

Аналитика онлайн-магазина (e-commerce)

Структура проекта

Описание проекта

Название: Аналитика данных онлайн-магазина

Цель: Построить модель базы данных, наполнить ее данными, и с помощью SQL-отчетов ответить на ключевые бизнес-вопросы.

Роль: Системный аналитик + SQL-аналитик.

Инструменты: PostgreSQL, DBeaver.

Модель базы данных

1. Таблица customers (Клиенты)

Назначение: Хранит информацию о клиентах.

Поле	Тип	Описание
customer_id	INT (PK)	Уникальный идентификатор клиента
name	VARCHAR	Имя клиента
email	VARCHAR	Email клиента
registration_date	INT	Дата регистрации клиента

2. Таблица categories (Категории товаров)

Назначение: Справочник категорий для товаров.

Поле	Тип	Описание
category_id	INT (PK)	Уникальный идентификатор категории
name	VARCHAR	Название категории

3. Таблица products (Товары)

Назначение: Информация о товарах, доступных к заказу.

Поле	Тип	Описание
product_id	INT (PK)	Уникальный идентификатор товара
name	VARCHAR	Название товара
price	INT	Цена товара
category_id	INT (FK)	Категория товара (ссылается на categories)

4. Таблица orders (Заказы)

Назначение: Информация о заказах, оформленных клиентами.

Поле	Тип	Описание
order_id	INT (PK)	Уникальный идентификатор заказа
customer_id	INT (FK)	Кто сделал заказ (ссылается на customers)
order_date	INT	Дата оформления заказа
status	VARCHAR	Статус заказа (например, "выполнен", "в обработке")

5. Таблица order_items (Позиции заказа)

Назначение: Конкретные товары, входящие в заказ.

Поле	Тип	Описание
order_item_id	INT (PK)	Уникальный идентификатор строки заказа
order_id	INT (FK)	Ссылается на заказ
product_id	INT (FK)	Ссылается на товар
quantity	INT	Количество единиц товара

Поле	Тип	Описание
price	INT	Цена на момент заказа

6. Таблица returns (Возвраты)

Назначение: Хранит информацию о возвращенных товарах.

Поле	Тип	Описание
return_id	INT (PK)	Уникальный ID возврата
order_item_id	INT (FK)	Ссылка на строку заказа
return_date	INT	Дата возврата
reason	VARCHAR	Причина возврата

7. Таблица payments (Платежи)

Назначение: Платежи по заказам.

Поле	Тип	Описание
payment_id	INT (PK)	Уникальный идентификатор платежа
order_id	INT (FK)	Заказ, за который был платёж
payment_date	INT	Дата платежа
amount	INT	Сумма платежа
payment_method	VARCHAR	Метод оплаты (например, "карта")

Связи между таблицами

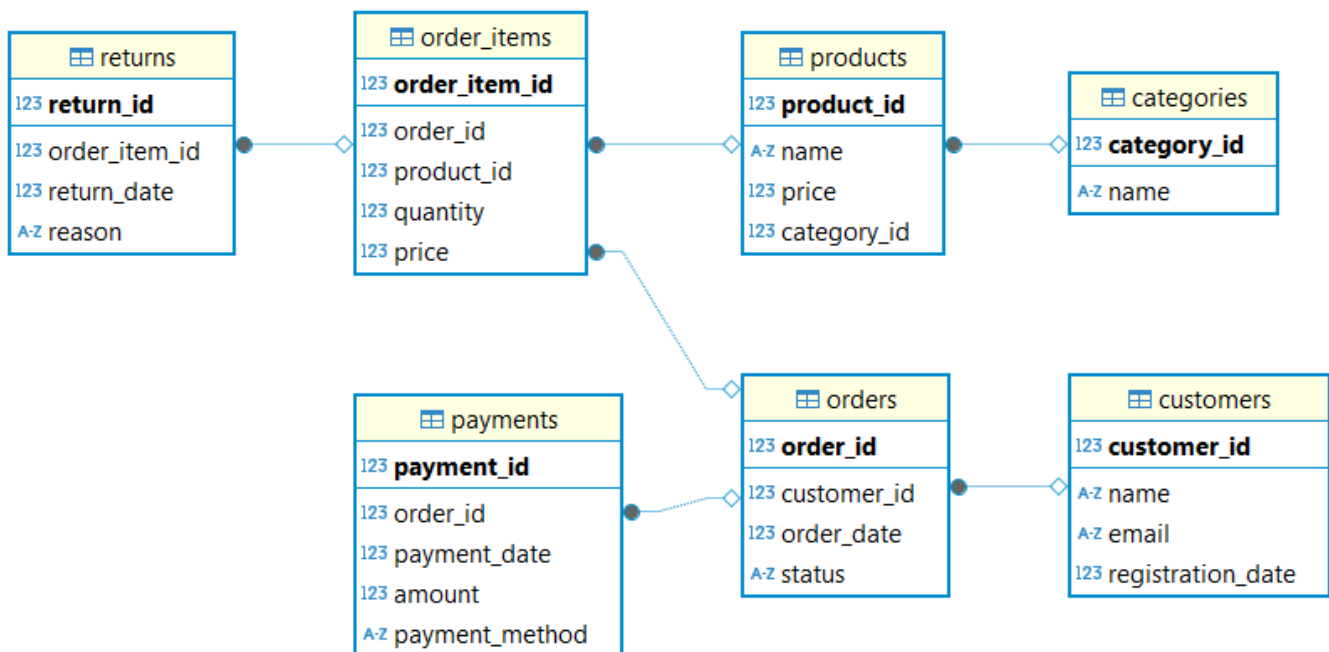
База данных построена на взаимосвязанных сущностях. Ниже представлены ключевые связи между таблицами:

- **orders.customer_id → customers.customer_id**

Каждый заказ связан с конкретным клиентом. Один клиент может сделать много

заказов (связь *один ко многим*).

- **order_items.order_id** → **orders.order_id**
Один заказ может включать в себя несколько товарных позиций. Это связь *один ко многим* между заказами и строками заказа.
- **order_items.product_id** → **products.product_id**
Каждая строка заказа содержит конкретный товар из таблицы товаров. Один товар может встречаться в разных заказах (*один ко многим*).
- **products.category_id** → **categories.category_id**
Каждый товар принадлежит определённой категории. Это *один ко многим* от категорий к товарам.
- **returns.order_item_id** → **order_items.order_item_id**
Возврат осуществляется по конкретной позиции заказа. Одна строка заказа может быть возвращена один раз.
- **payments.order_id** → **orders.order_id**
Платежи относятся к конкретному заказу. Один заказ — один или несколько платежей (в зависимости от бизнес-логики).



Тестовые данные, сгенерированные с помощью ChatGPT.

Всего сгенерировано:

- 10 записей для categories
- 50 для customers
- 100 для products
- 80 для orders
- 200 для order_items
- 20 для returns
- 80 для payments .

SQL запросы для анализа информации

Топ-5 продуктов по количеству проданных единиц

```
SELECT
    p.name AS product_name,
    SUM(oi.quantity) AS total_quantity
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
GROUP BY p.name
ORDER BY total_quantity DESC
LIMIT 5;
```

Клиенты с более чем 3 заказами

```
SELECT
    COUNT(*) AS customers_with_more_than_3_orders
FROM (
    SELECT customer_id
    FROM orders
    GROUP BY customer_id
    HAVING COUNT(order_id) > 3
) AS sub;
```

Средняя стоимость заказа по категориям

```
SELECT
    c.name AS category_name,
    AVG(oi.price * oi.quantity) AS average_check
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
```

```
JOIN categories c ON p.category_id = c.category_id
GROUP BY c.name;
```

Анализ причин возвратов

```
SELECT
    reason,
    COUNT(*) AS count
FROM returns
GROUP BY reason
ORDER BY count DESC;
```

Тренды месячного дохода

```
SELECT
    FLOOR(order_date / 100) AS month,
    SUM(oi.price * oi.quantity) AS revenue
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
GROUP BY month
ORDER BY month;
```

Распределение методов оплаты

```
SELECT
    payment_method,
    COUNT(*) AS count
FROM payments
GROUP BY payment_method
ORDER BY count DESC;
```

Заказы с возвратами

```
SELECT DISTINCT o.order_id
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
JOIN returns r ON oi.order_item_id = r.order_item_id;
```

Анализ доходов и возвратов

```
SELECT
    SUM(oi.price * oi.quantity) AS total_revenue,
    COALESCE(SUM(oi.price), 0) AS total_returns,
    SUM(oi.price * oi.quantity) - COALESCE(SUM(oi.price), 0) AS net_profit
FROM order_items oi
LEFT JOIN returns r ON oi.order_item_id = r.order_item_id
WHERE r.return_id IS NOT NULL;
```

Анализ данных электронной коммерции

Визуализация данных

E-commerce Analytics Dashboard

Customers with 3+ Orders



3

Loyal customers who have placed more than 3 orders

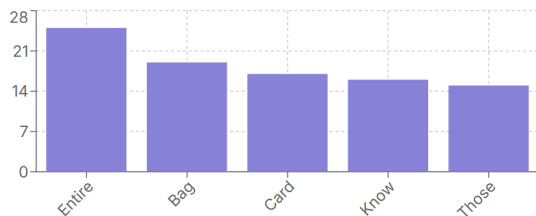
Store Profit

\$

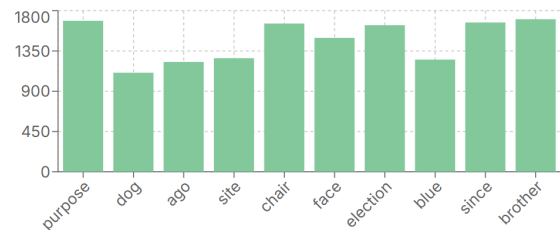
\$13 719

↗ 61.9% profit margin

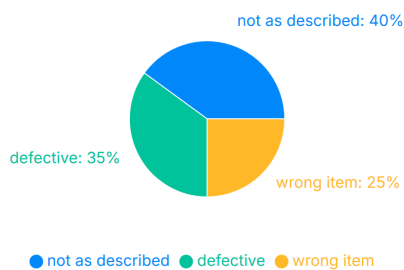
Top 5 Products by Sales



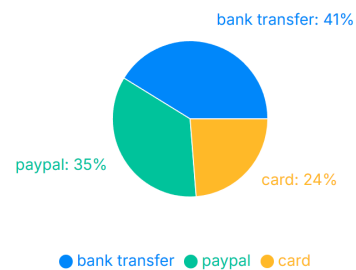
Average Check by Categories



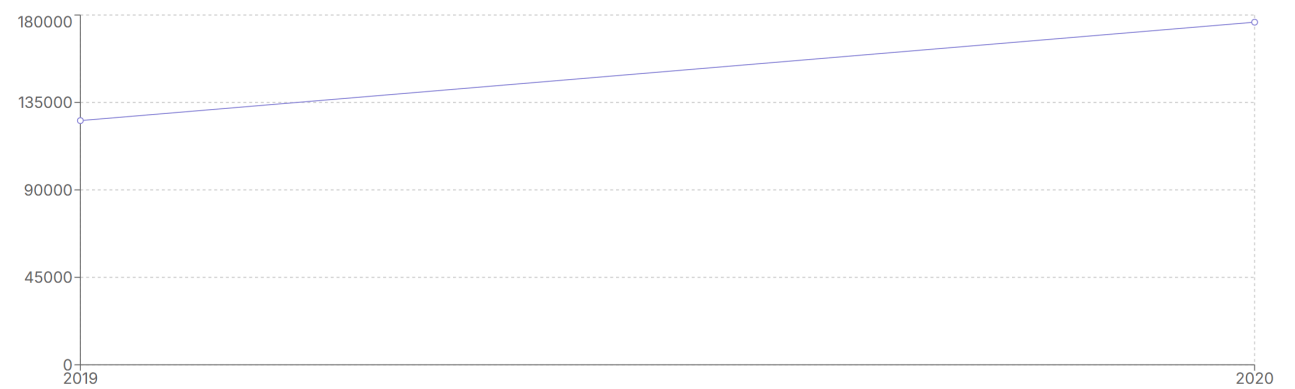
Returns by Reason



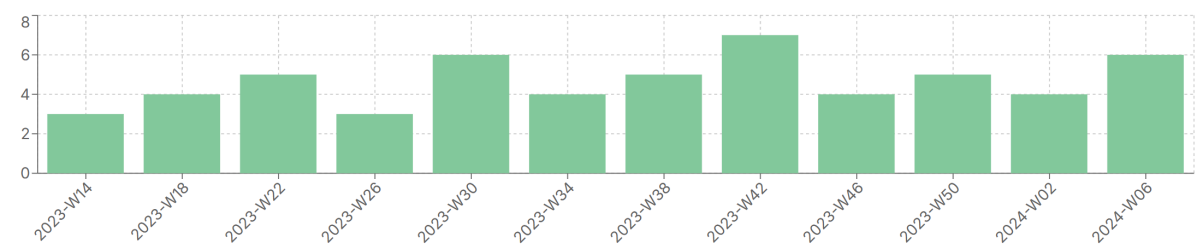
Payment Methods Distribution



Monthly Sales Dynamics



New Customers by Week



Orders with Returned Items

Order ID	Status
#54	Has returns
#70	Has returns
#63	Has returns
#31	Has returns
#59	Has returns
#62	Has returns
#42	Has returns
#41	Has returns
#53	Has returns
#61	Has returns
#77	Has returns

Выводы из анализа данных

1. Общий доход, возвраты и чистая прибыль

- Общий доход составил 22,150, что является хорошим показателем, но возвраты (8,431) значительно снижают прибыль. Чистая прибыль (13,719) отражает реальный доход после учета возвратов, что помогает понять, какой процент от общего дохода был потерян из-за возвратов.

2. Средняя стоимость заказа по категориям

- Стоимость заказов варьируется по категориям. Например, категория "Brother" имеет наивысшую среднюю стоимость заказа (1,704.25), что может указывать на более дорогие товары в этой категории. Напротив, категория "Dog" имеет более низкую среднюю стоимость (1,106.56), что может быть связано с дешевыми товарами или меньшими объемами покупок.

3. Причины возвратов

- Возвраты из-за того, что товары "не такие, как на описании", являются самой частой причиной (8 случаев). Это может свидетельствовать о необходимости улучшения качества описаний товаров или фотоматериалов, чтобы соответствовать ожиданиям покупателей. Следующие причины возвратов — дефекты товара (7 случаев) и неправильный товар (5 случаев), что требует дополнительного внимания к качеству продукции и логистике.

4. Распределение методов оплаты

- Метод оплаты через банковский перевод является самым популярным (33 случая), что может указывать на предпочтение клиентов использовать более традиционные и безопасные способы оплаты. PayPal и карта тоже используются, но в меньшем объеме (28 и 19 соответственно). Это может помочь для оптимизации платежных систем на сайте, чтобы предложить наиболее популярные способы оплаты.

5. Топ-5 продуктов по количеству проданных единиц

- Продукты "Entire", "Bag", "Card", "Know" и "Those" — это самые популярные товары. Возможно, стоит сконцентрироваться на этих продуктах для дальнейших маркетинговых стратегий и увеличения их доступности.

6. Месячные доходы

- Видно, что в месяцах 19 и 20 доходы значительно отличаются (125,670 и 176,319). Это может свидетельствовать о сезонных колебаниях в продажах или успешных акциях в определенные месяцы. Анализ таких трендов помогает адаптировать стратегии продаж.

7. Заказы с возвратами

- Список заказов с возвратами позволяет провести анализ проблемных позиций и товарных категорий, которые имеют высокие показатели возврата. Это поможет выявить слабые места в обслуживании и продукции, а также улучшить лояльность клиентов.

8. Количество клиентов с более чем 3 заказами

- Три клиента сделали более 3 заказов. Это подчеркивает важность привлечения и удержания лояльных клиентов, на которых можно строить долгосрочные стратегии, такие как программы лояльности и персонализированные предложения.

Рекомендации

1. Улучшение качества описаний товаров и фото

- Для уменьшения возвратов по причине "не такие, как описано" следует улучшить качество контента на сайте. Это может включать более точные описания, фотографии высокого качества и видеоролики.

2. Фокус на популярных категориях

- Стоит обратить внимание на категории с высокой средней стоимостью заказа, такие как "Brother", и разработать специальные предложения для клиентов, заинтересованных в дорогих товарах.

3. Оптимизация системы возвратов

- Стоит анализировать товары с высокой частотой возвратов и провести работу с производителями или поставщиками, чтобы снизить количество дефектных товаров.

4. Развитие программы лояльности

- Учитывая, что у нас есть лояльные клиенты, которые делают более 3 заказов, можно предложить им персонализированные скидки, бонусы или программы лояльности для повышения повторных покупок.

5. Анализ сезонности

- С учетом разницы в доходах между месяцами можно внедрить акции, скидки и другие маркетинговые мероприятия в периоды низких продаж для стимулирования покупок.

6. Оптимизация методов оплаты

- Можно предложить дополнительные способы оплаты, которые не так популярны в данный момент, чтобы увеличить конверсию, например, добавить поддержку криптовалют или других современных методов.