

## Аналитический отчет:

- эффективность маркетинговой кампании
  - факторы, способные повысить продажи
- 

\*на основании данных магазина спортивных товаров

Мария Смирнова



# Задачи:

---

- определить эффективность маркетинговой компании;
- выявить кластеры, на которые разбивается аудитория;
- предложить методы работы с каждым кластером;
- построить модель склонности клиента к покупке определённого товара.

# Исходные данные:

---

- Личные данные клиентов (*personal\_data*):

ID клиентов, их пол, возраст, образование, страна и город проживания;

- данные с персональными коэффициентами (*personal\_data\_coeffs*);

- данные о покупках (*purchases*):

ID клиентов, название товара, цвет, стоимость, гендерная принадлежность, скидка и дата покупки;

- данные об участии в маркетинговой компании (*ids\_first\_company\_positive.txt*).



Данные нуждались в улучшении качества.

Для решения этой задачи были :

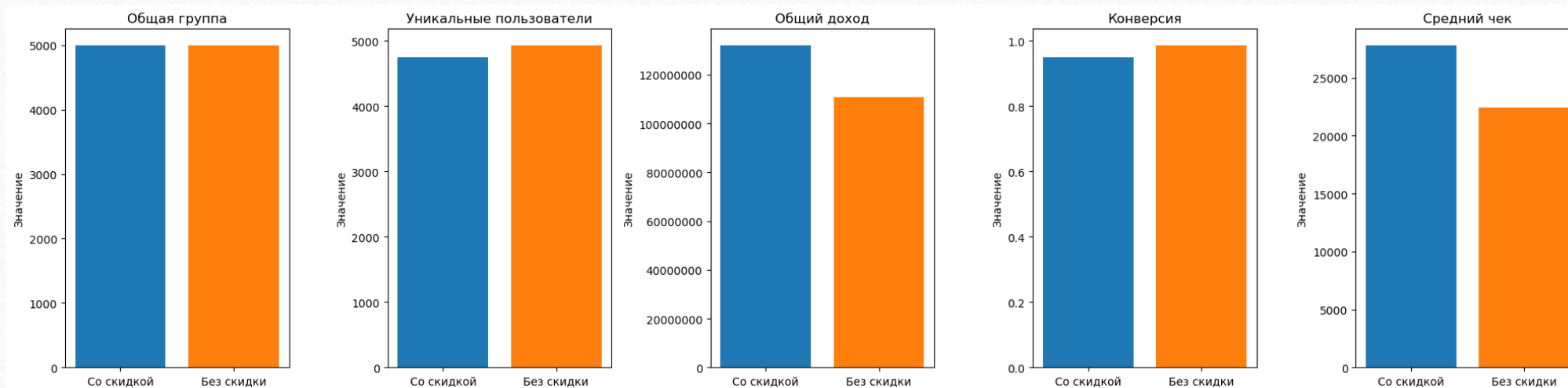
---

- восстановлены утерянные данные по части клиентов(пол клиентов);
- отфильтрованы данные только по стране с кодовым цифровым значением 32;
- обработаны пропуски в данных;
- выявлены и обработаны аномальные значения;
- данные о цвете товаров сделаны более однородными.

**Рекомендация:**

доработать процесс получения данных для улучшения их полноты и качества .

# Метрики для оценки эффективности маркетинговой компании





# Z-тест для конверсии

---

- Низкое p-значение ( $< 0.05$ ) указывает на то, что разница в конверсии между пользователями со скидкой и без скидки статистически значима.
- Отрицательная Z-статистика показывает, что конверсия в группе со скидкой ниже, чем в группе без скидки.
- Следовательно, несмотря на значительную долю пользователей со скидкой, пользователи без скидки конвертируются в покупки эффективнее.

# Т-тест для среднего чека

---

- Т-статистика равна -3.11 с р-значением 0.0019. Низкое р-значение ( $< 0.05$ ) указывает на то, что различие в среднем чеке между пользователями со скидкой и без скидки статистически значимо.
- Отрицательная Т-статистика показывает, что средний чек для пользователей со скидкой выше, чем для пользователей без скидки.
- Это свидетельствует о том, что пользователи, получившие скидку, тратят в среднем больше, что может свидетельствовать об эффективности скидок для увеличения дохода с покупок.



# Бизнес-рекомендации по результатам А/В-теста маркетинговой кампании

---

- *Конверсия*

**Вывод:** группа пользователей со скидкой показала значимо более низкую конверсию (94.98%) по сравнению с группой без скидки (98.70%) на основе Z-теста. Вероятно, предложение скидок, вероятно, не побудило больше пользователей совершить покупку, а в некоторых случаях могло снизить мотивацию.

**Рекомендация:** пересмотреть стратегию скидок. Возможно, скидка воспринимается не как стимул, а наоборот, как снижающее восприятие ценности продукта предложение. Альтернативным решением может стать тестирование других типов стимулов, например, программ лояльности, бесплатных подарков или бонусов.

- *Средний чек*

**Вывод:** средний чек среди пользователей со скидкой оказался выше (27,822.45) по сравнению с пользователями без скидки (22,442.42). Т-тест для этой метрики показал статистически значимую разницу (Т-статистика = -3.11, р-значение = 0.0019).

**Рекомендация:** хотя скидка не улучшила конверсию, она привела к увеличению среднего чека. Это может означать, что пользователи, решившие совершить покупку, предпочитают делать более крупные заказы. Для увеличения продаж можно внедрить предложения «купи больше, получи больше», пакетные предложения или скидки на следующий заказ.



# Бизнес-рекомендации по результатам А/В-теста маркетинговой кампании

---

- **Общий доход**

**Вывод:** Несмотря на более высокий средний чек в группе со скидкой, общий доход был выше в группе без скидки. Это связано с более высокой конверсией среди пользователей без скидок.

**Рекомендация:** Возможно, стоит пересмотреть политику предложений и учитывать влияние скидок на общее количество покупок. Например, тестирование динамических скидок в зависимости от суммы заказа или применение персонализированных предложений на основе поведения пользователей.

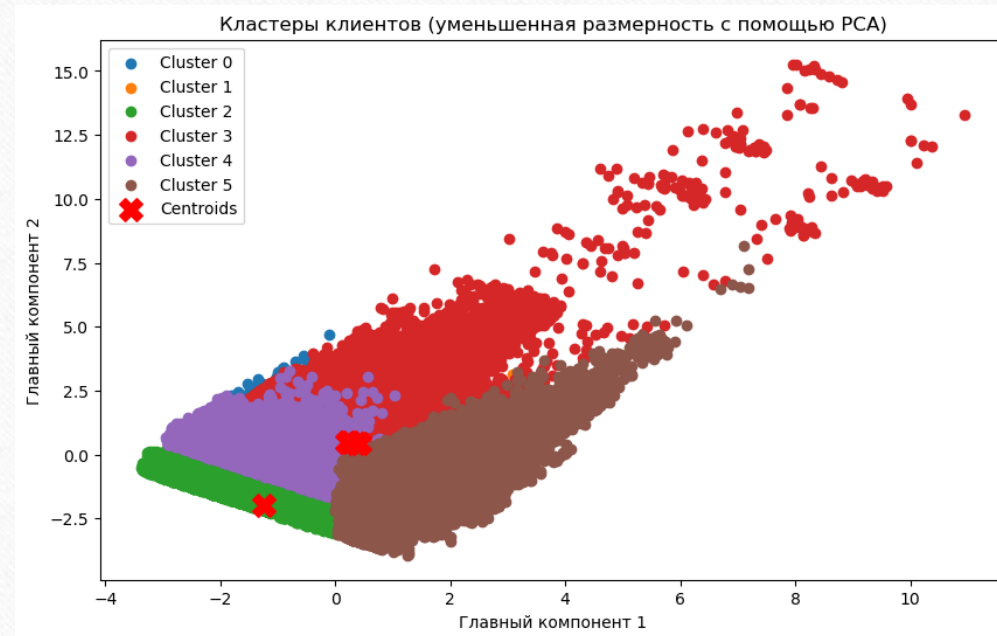
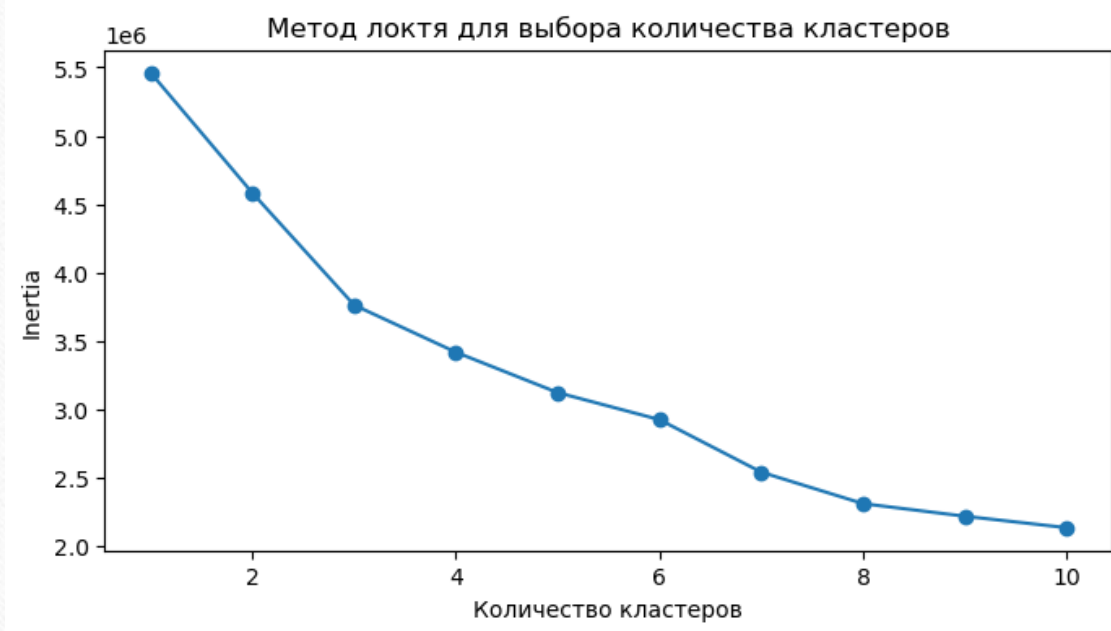
- **Общие выводы и стратегия**

Скидочная стратегия **неэффективна** для увеличения **объема покупателей**, но **эффективна** для повышения **среднего чека**.

Рекомендуется провести дополнительное **тестирование альтернативных стимулов**, которые будут влиять на конверсию, и при необходимости пересмотреть ценовую стратегию для долгосрочного увеличения прибыли.

Инвестиции в **дополнительные механизмы удержания** клиентов, такие как программы лояльности, могли бы повысить ценность клиента без необходимости значительного снижения цен.

# Кластерный анализ





Клас тер	Средний возраст (лет)	Средняя стоимость покупки	Предпочти тельные товары	Влияние скидки	Рекомендации
0	40.97	4451.12	Сумка FILA	Большинство покупок без скидки	Покупатели среднего возраста, предпочитающие аксессуары. Рекомендуется предлагать новые модели сумок, акцентировать внимание на качестве и функциональности.
1	41.68	4734.31	Брюки мужские Demix	Покупки почти равномерно со скидкой и без	Внедрить кампании со скидками, регулярно обновлять ассортимент мужской спортивной одежды.
2	42.76	3150.17	Футболка женская Demix	Все покупки совершаются по скидке	Регулярно предлагать акции на женскую спортивную одежду, особенно футболки.
3	41.21	7854.77	Брюки мужские Demix	Большинство покупок без скидки	Платежеспособная аудитория, предпочитающая мужскую одежду. Рекомендуется предлагать премиальные коллекции, подчеркивать качество товаров.
4	43.68	4592.21	Брюки женские Outventure	Все покупки совершаются без скидок	Ценят женскую одежду высокого качества. Можно предложить эксклюзивные коллекции и персонализированные подборки.
5	23.37	5763.02	Шорты мужские Demix	Покупки равномерно со скидкой и без	Молодая аудитория, интересующаяся мужской одеждой. Рекомендуется таргетированная реклама в соцсетях, акцент на новинках ассортимента.

# Модель склонности клиента к покупке определённого товара

---

- Аудитория:

жители страны 32 города 1 188

- Признаки для обучения модели:

возраст клиента 'age',

пол клиента 'gender',

дата покупки 'dt',

кластер клиентов 'cluster',

гендерная принадлежность потенциальных покупателей товара 'product\_sex'



# Вывод по данным модели

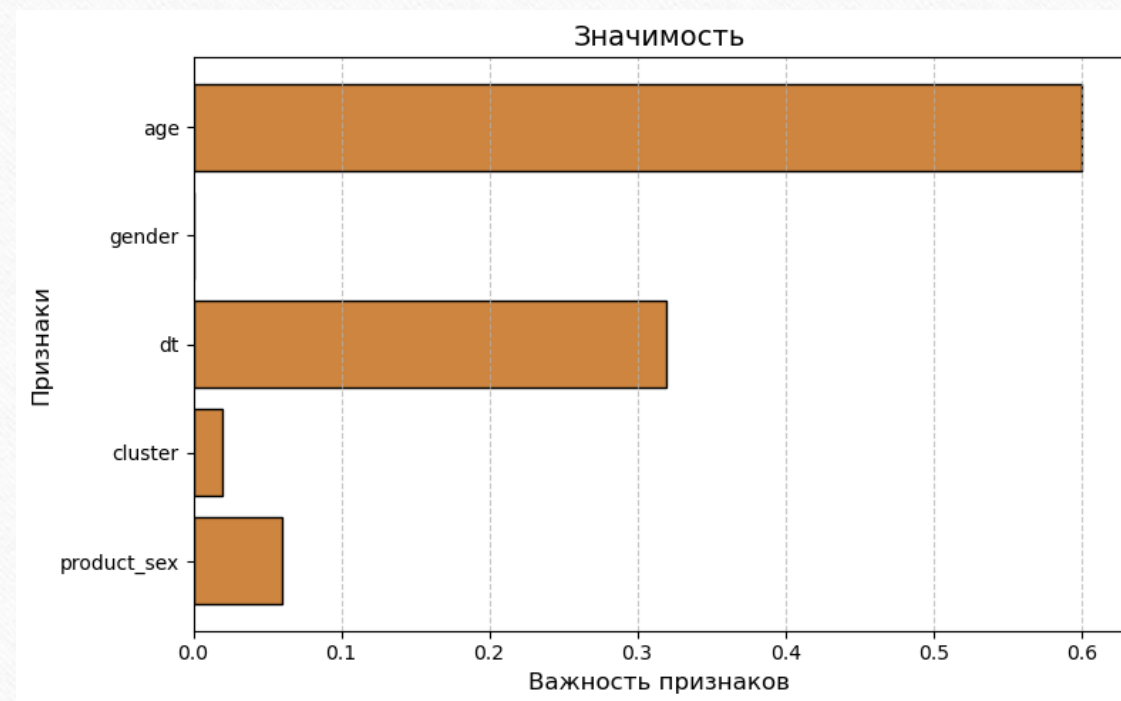
## Оценка точности модели:

**+** практически 1 на тестовых данных - отличная предсказательную способность;

**-** нулевая поддержка для кластера 0 - подозрение о наличии дисбаланса кластеров или недостаточной представленности кластера в данных.

## Кросс-валидация:

стратифицированная кросс-валидация показывает почти идеальную точность (все значения выше 0.9999) - модель стабильна и надежна на разных подвыборках данных.



# Рекомендации по улучшению модели:

---

- **Дисбаланс классов:**

несмотря на высокую точность, стоит уделить внимание кластеру 0, так как он либо отсутствует в тестовых данных, либо модель его игнорирует. Возможно, потребуется увеличение выборки для кластера 0 или использование методов, учитывающих дисбаланс кластеров.

- **Оптимизация признаков:**

можно рассмотреть исключение менее значимых признаков, таких как пол ('gender'), для упрощения модели.