

Отчёт: Анализ поведения пользователей и прогноз склонности к покупке

Проект для заказчика: Повышение эффективности маркетинговых коммуникаций

1. Цель исследования

Заказчик планирует запуск маркетинговой кампании в городе 1188 (страна 32).

Необходимо:

- Оценить поведение пользователей;
- Выделить группы клиентов (кластеризация);
- Построить модель склонности к отклику;

Сформировать рекомендации по персонализированному маркетингу.

2. Данные

Анализ проведён на основе объединённых данных о:

- Персональной информации о пользователях (пол, возраст, образование, город/страна проживания);
- Персональных коэффициентах клиентов;
- Покупках (продукт, цвет, гендерная принадлежность, наличие скидки, стоимость, дата);
- Истории участия в предыдущих маркетинговых кампаниях (A/B-тест).

Объём данных: 104 534 уникальных клиента и 786 260 покупок.

3. Анализ поведения пользователей

Ключевые наблюдения:

- Средний возраст покупателя — 36 лет;
- Большинство клиентов имеют среднее образование;
- В среднем совершается ~6 покупок, со средней суммой трат около 34 000 у.е.;
- Преобладают покупки одежды и обуви;
- Повышенный интерес к скидкам наблюдается у пользователей старших возрастов.

4. Анализ А/В-теста

Цель: оценить влияние email-рассылки на поведение.

Метрики и результаты:

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%

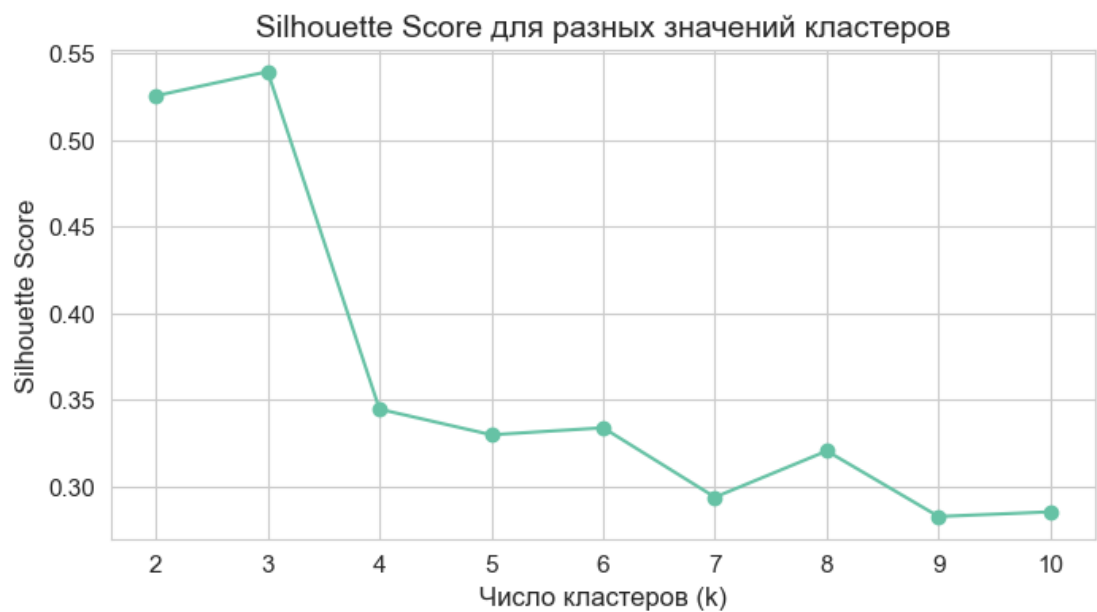
Т-тесты:

- Существенная разница найдена **в числе покупок** ($p\text{-value} = 0,0057$);
- Разницы по сумме трат и скидкам — **нет**.

Вывод: Email-рассылка увеличивает частоту покупок, но не влияет на средний чек.

5. Кластеризация клиентов

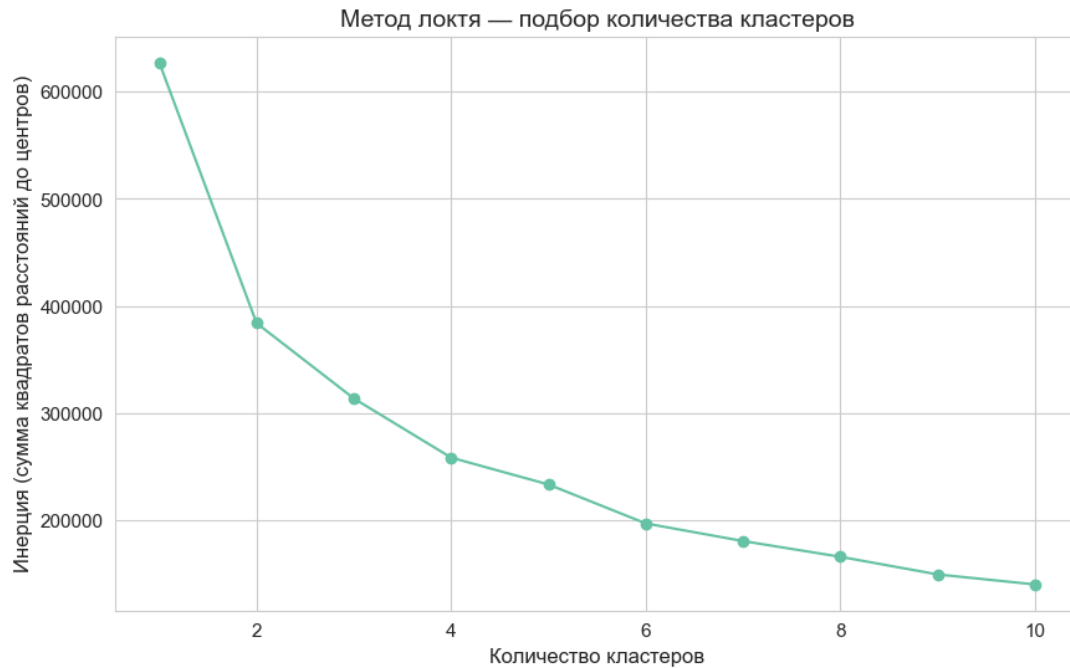
Выбранный метод: Cluster Silhouette (3 кластера)



Основные группы:

- Кластер 0 — Средних лет с низкой активностью, образование среднее;
- Кластер 1 — Молодые, активность низкая, образование высшее;
- Кластер 2 — Аномально активные, с очень высокими тратами;

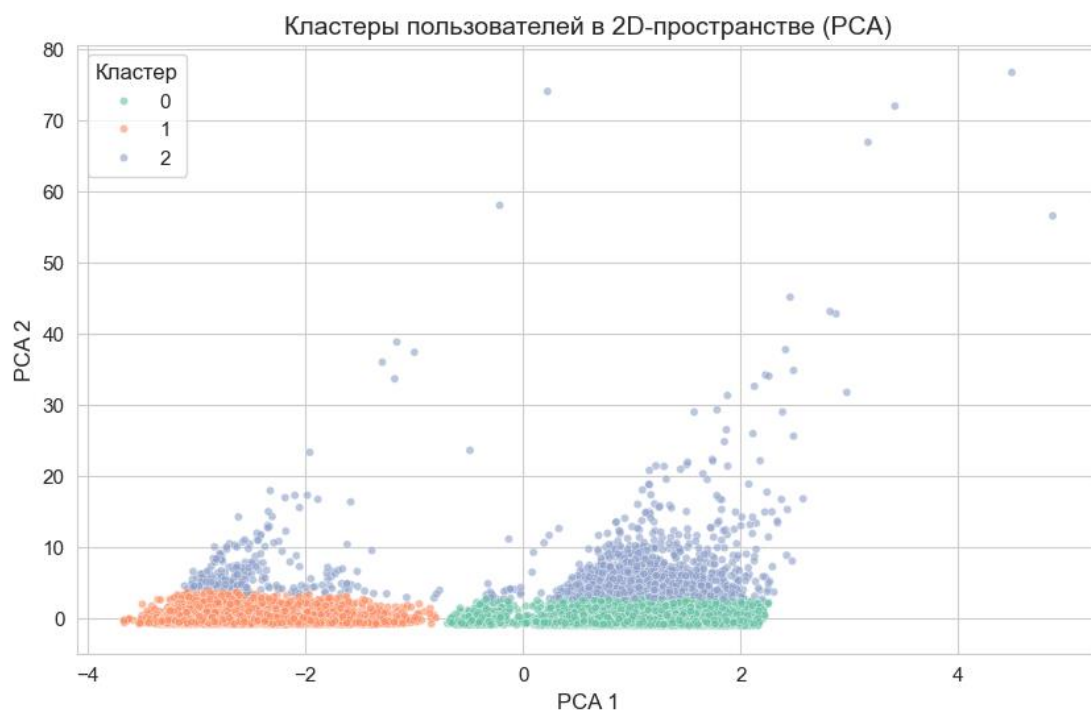
Протестированный дополнительно метод: K-Means (6 кластеров, по методу локтя)



Основные группы:

- Кластер 0 — Молодые с низкой активностью;
- Кластер 1 — Молодые активные;
- Кластер 2 — Аномально активные (возможно, корпоративные);
- Кластер 3 — Средний возраст, много скидок;
- Кластер 4 — Пожилые с умеренной активностью;
- Кластер 5 — С активностью выше среднего.

Сегменты различаются по возрасту, количеству покупок и тратам.



6. Модель склонности к покупке

Обучение на выборке из A/B-теста (target: группа тест/контроль)

Модель: RandomForestClassifier

Признаки: возраст, пол, траты, скидки, кластер, образование

Результаты:

Метрика	Значение
F1 для класса 1	0.71
Recall	0.80
AUC ROC	0.51 (низкая разделяющая способность)

Модель захватывает почти всех, кто откликнулся, но плохо отделяет неоткликнувшихся. Подходит для массовой охватной рассылки.

7. Портрет потенциального покупателя

Разбивка по вероятности отклика (qcut на 3 группы).

Группа “High” (наиболее вероятный отклик):

- Возраст: преимущественно 40+ (из-за доминирования кластера 0)
- Покупательская активность: умеренная (6–7 покупок)
- Средние траты: ~33–35 тыс. Р
- Образование: чаще среднее
- Каналы коммуникации: восприимчивы к контентным форматам
- Наиболее часто покупают:
 - Одежду
 - Обувь
 - Экипировку
- Скидка не является ключевым мотиватором.

8. Рекомендации

8.1. Основной фокус — на массовый кластер (Кластер 0)

Сегмент охватывает большинство high-response клиентов ($\approx 79\%$)

Характеристики: 40+, средний доход, базовая экипировка

Рекомендации:

- Использовать контентные форматы: подборки, советы, обзоры товаров
- Продвигать новинки в популярных категориях (одежда, обувь, экипировка)

- Выстраивать доверие к бренду через качественный сервис и гарантии
- Не строить упор на скидки — важнее ценность и простота выбора

8.2. Поддержка digital-коммуникаций для молодёжного сегмента (Кластер 1)

Активны, молоды, более восприимчивы к digital-форматам (push, email, приложение)

Рекомендации:

- Использовать визуальные, геймифицированные механики: квизы, баннеры, карусели товаров
- Отправлять персонализированные предложения через мобильные каналы
- Акцент на технологичность и удобство (например, быстрый подбор товаров, one-click покупка)

8.3. Индивидуальный подход к VIP-сегменту (Кластер 2)

Покупают редко, но на очень крупные суммы

Мало отзывчивы к скидкам, но восприимчивы к эксклюзивности

Рекомендации:

- Продвигать программы лояльности премиум-уровня
- Использовать приглашения на закрытые распродажи, персональные менеджеры
- Акцент на качество, сервис, статус, а не на массовые акции

8.4. Общие рекомендации

Сегментация по типу поведения даёт лучшие результаты, чем просто демография

При планировании акций не полагаться только на скидки — высокая чувствительность к ним наблюдается не у всех групп

Продумать разные механики под разные кластеры:

- Кластер 0: подборки, обзоры
- Кластер 1: digital-игры, мобильные уведомления
- Кластер 2: эксклюзив и персонализация

9. Ограничения в рамках исследования

- Небольшая обучающая выборка A/B;
- Отсутствие негативных исходов покупки;
- Возможные искажения от сезонности и города.

10. Вывод

Мы определили профили наиболее вероятных покупателей, выделили ключевые категории и подготовили модель для приоритезации коммуникаций.

Рекомендуем использовать полученные данные для таргетированной кампании и протестировать эффективность сегментированного подхода.