

# Аналитический отчёт к итоговому заданию по модулю “Метрики, гипотезы, точки роста” AIC-32

## Описание данных<sup>1</sup>

Исходные данные содержат 54214 записей о транзакциях интернет-магазина Великобритании. Имеется 8 столбцов данных со след информацией:

- InvoiceNo - номер счёта/транзакции
- StockCode - код акции
- Description - описание товара
- Quantity - количество товара
- InvoiceDate - дата и время счёта/транзакции
- UnitPrice - цена за штуку
- CustomerID - идентификатор пользователя/клиента
- Country - Страна

Столбцы CustomerID и Description содержат пропущенные значения: 13571 и 134 соответственно.

Столбец Description помимо пропусков содержит неполные данные, например, с символами “?” и “??”.

Тип данных для столбцов Quantity и CustomerID - числовой, для остальных - object. Столбец Quantity содержит как положительные, так и отрицательные значения.

## Описание примененного подхода к анализу данных

- Тип данных в столбце UnitPrice был приведён к числовому типу float64.
- Данные из столбца InvoiceDate были разложены на 2 дополнительных столбца InvoiceDateWithoutTime (Дата счёта/транзакции без учёта времени) и InvoiceTime (время счёта/транзакции), и столбец InvoiceDateWithoutTime был приведён к типу данных datetime64.
- Пропущенные значения в столбцах CustomerID и Description. Данные значения для анализа заменили на 0.
- Столбец Quantity с отрицательными значениями и соответствующие значения в столбце InvoiceNo с префиксом “C” интерпретировались как возврат товара.
- Дополнительно был рассчитан столбец “Quantity\*UnitPrice” для вычисления стоимости товаров. Рассчитывается как количество товара умноженное на цену за штуку.
- В данных с пропусками и аномалиями в столбце 'Description' имеют нулевую 0.0 цену и стоимость. Предполагаем, что данные транзакции не были проведены. Поэтому разобьём данный датасет на 2: с пропусками в описании ('Description' - df\_filter\_anomal\_description) и без пропусков (df). Для дальнейшего анализа, в т.ч. выручки, будем использовать датасет без пропусков в описании.

---

<sup>1</sup> Более полное описание данных и промежуточные выводы представлены в файле Google Colab:

[https://colab.research.google.com/drive/1sV7sVT58Bu7\\_KsL0z82UmML0KEkdz36d?usp=sharing](https://colab.research.google.com/drive/1sV7sVT58Bu7_KsL0z82UmML0KEkdz36d?usp=sharing)

- Анализ данных осуществлялся при помощи Google Таблиц, Google DataStudio и Google Colaboratory (Python: Pandas, NumPy, Matplotlib)
- Для полноценного анализа бизнеса имеющихся данных недостаточно, однако, возможно привлечение дополнительных источников данных по маркетингу и т.д.

## Описание бизнес-заказчика, его гипотетических болей и целей развития бизнеса

Теоретически, в роли бизнес-заказчика может выступать владелец данного интернет-магазина.

Варианты бизнес-задач:

- На решение проблем:
  - по уменьшению количества возвратов
  - по уменьшению оттока покупателей
  - по уменьшению количества “брошенных корзин”
- На развитие бизнеса:
  - увеличение количества продаж
  - увеличение выручки
  - увеличение среднего чека
  - увеличение количества товаров в чеке
  - расширение географии заказов
  - сокращение времени между повторными заказами

## Описание стейкхолдеров отчёта с обоснованием выбранных метрик

- Product-менеджер: для понимания того, как пользователи взаимодействуют с интернет-магазином
  - MAU/DAU (Количество пользователей в месяц/день)
  - Churn и Customer Retention Rate (Отток и коэффициент удержания клиентов)
  - Conversion Rate (Конверсия в покупки: какая доля посетителей сайта совершила покупку)
  - Unsubscribe (Отписки от рассылки)
  - Referral (Отслеживание реферальной программы)
- Маркетолог: для оценки эффективности привлечения аудитории
  - CPC (Стоимость клика)
  - CTR (Соотношение кликов и показов, а также их качество)
  - CPA (Стоимость за целевое действие пользователя)
  - ROAS (Прибыль с учетом рекламных расходов)
  - CPL (Цена за лид)
  - CPA (Стоимость за действие, или сколько стоило привлечение 1 пользователя, совершившего целевое действие на сайте)

- CPO (Стоимость за заказ, или стоимость подтвержденного заказа, который совершил пользователь на сайте)
- Владелец бизнеса: для оценки эффективности бизнеса
  - ARPU (Средняя выручка с пользователя). Дополнительно, можно рассматривать ARPU как  $ARPPU$  (Средняя выручка с платящего пользователя) \* Conversion Rate.
  - LTV (Выручка или доход на одного пользователя за всё время взаимодействия с продуктом)
  - CPS (Сколько компании стоит привлечение 1 покупателя)
  - ROI (Прибыль с учетом маркетинговых расходов)
  - EBIT (Прибыль до вычета процентов и налогов)
  - Частота покупок
  - Количество товаров в чеке
  - Transactions by user (Количество транзакций у пользователя)
  - Средний чек

## Рекомендации по модели данных бизнес-заказчика

Для увеличения прибыли интернет-магазина Великобритании можно воспользоваться следующими метриками, согласно фреймворку AARRR:

### Привлечение (Acquisition)

Тематические каналы в популярных социальных сетях:

- Instagram
- Facebook
- Twitter
- YouTube

Сторонние онлайн-каналы:

- реклама в поисковых системах: Google, Yahoo и др.

Необходимо анализировать окупаемость затрат, вложенных в поддержание имеющихся каналов, с помощью следующих метрик: CPC (переходы на сайт), CTR (соотношение кликов и показов, а также их качество), ROI (прибыль с учетом маркетинговых расходов), ROAS (прибыль с учетом рекламных расходов).

### Активация (Activation)

Дополнительной ценностью интернет-магазина может быть возможность для пользователя перейти на сайт за подробным описанием товара, дополнительной информацией о производителе и процессе производства товара, лайфхаками по его использованию. Затем потенциальный покупатель может перейти к разделам сайта с описанием системы скидок и условиями доставки. При этом полезно учитывать частоту такой активности в месяц (MAU).

Аргументом для активации на сайте является накапливаемая скидка, которая действует при авторизации с помощью аккаунта в Facebook, Google, Yahoo или через электронную почту. Таким образом пользователь становится потенциальным клиентом (CPL). При авторизации с помощью почты необходимо перейти на сайт по ссылке в

письме. Можно посчитать, сколько человек в итоге перешли по данному типу ссылок с помощью метрики CTR или Conversion Rate.

После авторизации на сайте, пользователь автоматически подписывается на общую рассылку издательства. Можно измерить, какое количество пользователей отписывается от нее, с помощью метрики Unsubscribe.

Авторизованный пользователь может добавлять товары в WishList (список желаний) и в корзину, за эти показатели отвечает метрика CPO. Если потенциальный клиент делится найденными товарами в социальных сетях, заполняет форму обратной связи на сайте, то стоимость активных действий можно проанализировать с помощью метрики CPA.

## Удержание (Retention)

Удержание пользователей происходит с помощью тематических рассылок с полезной информацией, акциями. Переходы по рекламным ссылкам можно измерить метрикой CTR.

При таком активном использовании канала почты необходимо измерять количество пользователей, которые сами подписались на рассылки (SignUp) или отписались от них (Churn).

Для анализа возвращающихся пользователей можно использовать Retention и WAU, DAU (в неделю и день).

Метрикой ROI можно проверить гипотезы, оправдывающие затраты на маркетинг и на поддержание удобства сайта.

## Доход (Revenue)

После совершенной пользователем покупки можно проанализировать совокупный доход на каждого пользователя (LTV), средний ежемесячный доход (ARPU), прибыль до вычета процентов и налогов (EBIT).

Также, для анализа имеющихся совершённых сделок, интересно посмотреть на метрики отражающие частоту покупок, количество товаров в чеке, средний чек и количество транзакций у пользователя (Transactions by user).

На данном этапе также можно проанализировать соотношение доходов с затратами на рекламу на высоких уровнях воронки - CPS, ROI.

## Рекомендация (Referral)

Пользователь может поделиться найденными товарами или статьями на сайте в социальных сетях: Facebook, Twitter, Instagram. В случае личных рекомендаций, клиент может приобрести в подарок сертификат на покупку. Стоимость перечисленных действий можно оценить показателем CPA.

За реферальной программой в интернет-магазине можно проследить с помощью метрики Referral.

## Выводы и рекомендации для бизнес-заказчика

### Выводы:

- Основной сегмент покупателей и заказов приходится на Великобританию (более 80% от общего объема).
- 98% (13296) неидентифицированных пользователей приходится на United Kingdom.
- Итоговая стоимость всех транзакций (1000095 у.е) имеет положительное значение, следовательно, возвратов по заказам меньше, чем самих заказов.
- Данные по суммарной стоимости заказов по странам показывают, что наибольший доход приносят заказы из United Kingdom (более 80%). Топ-5 стран по выручке: United Kingdom, EIRE, Netherlands, Germany, France
- Топ-5 стран по количеству пользователей: United Kingdom, Germany, France, Spain, Belgium
- Подавляющее количество транзакций - из Великобритании. Также, несмотря на то, что Ирландия имеет небольшое количество пользователей (4), у неё высокий процент транзакций. Поэтому можно предположить, что в Ирландии высокий уровень лояльности клиентов к данному интернет-магазину. К тому же, самый активный пользователь также находится в Ирландии, и совершил уже 179 транзакций на протяжении года (с декабря 2019 по декабрь 2020).
- В USA количество возвращённого товара больше, чем купленного. Однако Выручка по данному региону имеет положительное значение. Возможно, сказывается неполнота данных, и часть товара была приобретена либо в другой период (не отражённый в данных), либо другим способом (оффлайн).
- У интернет-магазина есть постоянные клиенты (более 60% от общего числа пользователей), которые совершали транзакции более 1 раза
- Наибольшее количество транзакций приходится на конец 2020 года
- Самый дорогой заказ также приходится на конец 2020 года.
- Помимо проведённых транзакций на возврат, также имеются данные по транзакциям с нулевой стоимостью возврата при уменьшении количества товара. Возможно, в данном случае речь идёт о замене товара без возврата денег.
- Сумма возвратов составляет около 5% от суммы всех заказов.

### Рекомендации:

1. продолжить сбор данных по продажам для возможности в последующем определения сезонности в продажах
2. в связи с большим количеством неидентифицированных пользователей из Великобритании следует подумать над иными способами их идентификации
3. в связи с большим количеством возвратов из USA следует провести анализ структуры продаж внутри данного сегмента для выявления и устранения причин подобных возвратов

4. для расширения продаж в другие страны, помимо Великобритании, следует сосредоточиться на странах из Топ-5 по выручке и Топ-5 по количеству пользователей
5. обогатить данные по маркетингу, для более успешного привлечения аудитории и повышения конверсии в продажи