

Анализ клиентского поведения и построение модели склонности к покупке

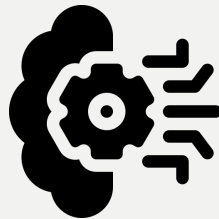
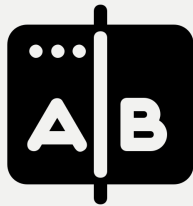
Курс: Data Analyst. Junior

Дата: 20.06.2025

Студент: Сизиков Никита

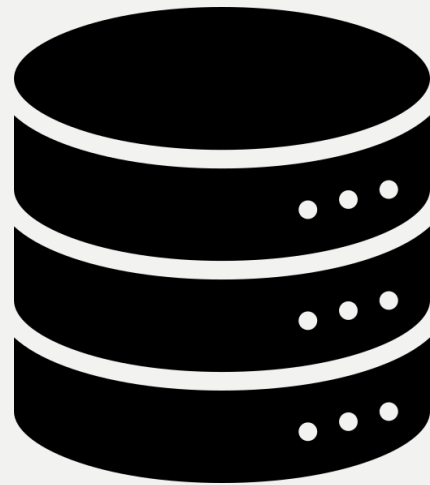
Цель проекта

- Понять поведение клиентов и выявить факторы, влияющие на покупки
- Провести A/B-тест новой модели ETA
- Построить модель склонности к покупке для города 1188, страны 32
- Подготовить рекомендации по маркетинговой стратегии



Данные

- Кол-во клиентов 104 534, транзакций 786 260
- Источники: профили клиентов, история покупок, товары, маркетинговые кампании
- Обработка: очистка, категоризация, кодирование



Восстановление пола пользователей

- Цель: улучшить поведенческий анализ
- Подход: модель случайного леса
- Качество: F-мера = 0,62
- Вывод: позволило уточнить профили клиентов



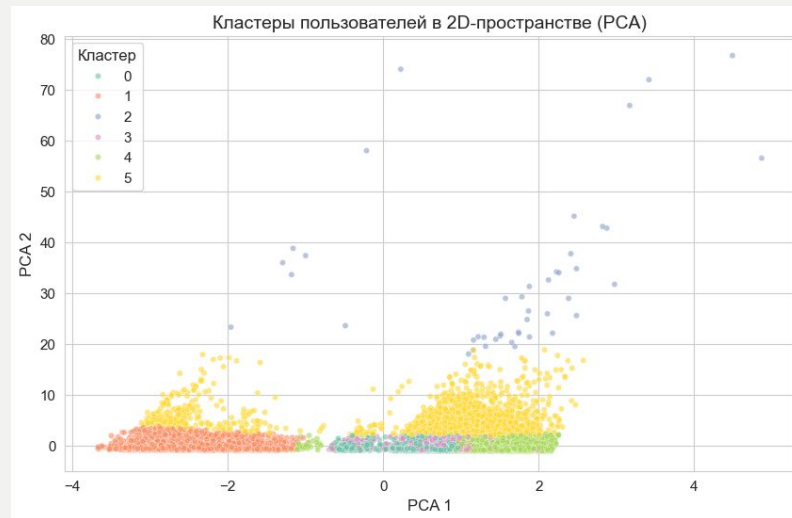
Результаты А/В-тестирования

- Группы: А — контроль, В — тест
- Метрики: количество покупок, средний чек, скидки
- Вывод: скидка повлияла на частоту покупок, но не на общую выручку

| Метрика | Контроль | Тест |
|-----------------------|----------|--------|
| Конверсия | 1,00 | 1,00 |
| Среднее число покупок | 11,02 | 12,80 |
| Средняя сумма трат | 62 495 | 65 695 |
| Доля со скидкой | 31,7% | 31,9% |

Кластеризация клиентов

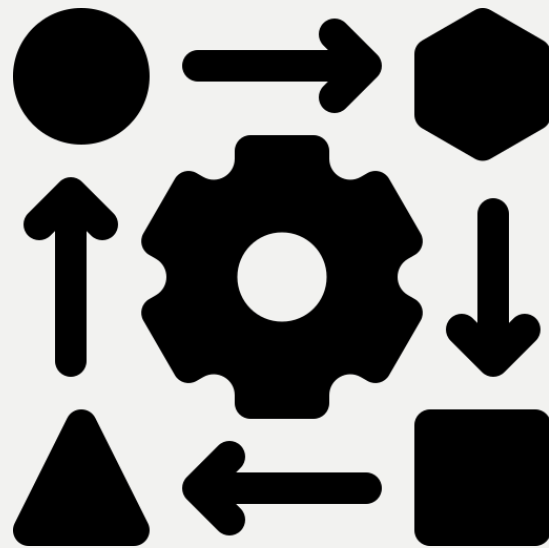
- Метод: KMeans, 6 кластеров
- Основные характеристики кластеров: возраст, образование, кол-во покупок, реакция на скидки



- Вывод: разные сегменты требуют разных стратегий

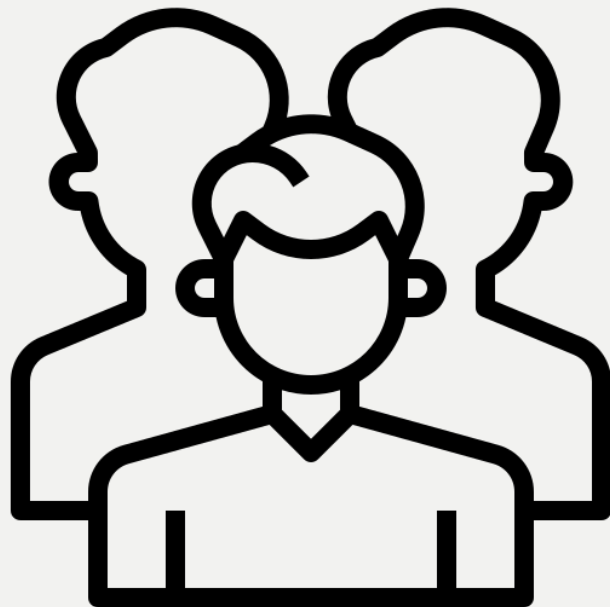
Модель склонности к покупке

- Целевая аудитория: страна 32, город 1188
- Алгоритм: RandomForestClassifier
- ROC-AUC = 0,51
- Фичи: возраст, пол, образование, категория товара, скидка, кол-во покупок и трат



Портрет целевого покупателя

- Молодые (до 30 лет)
- Делают больше покупок
- Реже используют скидки
- Наиболее часто покупают:
 - Одежду
 - Обувь
 - Экипировку
- Скидка не является ключевым мотиватором



Рекомендации

- Запуск коммуникации через [email / push / баннеры]
- Акцент на товары категории X и Y
- Использовать персонализированные скидки
- Повторные кампании в определенные дни



Заключение

- Определен профиль потенциальных покупателей
- Выделены ключевые категории покупателей
- Построена модель приоритезации коммуникаций

