

MISM 6201 PROJECT

A Typical E-Commerce Database

Changjiang Song, Chuwen Feng, Hongyi Hu and Zijing Bai

12/11/2022

Introduction

In the era where everything is associated with the internet, online shopping is no longer the latest concept. “In 2021, the number of global digital buyers was 2.14 billion. As of October 2022, the world population comprises over 7.9 billion people. That means at least 27% of the worldwide population shop online to fulfill their needs.” (Maisha & Jordana, 2022) This refers to the fact that out of four people, one of them describes themselves as an online shopper.

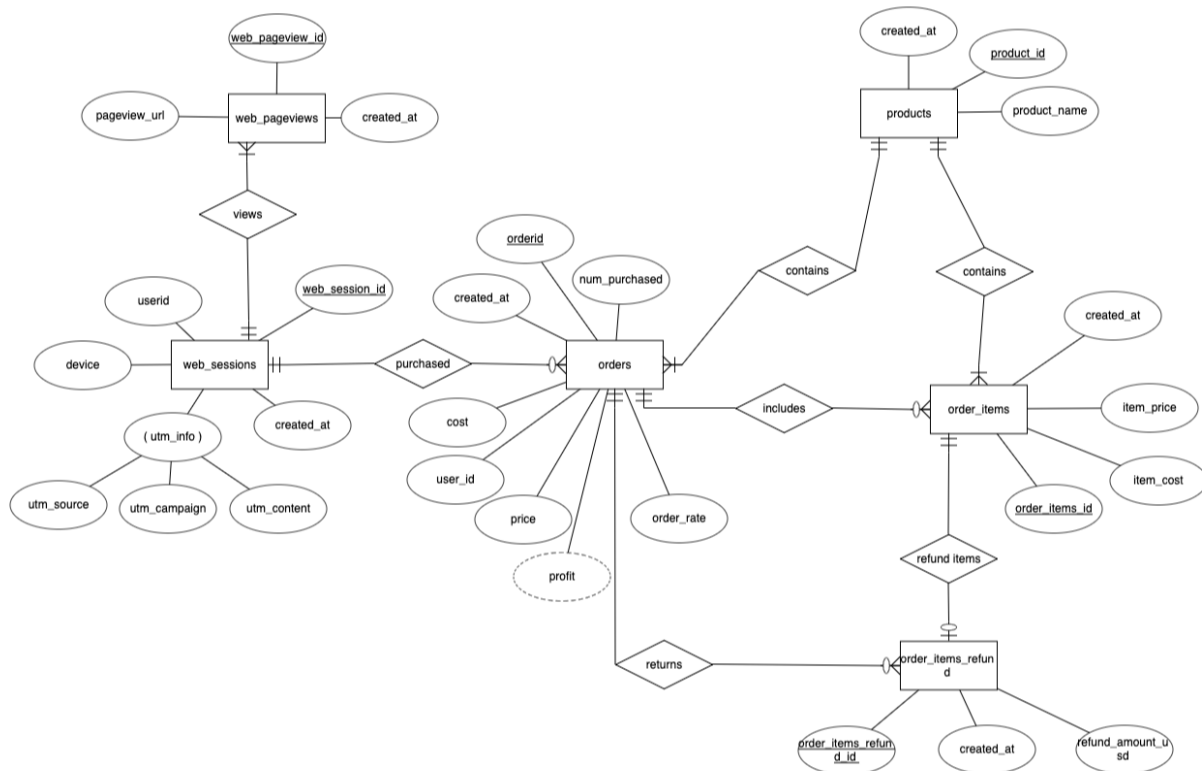
On the other hand, data, as a concept, has been immersed in our life as well. When A traditional field got transformed into the form of digital, the implied features of such a field would be embodied as certain types of data which, in general, are sizing enormously. So then, the measurement degree of key performance indicators in such a field has evolved, requiring a better platform for all stakeholders to analyze and to be able to utilize. The same statement was made by Prince Bharadwaj and other associated authors in 2022: “One of the biggest benefits of using a database for e-commerce is structuring vast amounts of shop data.”

As the conclusion of the introduction section, the need of handling the massive amount of data along with its form of being digital ensured the feasibility of the database management system. In this paper, we have introduced a classic, typical and default version of the database

management system structure, as well as some possible business questions along with the practical queries to each question.

Database Management System Structure

The ER Diagram



- **Web_session Relation**

- It keeps track of web session ID (unique), User ID, device, timestamp of its creation, utm information (composed by source, campaign and content).
- In each web_session, there are one or more page views.
- In each web_session, there are zero or more orders.

- **Web_view Relation**

- It keeps track of pageview ID (unique), pageview URL and timestamp of its creation.
- For each web view, it has to be viewed through one web session.

- **Order Relation**

- It keeps track of order ID (unique), number of purchases, price, user ID, cost, order rating and timestamp of its creation.
- Each order contains only one product.
- An order can be assigned as one or more order items.
- In each order, there are zero or more returned order items.
- For each order, it has to be made in one web session.

- **Products Relation**

- It keeps track of product ID (unique), product name and timestamp of its creation.
- Each product is assigned differently as an order item between orders.
- A product can be purchased in one order or many different orders.

- **Order_items Relation**

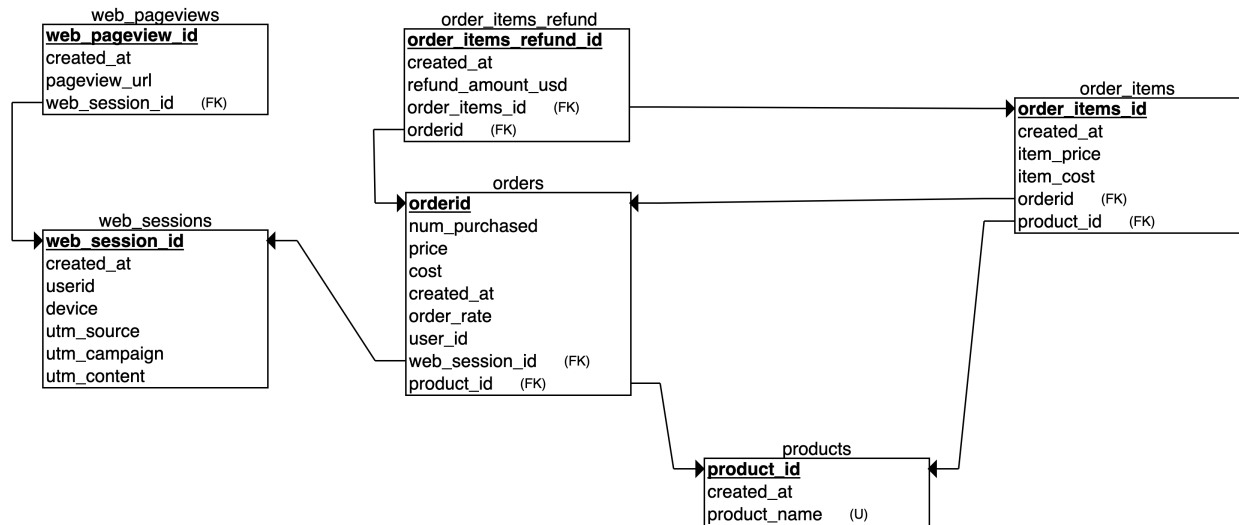
- It keeps track of order items ID (unique), item price, item cost and timestamp of its creation.
- Each order item itself is uniquely refers to a product.
- An order item is either refunded or not.
- The relationship between order and order items is exclusive between orders.

- **Order_items_refund Relation**

- It keeps track of order items refund ID (unique), refund amount in US dollars and timestamp of its creation.

- Each refunded item is an ordered item.
- Each refunded item is refunded in a certain order.

The Schema



This schema shows a clear vision of how tables are related. In terms of data inputting, we want to discuss it by the actual level of users' engagement which is also relevant to the hierarchical logic we had in the next section when we listed some queries for solving practical business questions.

- Informational table: Products
- Level one: A user who clicked the link will be recorded in web_sessions and web_pageviews.
- Level two: A user who has gone through to the last page and made purchases would have records left in orders, and order_items.
- Level three: An order returned by a user were labeled and recorded in order_items_refund.

Queries for Solving Practical Business Questions

The database designed can help grow the business and analyze performance along the way. The following questions are the most common questions that Marketing Director and Website Manager want to find out in an e-commerce company.

1. Which day in a week has the greatest number of orders?

The query above can help us understand which week has the greatest number of order so that as a Marketing director can learn the sale on the time trend of weekday. In this way, the Marketing strategies will be more targeted in the time dimension. The methodology of the query can also be applied in the multi-time dimension to observe which specific time range has the least and the greatest number of orders.

The query technique included:

The standard SQL query language such as select, group by, aggregation function such as 'count', Date transform.

Result of Query 1:

The greatest number of orders, five orders, occurred on Tuesday.

2. What is the net profit of each product in March?

Net profit is another important parameter that determines the financial health of your business. It shows whether the business can make more than what it spends. Net profit is able to help manager decide when and how to work towards expanding the e-commerce business and when to reduce the expenses.

The query technique included:

The standard SQL query language such as select, group by, aggregation function such as 'sum'.

Result of Query 2:

The net profit of each product in March is \$123, \$82, \$58.5 and \$63, respectively.

3. To find out the products that have been ordered most.

This query analyzes the performance of the business in perspective of products. Product sales performance is one of the indicators that every company most expect to learn.

The query technique included:

The standard SQL query language such as select, group by, aggregation function such as 'count', Window Function "dense_rank", and Subquery technique like CTE.

Result of Query 3:

Product 1 and 4 have been ordered the most, and each has been ordered 6 times.

4. What products are most likely to be returned?

This query also analyzes business performance from a product perspective. Knowing the return status of a product is very important. This helps companies locate the shortcomings of products with poor sales performance in order to improve targeted products and design more efficient marketing strategies in the future.

The query technique included:

The standard SQL query language such as select, group by, aggregation function such as 'count', and Inner Join.

Result of Query 4:

Product 1 has been returned the most, which is 3 times.

5. Which page of our online store has gone wrong? What was the conversion rate between pages? (Building Conversion Funnel)

Conversion funnel analysis is about understanding and optimizing each step of your user's experience on their journey toward purchasing your products.

There are three common use cases.

- Identifying the most common paths customers take before purchasing your products
- Identifying how many of your users continue to each next step in your conversion flow, and how many users abandon at each step
- Optimizing critical pain points where users are abandoning, so that you can convert more users and sell more products

When we perform conversion funnel analysis, we will look at each step in our conversion flow to see how many customers drop off and how many continue at each step.

- Step 1: select all pageviews for relevant sessions and identify each relevant pageview as the specific funnel step
- Step 2: Create the session - level conversion funnel view
- Step 3: Aggregate the data to assess funnel performance

The query technique included:

The standard SQL query language such as select, group by, aggregation function such as 'count', 'max', left Join, case, and Subquery technique like CTE.

Result of Query 5:

Clicked Number:

Count	To_product	To_mrfuzzy	To_cart	To_shipping	To_billing	To_thankyou
24	14	11	10	9	9	8

Clicked Rate:

To_product	To_mrfuzzy	To_cart	To_shipping	To_billing	To_thankyou
0.5833	0.7857	0.9091	0.9000	1.0000	0.8889

Conclusion

Even though e-commerce and data are widely accepted concepts, the actual usage of database management systems in the e-commerce industry is critical. Tools like conversion funnel are meant to seek the underlying logic of users' behaviors; It is very different from running a business physically in which more people skills are applied, the only clue in the e-commerce industry is the massive influx of data generated every moment. Thus, understanding the message behind the data and exploring the possibility of the database management system is the skill which all e-commerce practitioners need to have.

Reference

- Bharadwaj, P., Search for more papers by this author, Kumar, S., Raj, A., & Hudnurkar, M. (n.d.). *Chapter 15: Role of database management in e-commerce firms*. World Scientific. Retrieved December 12, 2022, from https://worldscientific.com/doi/10.1142/9789811245992_0015
- R., M. (2022, December 8). *25 ecommerce statistics and how to capitalize on them in 2022*. Hostinger Tutorials. Retrieved December 12, 2022, from <https://www.hostinger.com/tutorials/ecommerce-statistics>