	0.097	0.046	-0.03	-0.015	1

Работа с выбросами



МЛАДШЕ 7 ЛЕТ

Не могут пользоваться сервисом

5 И 95 ПЕРЦЕНТИЛИ

Для однородности данных

ДУБЛИКАТЫ

Произошло ошибочное дублирование данных
или
Игра покупалась не для себя

Удалено **14%** данных

Исследовательский вопрос

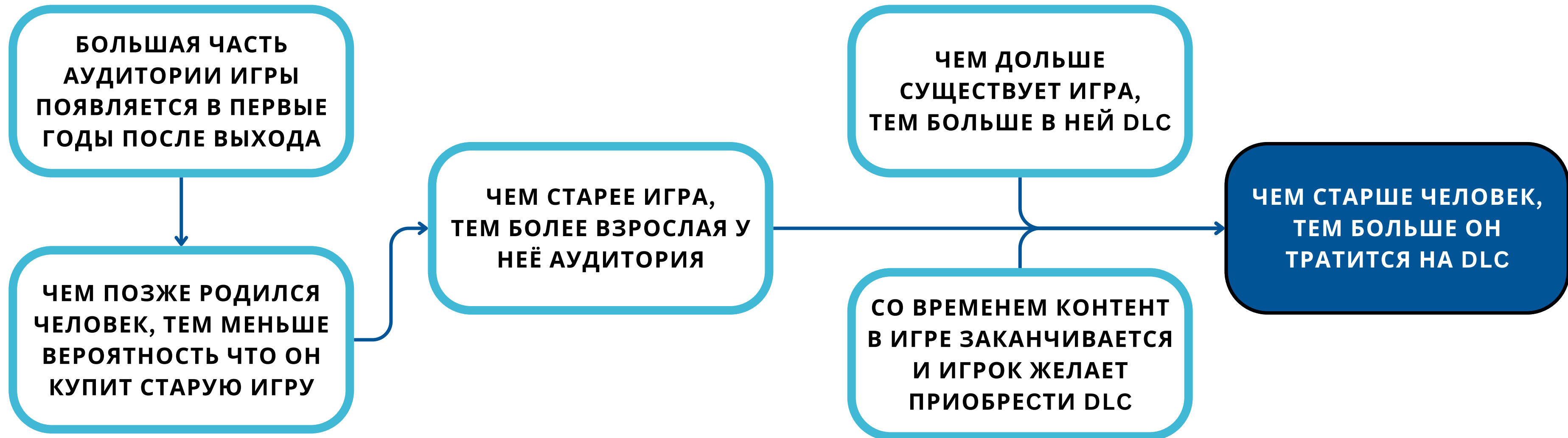


**КАК ВЛИЯЕТ ВОЗРАСТ
ЧЕЛОВЕКА НА
ПОКУПКУ DLC**

Гипотеза

**ЛЮДИ СТАРШЕГО
ПОКОЛЕНИЯ
ПРЕДПОЧИТАЮТ
ПОКУПКУ DLC,
НЕЖЕЛИ ПОКУПКУ
НОВОЙ ИГРЫ**

Механизм



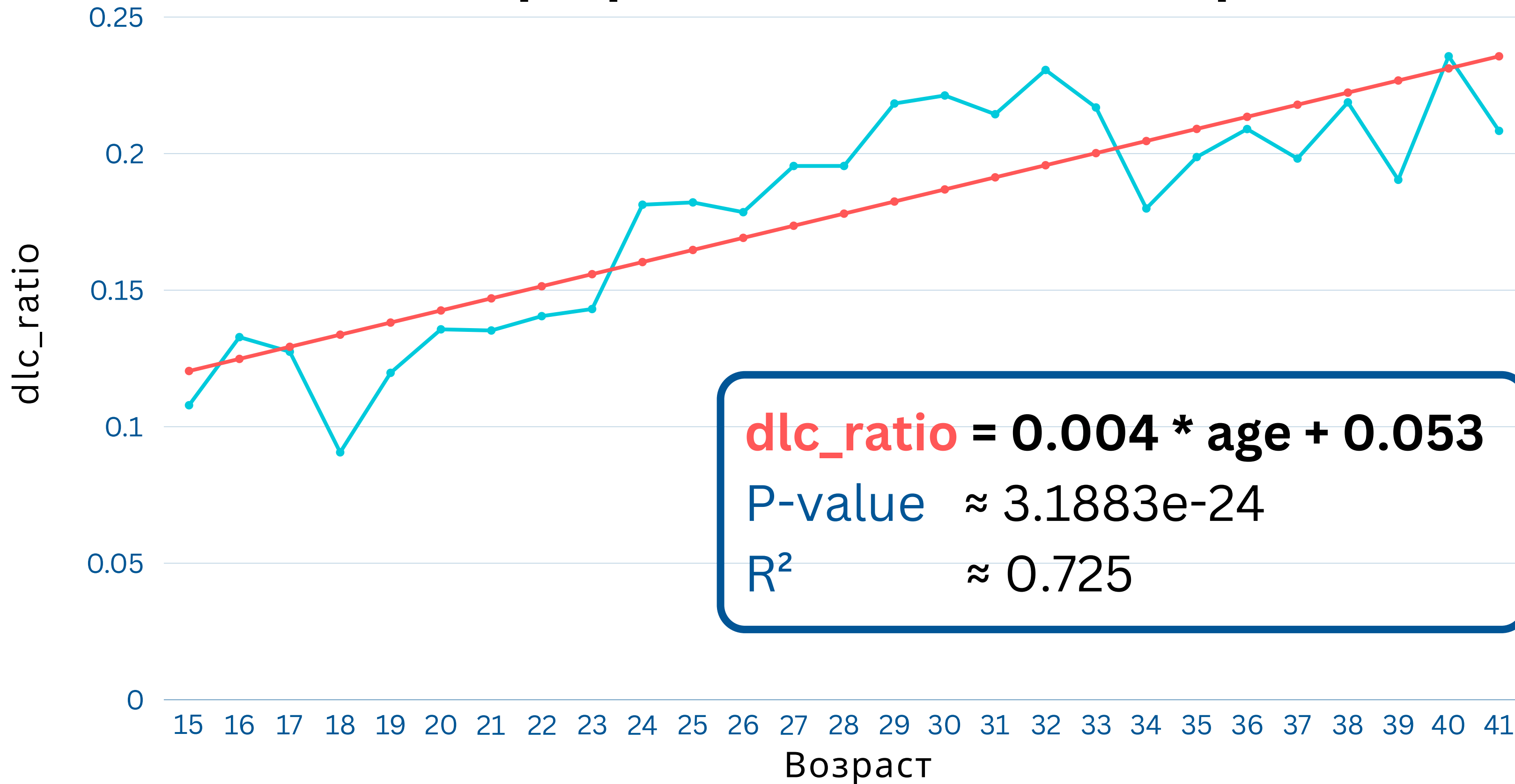
Параметр `dlc_ratio`

$$dlc_ratio = \frac{\text{количество заказов DLC для этого возраста}}{\text{общее количество заказов для этого возраста}}$$



**КОЛИЧЕСТВО ЗАКАЗОВ DLC НАХОДИЛОСЬ
С ПОМОЩЬЮ НАЗВАНИЙ ПО КЛЮЧЕВЫМ СЛОВАМ
“DLC” И “РАСК” В НАЗВАНИИ ИГРЫ, БЕЗ УЧЁТА
ВНУТРИИГРОВЫХ ПОКУПОК**

Линейная регрессия dlc_ratio от возраста



Интерпретация

Статистическая значимость
мат. модели подтвердилась:
р-значение < **0.01**
(высокий порог доверия)

Чем старше человек, тем
вероятнее, что он купит DLC
вместо новой игры

Процент заказов DLC на **72.5%**
описывается возрастом
покупателя

В среднем доля заказов DLC
составляет **17%**
Точка минимума - **18 лет, 9%**
Точка максимума - **40 лет, 24%**

Перспективы



Расширить исследование на все
внутриигровые покупки



Глобализировать результаты, проведя
исследования с другими сервисами



Исследовать предыдущий игровой опыт
пользователей

Ограничения



Перечень игр и DLC ограничен
представленными в сервисе



Отсутствуют данные до 2022 года

Применение

1

Позволяет сервису корректировать стратегию продвижения продукта

Policy Implication

Перераспределение рекламных предложений людям старшего возраста в сторону DLC.

Например частичный пересказ контента DLC, предоставляемый через сториз в приложении Тинькофф , пользователям , уже купившим материнскую игру



**Пыльников
Владимир**

Программист-аналитик
мат.модель

*Действующий
финалист DANO*



**Зуев
Альберт**

Программист

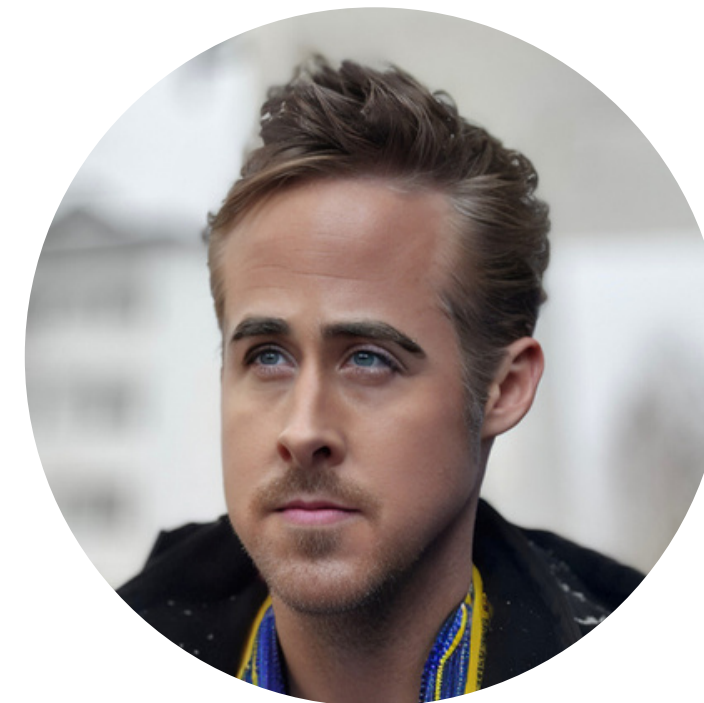
*Действующий
финалист DANO*



**Бубеиков
Константин**

Тимлид, критик,
оратор

*Без Действующий
финалист DANO*



**Роттэ
Матвей**

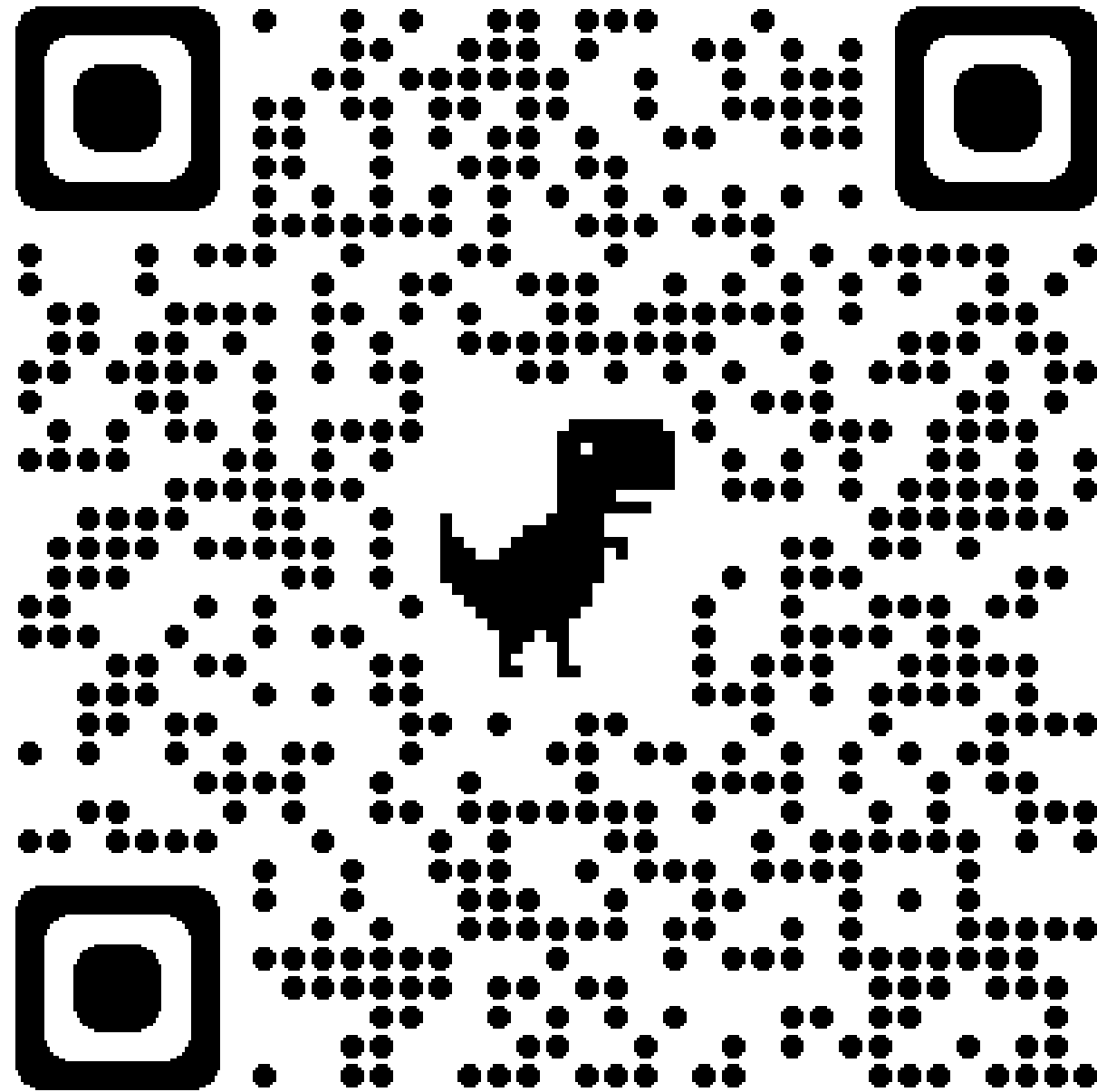
Аналитик
дизайнер

*Действующий
финалист DANO*



Приложение

GitHub



https://github.com/MatveyPRO3/DANO_23_project_tour

Добавленные/изменённые переменные

education_level

Значения были заменены на числа от 0 до 4 в следующем порядке:

- Школьный - 0
- Неполное высшее - 1
- Высшее - 2
- Двойное высшее - 3
- Академическая степень - 4

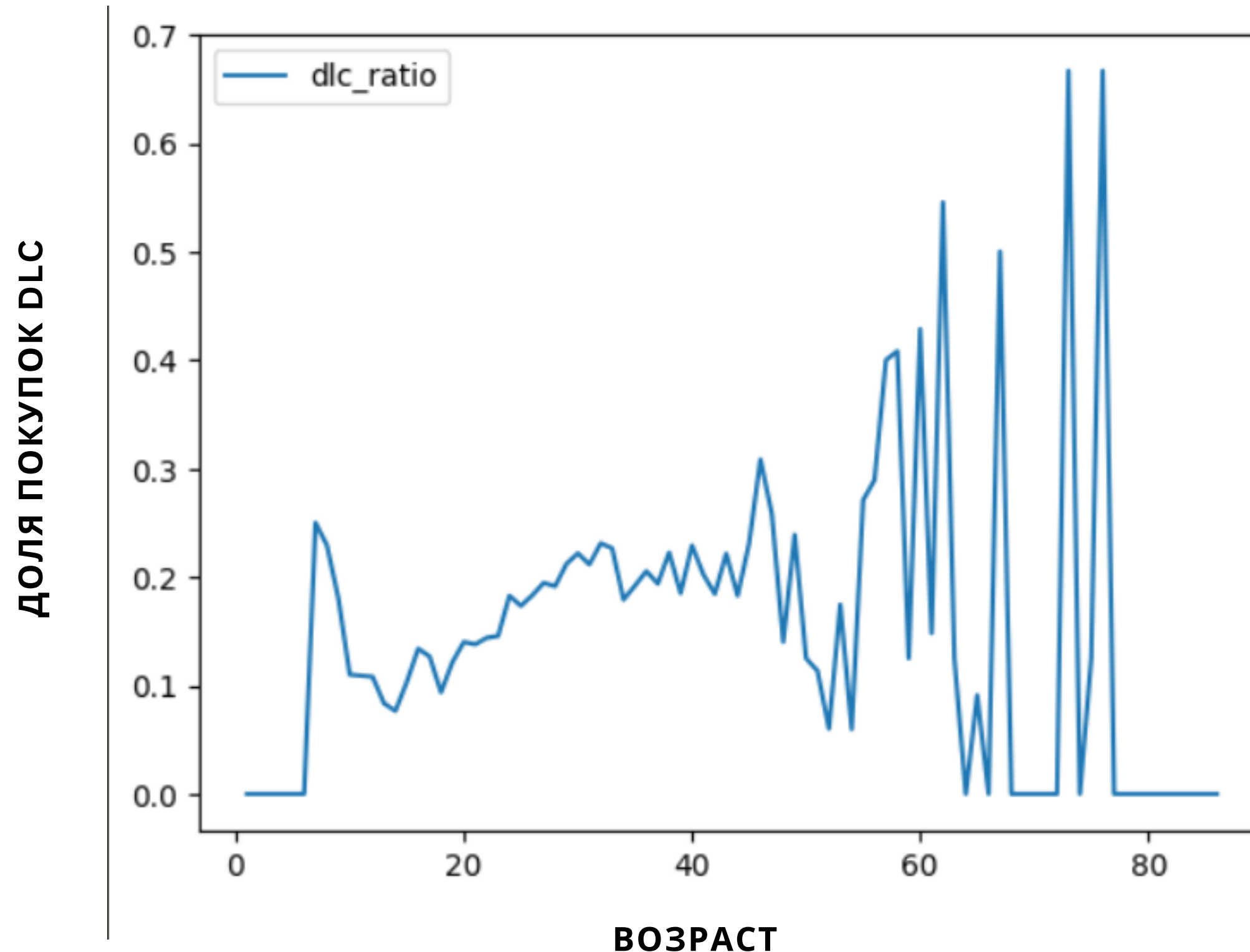
order_timestamp

Дата заказа, переведённая в формат временной метки Unix

is_dlc

Отображает тот факт, является ли позиция DLC. Механика описана на с. 9

Отношение количества покупок DLC к возрасту в датасете с неотфильтрованными выбросами



Распределение абсолютного количества заказов по возрастам

