# **Аналитический отчет**

## **Резюме**

**Цели исследования:**

Анализ маркетинговых компаний посредством A/B-теста, выявление сегментов клиентов с помощью кластерного анализа, построение модели склонности клиента к покупке.

**Ключевые выводы:**

* Основная цель первой маркетинговой компании выполнена – люди стали покупать больше товаров. Но, вместе с этим, упал и средний чек пользователя.
* Выявлено 4 основных сегмента пользователей, различающихся по различным признакам, таким как: пол, возраст, кол-во купленных товаров по скидке, кол-во купленных товаров.
* Была построена модель прогнозирования склонности пользователей к покупке, ее точность – 71%, что позволит довольно успешно предсказывать поведение клиентов.

**Общие рекомендации:**

* Продолжить проводить маркетинговые компании, однако пересмотреть формат скидок.
* Стимулировать увеличение среднего чека у пользователей через кросс-продажи и upselling.
* Запускать больше маркетинговых компаний со скидками, ориентируемых на женщин.
* Запустить программу лояльности или накопительные скидки для более зрелых клиентов.

## **Описание данных и методы обработки**

Мною были собраны персональные данные о клиентах, данные об их покупках и данные с индивидуальными коэффицентами. Проведена обработка наборов данных: заполнены пропуски, выявлены и удалены аномалии, выполнена проверка на дубликаты, обработаны некоторые признаки.

Важно отметить, что для заполнения информации о половой принадлежности среди потерянных данных о клиентах была настроена модель определение пола клиента по другим данным. Модель получилась очень точной (97% точности)

## **Анализ A/B-теста**

Был проведен эксперимент, анализирующий эффективность первой маркетинговой компании. Эта кампания включала в себя предоставление персональной скидки 5 000 руб. клиентам через email-рассылку.

В тестовую группу попали клиенты, которым была предложена скидка.

В контрольную группу попали клиенты со схожими социально-демографическими признаками и покупками, но которым скидку не предложили.

Метриками были выбраны: среднее количество заказов, средний чек.

Для проверки статистической значимости был выбран T тест Стьюдента.

**Результаты:**

* В тестовой группе наблюдается, что среднее количество заказов больше, разница статистически значима.

* В тестовой группе наблюдается статистически значимое уменьшение среднего чека.

**Интерпретация результатов:**

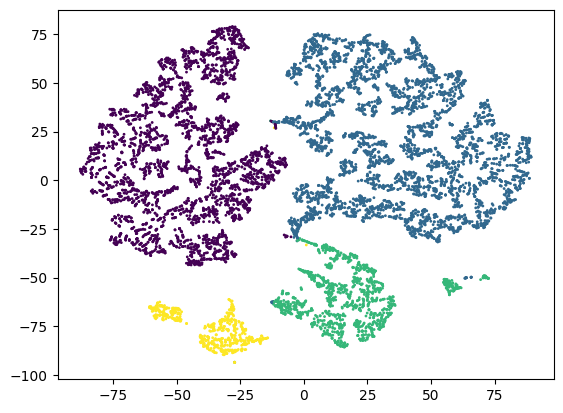
Основная цель маркетинговой компании достигнута, люди стали покупать больше товаров, однако упал средний чек на пользователя, из чего следует, что люди пусть и стали покупать больше товаров, но они концентрировались на более дешевых товарах.

**Бизнес-рекомендация:**

С этим очень хорошо помогла бы процентная скидка: чем дороже товар - тем выгоднее клиенту покупать его. Так же можно повысить средний чек через cross-selling и upselling, когда клиентам, при покупке товара, предлагается такой же товар, но качеством лучше, либо мелкие товары, дополняющие основной товар.

## **Кластерный анализ**

Для сегментации данных на кластеры был проведен кластерный анализ разными методами, наиболее полезным выбран метод k-средних (K-means). Были получены 4 кластера, которые по большей части делят людей по их полу и возрасту:  
1 и 4 кластеры – это зрелые и более молодые женщины соответсвенно  
2 и 3 кластеры – это зрелые и более молодые мужчины соответсвенно



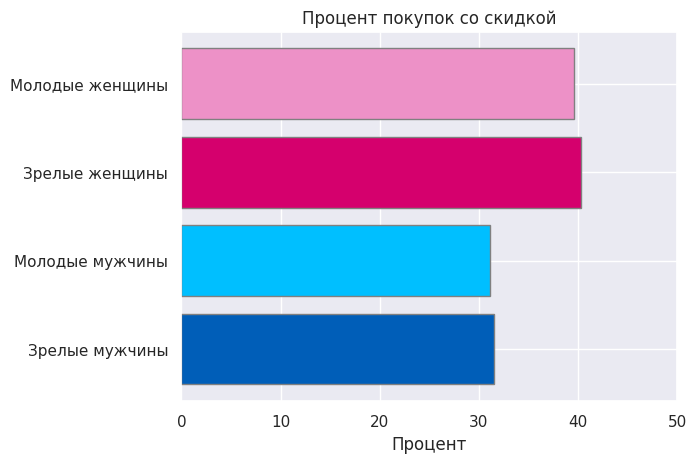
На рисунке наглядно показывается разбиение клиентов на кластеры.

Два больших кластера – мужчины,

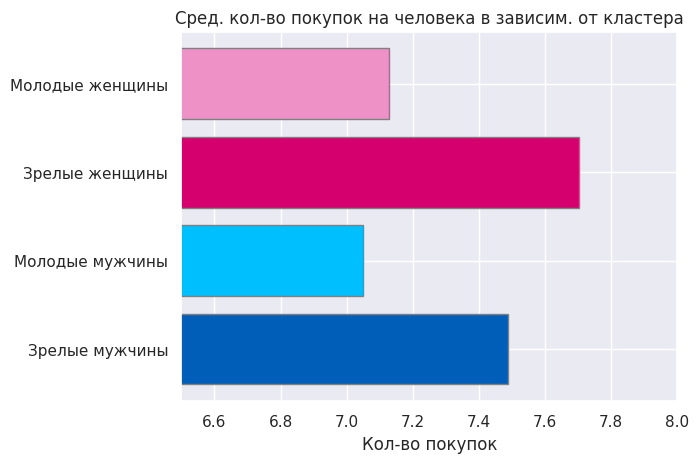
Два меньших кластера - женщины

В результате анализа было выяснено следующее:

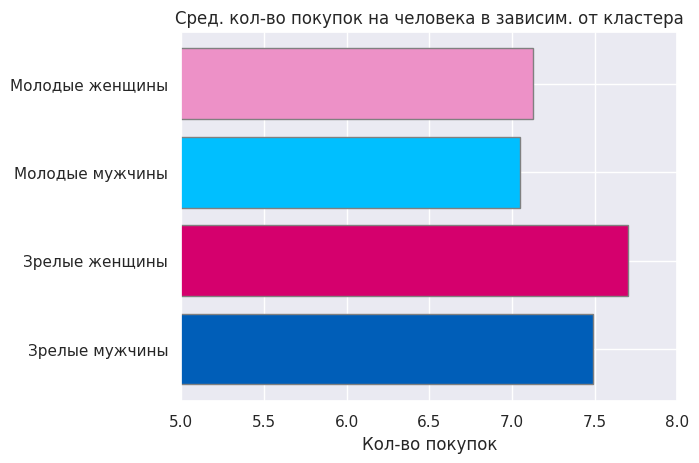
1. Женщины обоих возрастных категорий чаще покупают товары по скидке, чем мужчины.



1. Более зрелые люди делают больше покупок чем молодые.



1. Агрегируя по возрастным категориям, женщины делают больше покупок нежели мужчины.



**Рекомендации по работе с кластерами:**

* Пополнить ассортимент товаров для женских кластеров. В особенности, это пляжные принадлежности (купальник, плавки) и повседневные вещи (брюки, лиф, футболка). Также, так как они более склонны к покупкам по скидке, для них стоит вводить больше персональных предложений и акций разного рода.
* Учитывая меньшую активность мужчин в покупках, предложить им специальные предложения на товары, которые они чаще всего выбирают. Для молодых мужчин фокус стоит сделать на шорты и брюки спортивного стиля, а для зрелых мужчин — на брюки и футболки более универсального характера.

## **Предсказательная модель склонности к покупке**

Был создан датасет с различными данными пользователей, включая персональные данные, данные о покупках и данные об участии в маркетинговых компаниях. Целевой переменной был обозначен признак, принимающий только два числа. Он принимал значение 1 если человек совершал покупок на сумму больше среднего чека, значение 0 в ином случае.

Эти данные были закодированы с помощью метода LabelEncoder, затем нормализованы с помощью метода StandardScaler. Далее данные были разделены на тренировочную и тестовую выборки, где тестовая – выборка посетителей из города 1188, а контрольная – выборка посетителей из всех других городов.

Моделью была выбрана логистическая регрессия (LogisticRegression).

Модель была обучена и проверена с помощью кросс-валидации. Получена следующая точность модели – 71%.

## **Заключение**

В ходе проведенного анализа получены значимые результаты, которые можно использовать для повышения эффективности маркетинговых кампаний и оптимизации бизнес-процессов.

1. **А/B тестирование** продемонстрировало, что маркетинговая кампания с фиксированной скидкой привела к увеличению среднего количества заказов в тестовой группе с *11.2* до *12.9* покупок. Однако средний чек снизился с *5831* до *5596* единиц, что указывает на смещение спроса в сторону более доступных товаров. Рекомендуется пересмотреть формат предоставления скидок, заменив фиксированную сумму на процент от стоимости товара, чтобы стимулировать покупки более дорогих позиций. Также целесообразно внедрить стратегии кросс-продаж и upselling для увеличения среднего чека.
2. **Кластерный анализ** позволил выделить четыре ключевых сегмента клиентов, отличающихся по полу и возрасту. Выявлено, что женщины обеих возрастных категорий совершают больше покупок, особенно со скидками. Рекомендуется расширить ассортимент популярных товаров, таких как купальники и легинсы, а также активизировать персонализированные предложения для женской аудитории. Для мужских сегментов можно внедрить более целевые промоакции, акцентируя внимание на востребованных товарах, таких как спортивные брюки и шорты.
3. **Модель прогнозирования склонности к покупке** продемонстрировала точность на уровне 71%, что делает её применимой для сегментации клиентов и более точного таргетинга маркетинговых акций. Дальнейшее повышение точности возможно за счет увеличения числа признаков и интеграции данных из дополнительных источников.