**Описание контекста кейса для маркетплейса “Малинка”**

Маркетплейс «Малинка» объединяет тысячи поставщиков и широкий ассортимент товаров, включая электронику, одежду, товары для дома и продукты питания. Пользователи платформы разделены на сегменты по активности и лояльности, а все их действия – просмотры, добавления в корзину, покупки и возвраты – фиксируются для анализа поведения.

Продажи и транзакции связаны с конкретными товарами, поставщиками и пользователями, а рекламные кампании отслеживаются по доходу, расходам, кликам и показам. Складские остатки и тикеты службы поддержки позволяют контролировать наличие товаров и качество обслуживания клиентов. Дополнительно, маркетплейс собирает данные о каналах трафика и устройствах пользователей, что позволяет оценивать эффективность рекламы и оптимизировать привлечение аудитории.

Данные представлены за 2025 год.

**Описание набора данных маркетплейса “Малинка”**

Набор данных состоит из 10 таблиц, охватывающих продажи, пользователей, товары, поставщиков, рекламу, трафик, возвраты, склад и поддержку клиентов.

1. **suppliers.csv – Справочник поставщиков**

**Назначение:** хранение информации о поставщиках.

**Поля:**

* supplier\_id – уникальный идентификатор поставщика
* supplier\_name – название поставщика
* region – регион поставщика
* rating – рейтинг поставщика

1. **products.csv – Справочник товаров**

**Назначение:** описание товаров маркетплейса.

**Поля:**

* product\_id – уникальный идентификатор товара
* product\_name – наименование товара
* category – категория товара
* price – цена
* supplier\_id – идентификатор поставщика (связь с suppliers)

1. **user\_segments.csv – Сегментация пользователей**

**Назначение:** информация о пользователях и их сегментах.

**Поля:**

* customer\_id – уникальный идентификатор пользователя
* segment – сегмент пользователя полученный от ML модели (new – новый пользователь, returning – повторно активный, loyal – лояльный, churn\_risk – склонный к уходу, high\_spender – крупный покупатель, discount\_hunter – ищущий скидки).
* region – регион пользователя
* registration\_date – дата регистрации

1. **sales.csv – Продажи и транзакции**

**Назначение:** информация о покупках пользователей.

**Поля:**

* transaction\_id – уникальный идентификатор транзакции
* customer\_id – идентификатор пользователя
* product\_id – идентификатор товара (связь с products)
* quantity – количество товара в транзакции
* payment\_method – способ оплаты
* transaction\_date – дата и время транзакции

1. **events.csv – События пользователей**

**Назначение:** действия пользователей на платформе.

**Поля:**

* event\_id – уникальный идентификатор события
* customer\_id – идентификатор пользователя (связь с user\_segments)
* event\_type – тип события (view – просмотр товара, add\_to\_cart – добавление в корзину, purchase – покупка, search – поиск, wishlist – добавление в избранное, click – клик по товару)
* event\_timestamp – дата и время события

1. **ad\_revenue.csv – Рекламный доход и расходы**

**Назначение:** информация о рекламных кампаниях и связанных товарах.

**Поля:**

* ad\_id – уникальный идентификатор кампании
* campaign\_name – название кампании
* product\_id – идентификатор товара (связь с products)
* spend – расходы на рекламу
* revenue – доход с кампании
* impressions – количество показов
* clicks – количество кликов
* date – дата кампании

1. **returns.csv – Возвраты**

**Назначение:** информация о возвратах товаров.

**Поля:**

* return\_id – уникальный идентификатор возврата
* transaction\_id – идентификатор транзакции (связь с sales)
* customer\_id – идентификатор пользователя (совпадает с sales)
* product\_id – идентификатор товара (совпадает с sales)
* reason – причина возврата

1. **traffic.csv – Данные по каналам трафика**

**Назначение:** хранит информацию о посещениях пользователей маркетплейса, включая канал привлечения, устройство и регион.

**Поля:**

* traffic\_id – уникальный идентификатор сессии
* customer\_id – идентификатор пользователя (связь с user\_segments)
* channel – канал трафика (malinka\_ads – собственная реклама маркетплейса, yandex – переход через Яндекс, mail.ru\_search – переход через поиск Mail.ru, vk – переход через социальную сеть VK, organic – органический трафик)
* session\_start – время начала сессии
* device – устройство, с которого был визит (desktop, mobile, tablet)

1. **inventory.csv – Остатки на складе**

**Назначение:** информация о текущих запасах товаров.

**Поля:**

* product\_id – идентификатор товара (связь с products)
* warehouse\_id – идентификатор склада
* stock\_quantity – количество на складе
* last\_updated – дата последнего обновления

1. **customer\_support.csv – Поддержка клиентов**

**Назначение:** информация о тикетах поддержки.

**Поля:**

* ticket\_id – уникальный идентификатор тикета (Тикет – это отдельная запись или запрос в системе поддержки клиентов, который фиксирует проблему или вопрос пользователя. Каждому тикету присваивается уникальный идентификатор (ticket\_id), чтобы отслеживать процесс решения, назначать ответственных и фиксировать статус (решён/не решён))
* customer\_id – идентификатор пользователя (связь с user\_segments
* issue\_type – тип проблемы (refund – возврат или возврат денег, delivery\_delay – задержка доставки, product\_question – вопрос о товаре, account – проблемы с аккаунтом)
* resolution\_time\_minutes – время решения тикета
* resolved – признак решения тикета (True/False)
* support\_date – дата тикета

**Описанные взаимосвязи таблиц**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Таблица 1** | **Колонка** | **Таблица 2** | **Колонка** | **Тип связи** |
| sales | product\_id | products | product\_id | Многие к одному |
| sales | customer\_id | user\_segments | customer\_id | Многие к одному |
| returns | transaction\_id | sales | transaction\_id | Многие к одному |
| returns | product\_id | sales | product\_id | Многие к одному |
| returns | customer\_id | sales | customer\_id | Многие к одному |
| events | customer\_id | user\_segments | customer\_id | Многие к одному |
| ad\_revenue | product\_id | products | product\_id | Многие к одному |
| traffic | customer\_id | user\_segments | customer\_id | Многие к одному |
| customer\_support | customer\_id | user\_segments | customer\_id | Многие к одному |
| inventory | product\_id | products | product\_id | Многие к одному |
| products | supplier\_id | suppliers | supplier\_id | Многие к одному |