**INTRODUCTION**

As a data analyst, one of the most interesting thing of my job is uncovering insights about customer behaviour using database analysis. I got an opportunity to work on a Case Study that involved analyzing multiple tables in a database to understand customer behaviour for a popular online retailer.

**TABLE CONSIST FOLLOWING TABLES:**

1. SALE\_PROJECT
2. PRODUCTS PROJECT \_
3. USER\_NAME\_PROJECT
4. USERS\_PROJECT
5. GOLD\_USER\_SIGNUP

Each table contains unique information about the customers and products sold.

1. **SALE\_PROJECT**: This table provides details about each Sale, including the date of the sale, the product sold, the price, and the user who made the purchase. This table help us to analyse about top selling product and most active user.
2. **PRODUCTS PROJECT:** The product table contained information about each product sold on the website, including the product ID, Product Name and price of a Product. . Using this table, we were able to identify the prices for each product and compare them to determine which products are more expensive and which are less expensive.
3. **USER\_NAME\_PROJECT:** The username table contains information about each user’s UserID and Username. We utilized this table to link with other tables and use their User ID to join with other tables. Specifically, the table includes the UserID and the corresponding Names of the users.
4. **USERS\_PROJECT:** The users table provided information about each customer who created an account on the website, including their name, Signup Date. By analyzing this table, we were able to identify trends in customer demographics.
5. **GOLD\_USER\_SIGNUP:** The goldusers\_signup table contained data about customers who signed up for the website’s premium membership program. We were able to use this table to identify the demographics of customers who were most likely to sign up for the program and which benefits of the program were the most appealing to customers.

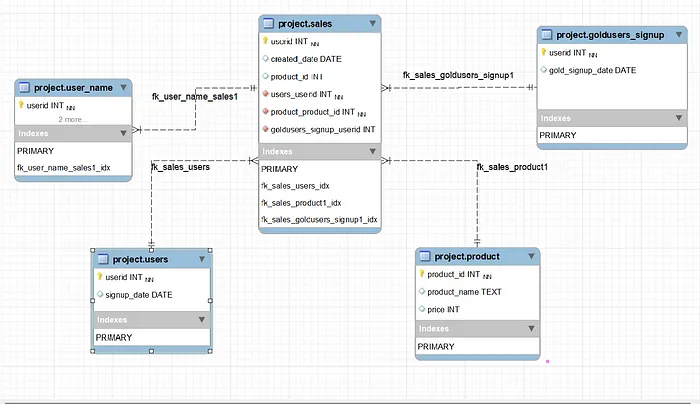


Fig. Entity Relationship Diagram (ERD)

Following are the tables:

Table-1: Sales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| USER\_ID | CREATED\_DATE | PRODUCT\_ID | QUANTITY\_SOLD | SALE\_PRICE |
| 1 | 2017-04-19 | 2 | 5 | 50000 |
| 3 | 2019-12-18 | 1 | 7 | 60000 |
| 2 | 2020-07-20 | 3 | 8 | 40000 |
| 1 | 2019-10-23 | 2 | 4 | 50000 |
| 1 | 2018-03-19 | 3 | 12 | 80000 |
| 3 | 2016-12-20 | 2 | 2 | 20000 |
| 1 | 2016-11-09 | 1 | 4 | 60000 |
| 1 | 2016-05-20 | 3 | 4 | 80000 |
| 2 | 2017-09-24 | 1 | 9 | 30000 |
| 1 | 2017-03-11 | 2 | 2 | 70000 |
| 1 | 2016-03-11 | 1 | 3 | 80000 |
| 3 | 2016-11-10 | 1 | 7 | 90000 |
| 3 | 2017-12-07 | 2 | 8 | 70000 |
| 3 | 2016-12-15 | 2 | 6 | 10000 |
| 2 | 2017-11-08 | 2 | 4 | 20000 |
| 2 | 2018-09-10 | 3 | 5 | 60000 |
| 4 | 2019-05-01 | 1 | 8 | 50000 |
| 5 | 2018-11-23 | 3 | 10 | 70000 |
| 6 | 2017-06-30 | 9 | 8 | 90000 |
| 7 | 2018-08-12 | 8 | 15 | 120000 |

Table-2:

Table-2: Product

|  |  |  |
| --- | --- | --- |
| PRODUCT\_ID | PRODUCT\_NAME | PRICE |
| 1 | HEADPHONES | 30000 |
| 2 | MICROWAVE OVEN | 40000 |
| 3 | PRINTER | 90000 |
| 4 | SMARTPHONE | 80000 |
| 5 | LAPTOP | 1000 |
| 6 | COFFEMAKER | 1500 |
| 7 | BLENDER | 2000 |
| 8 | FOOD PROCESSOR | 2500 |
| 9 | WATER PURIFIER | 50000 |
| 10 | VACUUM CLEANER | 5000 |
| 11 | TABLET | 45000 |
| 12 | COMPUTER | 7000 |
| 13 | AIR CONDITIONER | 2000 |
| 14 | GEYSER | 18000 |
| 15 | DIGITAL CAMERA | 25000 |
| 16 | WEB CAMERA | 15000 |
| 17 | CEILING FAN | 3000 |
| 18 | WASHING MACHINE | 2500 |
| 19 | DISH WASHER | 1999 |
| 20 | PROJECTOR | 2499 |

Table-3: User\_name

|  |  |
| --- | --- |
| ID | NAMES |
| 1 | Rohan |
| 2 | Dezy |
| 3 | Akila |
| 4 | Suman |
| 5 | Chander |
| 6 | Vipina |
| 7 | Ayushi |
| 8 | Payal |
| 9 | Shreyash |
| 10 | Sharad |
| 11 | Sakhi |
| 12 | Sara |
| 13 | Apurva |
| 14 | Jyotsna |
| 15 | Neha |
| 16 | Prajakta |
| 17 | Krushna |
| 18 | Pooja |
| 19 | Ritik |
| 20 | Himanshu |

Table-4: User\_project

|  |  |
| --- | --- |
| USER\_ID | SIGNUP\_DATE |
| 1 | 02/09/14 |
| 2 | 15/01/15 |
| 3 | 11/04/14 |
| 4 | 17/11/15 |
| 5 | 02/01/16 |
| 6 | 02/01/16 |
| 7 | 02/04/13 |
| 8 | 15/12/13 |
| 9 | 08/09/15 |
| 10 | 13/07/14 |
| 11 | 15/02/14 |
| 12 | 15/09/16 |
| 13 | 10/05/13 |
| 14 | 20/10/15 |
| 15 | 16/06/14 |
| 16 | 10/09/14 |
| 17 | 14/12/13 |
| 18 | 16/03/13 |
| 19 | 18/06/15 |
| 20 | 20/07/16 |

Table-5: GOLD\_USER\_SIGNUP

|  |  |  |  |
| --- | --- | --- | --- |
| USER\_ID | SIGNUP\_DATE | MEMBERSHIP\_TYPE | SALE\_PRICE |
| 1 | 21/02/22 | REGULAR | 400000 |
| 2 | 23/12/21 | GOLD | 500000 |
| 3 | 27/02/22 | GOLD | 700000 |
| 4 | 24/02/22 | REGULAR | 250000 |
| 5 | 03/01/24 | GOLD | 90000 |
| 6 | 04/01/24 | REGULAR | 79000 |
| 7 | 05/01/24 | GOLD | 300000 |
| 8 | 29/01/24 | GOLD | 800000 |
| 9 | 30/01/24 | REGULAR | 500000 |
| 10 | 31/01/24 | GOLD | 700000 |
| 11 | 01/02/24 | REGULAR | 500000 |
| 12 | 02/02/24 | GOLD | 690000 |
| 13 | 03/02/24 | REGULAR | 78000 |
| 14 | 04/02/24 | GOLD | 980000 |
| 15 | 31/05/24 | GOLD | 650000 |
| 16 | 01/06/24 | GOLD | 780000 |
| 17 | 02/06/24 | REGULAR | 78000 |
| 18 | 03/06/24 | GOLD | 850000 |
| 19 | 04/06/24 | GOLD | 950000 |
| 20 | 05/06/24 | REGULAR | 78900 |

**Q1. What is the total sales revenue generated by each product?**

To determine the total sales revenue generated by each product, we need to analyse the sales data in the database. By using SQL queries, we can aggregate the sales data by product and calculate the total revenue generated by each product.

QUERY: SELECT

Product\_ID,

SUM(Quantity\_Sold \* Sale\_Price) AS TotalRevenue

FROM

Sales\_PROJECT

GROUP BY

Product\_ID;

# Q2. Which 3 product has the highest sales revenue?

To determine the 3 products with the highest sales revenue, we need to analyze the sales data in the database. By using SQL queries, we can aggregate the sales data by product and sort them by the revenue generated to identify the top-selling products**.** The answer to this question cannot be determined based on the information provided in the original table. However, once we have access to the sales data, we can use SQL to calculate the total revenue generated by each product and sort them in descending order to identify the top 3 products with the highest sales revenue.

**QUERY:** SELECT

Product\_ID,

SUM(Quantity\_Sold \* Sale\_Price) AS TotalRevenue

FROM

Sales\_PROJECT

GROUP BY

Product\_ID

ORDER BY

TotalRevenue DESC

LIMIT 3;

|  |  |
| --- | --- |
| product\_id | total revenue |
| 3 | 2600000 |
| 1 | 2200000 |
| 8 | 1800000 |

# Q3. How many users have signed up for the service and has taken the gold membership?

Knowing the number of users who have taken the gold membership can provide valuable insights into the success of the loyalty program and the overall business performance. A high number indicates an effective program leading to increased customer retention, repeat purchases, and positive referrals. A low number suggests a need for program improvement, which can be done by offering better rewards or simplifying the redemption process. This information can help businesses identify opportunities to improve the loyalty program and ultimately drive growth and profitability.

QUERY: SELECT

COUNT(\*) AS Gold\_Membership\_Count

FROM

GOLD\_USER\_SIGNUP

WHERE

UPPER(Membership\_Type) = 'GOLD';

|  |
| --- |
| gold\_membership\_count |
| 12 |
|  |

# Q4. What is the revenue generated from gold users?

By analyzing the revenue generated from gold users, businesses can evaluate the impact of the loyalty program on their overall business performance. This information can help businesses make informed decisions about how to allocate resources to improve customer retention and increase revenue. Additionally, it can provide insights into which products or services are most popular among gold users and identify potential areas for expansion or improvement. Ultimately, understanding the revenue generated from gold users can help businesses maximize the effectiveness of their loyalty program and drive growth and profitability.

Query: SELECT

SUM(Sale\_PRICE) AS GoldUserRevenue

FROM

GOLD\_USER\_SIGNUP AS S

JOIN

Users\_PROJECT AS U ON S.User\_ID = U.User\_ID

WHERE

S.Membership = 'GOLD';

# Q5. What is the total revenue generated from gold users?

If the revenue generated from gold users is significantly higher than that of non-gold users, then it suggests that the loyalty program is successful in driving sales and customer loyalty. On the other hand, if the revenue generated from gold users is not significantly different from that of non-gold users, then it may suggest that the rewards offered by the loyalty program are not attractive enough to incentivize customers to spend more.

Query: SELECT

User\_ID,

Membership,

SIGNUP\_DATE,

CURRENT\_DATE - SIGNUP\_DATE AS DaysAsGoldMember

FROM

GOLD\_USER\_SIGNUP

WHERE

Membership = 'GOLD';

# Q6. Which users has been a gold user for the How much of time?

Identifying which users have been gold users for how long can provide valuable insights into the loyalty of customers and the effectiveness of the loyalty program. By analyzing this data, businesses can identify their first gold user and use this information to better understand their most loyal customers. Additionally, by knowing how long each gold user has been a member, businesses can identify trends in loyalty and use this information to optimize their loyalty program.

Query:

SELECT

S.Product\_ID,

SUM(S.Quantity\_SOLD) AS TotalQuantitySold

FROM

Sales\_PROJECT AS S

JOIN

GOLD\_USER\_SIGNUP AS U ON S.User\_ID = U.User\_ID

WHERE

U.Membership = 'GOLD'

GROUP BY

S.Product\_ID

ORDER BY

TotalQuantitySold DESC

LIMIT 1;

# Q7. What is the most popular product among gold users?

Identifying the most popular product among gold users is crucial in determining which products to prioritize in our menu and marketing campaigns. Gold users are our premier customers, and their purchasing behavior can provide valuable insights into their preferences. By analyzing their order history, we can identify the product that is most frequently ordered by gold users, and use this information to develop strategies to increase sales and customer loyalty.

QUERY:

SELECT p.Product\_Name, SUM(s.Quantity\_SOLD) AS TotalQuantitySold FROM Sales\_PROJECT AS s

JOIN USERS\_PROJECT u ON s.User\_ID = u.User\_ID

JOIN GOLD\_USER\_SIGNUP AS G ON U.USER\_ID = G.USER\_ID

JOIN Products\_PROJECT p ON s.Product\_ID = p.Product\_ID

WHERE G.Membership = 'GOLD'

GROUP BY p.Product\_Name

ORDER BY TotalQuantitySold DESC

LIMIT 1;

# Q8. What is the total sales revenue generated in each year?

This is an important analysis to determine the pattern of sales over the years, identifying which years had the highest and lowest sales revenue. This information can help us understand which factors contribute to sales fluctuations and predict future sales trends.

QUERY:

SELECT

EXTRACT(YEAR FROM CREATED\_DATE) AS Year,

SUM(Quantity\_Sold \* Sale\_Price) AS TotalRevenue

FROM

SALES\_PROJECT

GROUP BY

Year

ORDER BY

Year;

# Q9. How has the sales revenue changed over the years?

To analyze the change in sales revenue over the years, we can calculate the percentage increase or decrease in revenue each year. This will help us understand the trends in our business and identify which years have shown the most growth or decline in sales. By studying this data, we can make informed decisions about where to allocate resources and focus our efforts to maximize revenue in the future.

QUERY:

SELECT

EXTRACT(YEAR FROM CREATED\_DATE) AS Year,

SUM(Quantity\_Sold \* Sale\_Price) AS TotalRevenue

FROM

SALES\_PROJECT

GROUP BY

Year

ORDER BY

Year;

# Q10. What is the average Gold-signup compare to just sign up for the users?

From an average perspective, we can understand what percentage of our total users have signed up for the gold membership compared to those who haven’t, which can give us insights into the popularity and effectiveness of our loyalty program.

QUERY:

SELECT

COUNT(\*) AS GoldSignups,

(SELECT COUNT(\*) FROM Users\_PROJECT) AS TotalSignups,

((SELECT COUNT(\*) FROM GOLD\_USER\_SIGNUP WHERE Membership = 'GOLD') \* 1.0 / (SELECT COUNT(\*) FROM Users\_PROJECT)) AS AverageGoldSignupRatio

FROM

GOLD\_USER\_SIGNUP

WHERE

Membership = 'GOLD';

# ****Q11. How many gold members users have order how many numbers of time ?****

“We can analyze the difference in the number of orders made by gold members before and after signing up for the membership. This will help us understand whether there is a significant difference in ordering behavior and whether there are any weaknesses in our service that we need to address. Additionally, we can also identify how many times on average gold members have placed an order to assess their loyalty and engagement with our platform.”

QUERY:

SELECT U.User\_ID, COUNT(S.Product\_ID) AS OrderCount FROM Users\_PROJECT AS U

JOIN GOLD\_USER\_SIGNUP AS G ON U.User\_ID = G.User\_ID

JOIN SALES\_PROJECT AS S ON U.User\_ID = S.User\_ID

WHERE G.Membership = 'GOLD' GROUP BY U.User\_ID;

# Q12. What is the total amount each customer spend on Online Food Delivery?

From this, we can understand the total amount that each customer spends on online food delivery, which can help us identify high-spending customers and tailor our services accordingly.

QUERY:

SELECT

USER\_ID,

SUM(SALE\_PRICE) AS TotalAmountSpent

FROM

SALES\_PROJECT

GROUP BY

USER\_ID;

# Q13. What is the frequency of customer visits to the online platform?

By analyzing the frequency of customer visits, we can gain insights into user engagement and loyalty towards the platform. We can use this information to identify areas for improvement, such as providing personalized recommendations, improving the user interface, or offering more discounts or loyalty rewards. Additionally, we can use this data to predict future customer behavior and make data-driven decisions.

QUERY:

SELECT USER\_ID, COUNT(PRODUCT\_ID) AS VisitFrequency FROM SALES\_PROJECT GROUP BY USER\_ID;

# Q14. What was the first order purchase by each customer ?

Knowing the first order purchase by each customer can help us understand their taste and preferences, which can help us personalize their experience and offer them relevant recommendations. This can lead to increased customer satisfaction and loyalty. Additionally, analyzing the data can help identify popular first orders and assist in developing marketing strategies to promote these items.

QUERY: SELECT USER\_ID, MIN(CREATED\_DATE) AS FirstOrderDate FROM SALES\_PROJECT GROUP BY USER\_ID;

# Q15. What is the most purchase item on the menu and how many times was it purchased by all customers?

By identifying the most purchased item on the menu and the number of times it has been purchased by all customers, we can make informed decisions about which menu items to feature prominently on the platform. This can help increase sales and customer satisfaction by making it easier for customers to find and order their favorite items. Additionally, by analyzing the most popular menu items, we can identify trends in customer preferences and adjust our menu offerings accordingly to meet the changing demands of our customers.

QUERY:

SELECT

p.Product\_Name,

SUM(s.QUANTITY\_SOLD) AS TotalQuantitySold

FROM

SALES\_PROJECT AS s

JOIN

PRODUCTS\_PROJECT AS p ON s.Product\_ID = p.Product\_ID

GROUP BY

p.Product\_Name

ORDER BY

TotalQuantitySold DESC

LIMIT 1;

# Q16. Which item was most popular for each customer ?

Analyzing the most popular item for each customer can help us personalize their experience by recommending their favorite items or creating personalized offers. It can also help in predicting future demand for certain items and optimizing inventory management. Furthermore, it can provide insights into the taste preferences of our customers and help us develop targeted marketing strategies.

QUERY:

SELECT USER\_ID, Product\_ID, SUM(QUANTITY\_SOLD) AS TotalQuantity FROM SALES\_PROJECT AS S GROUP BY USER\_ID, Product\_ID

HAVING

SUM(QUANTITY\_SOLD) = (SELECT MAX(TotalQuantity)FROM (SELECT SUM(QUANTITY\_SOLD) AS TotalQuantity

FROM Sales\_PROJECT AS innerSales WHERE innerSales.USER\_ID = S.USER\_ID GROUP BY Product\_ID) AS subquery);

# Q17. Which item was purchase first by the customer after they become a member ?

Analyzing the first order purchased by the customer after they become a gold member can give us insight into their taste preferences and what items they prioritize in their food choices. It can also give us an idea of what types of promotions or offers are most likely to attract gold members to make a purchase. Additionally, tracking the first purchase of gold members can help us identify any patterns or trends in their behavior that we can use to improve our services or offerings.

QUERY:

SELECT S.user\_ID, MIN(S.CREATED\_DATE) AS FirstOrderDate, S.Product\_ID FROM SALES\_PROJECT AS S

JOIN USERS\_PROJECT AS U ON S.USER\_ID = U.USER\_ID

JOIN GOLD\_USER\_SIGNUP AS G ON U.USER\_ID = G.USER\_ID

WHERE

S.CREATED\_DATE >= G.SIGNUP\_DATE GROUP BY S.user\_ID, S.Product\_ID ORDER BY S.user\_ID, FirstOrderDate;

# Q18. Which item was purchase before the customer become a member ?

By knowing the last product purchased before the customer became a member, we can analyze if there is any specific product that encouraged the customer to become a member. This can help us create targeted marketing strategies to attract more customers to become members.

QUERY:

SELECT s.USER\_ID, p.Product\_Name, s.CREATED\_DATE FROM SALES\_PROJECT AS s

JOIN USERS\_PROJECT AS u ON s.USER\_ID = u.User\_ID

JOIN GOLD\_USER\_SIGNUP AS G ON G.USER\_ID = U.USER\_ID

JOIN PRODUCTS\_PROJECT AS p ON s.Product\_ID = p.Product\_ID

WHERE s.CREATED\_DATE < G.SIGNUP\_DATE;

# Q19. What is the total orders and amount spent for each member before they become a member ?

It can also help in identifying potential loyal customers who are likely to become gold members in the future.

QUERY: SELECT

s.USER\_ID,

COUNT(s.PRODUCT\_ID) AS TotalOrders,

SUM(s.SALE\_PRICE) AS TotalAmountSpent

FROM

SALES\_PROJECT AS s

JOIN

USERS\_PROJECT AS u ON s.USER\_ID = u.USER\_ID

JOIN

GOLD\_USER\_SIGNUP AS G ON u.USER\_ID = G.USER\_ID

WHERE

s.CREATED\_DATE < G.SIGNUP\_DATE

GROUP BY s.USER\_ID;

# Q20. Rank all the transactions for each member whenever they are a XYZ gold member for every non gold member Transaction marks as na ?

Ranking all the transactions for each member during their gold membership can help us to identify the most popular items purchased by gold members, as well as their spending patterns during their gold membership period. This information can be used to optimize our menu, promotions, and discounts for gold members to increase their satisfaction and loyalty to our platform. Additionally, comparing the ranked transactions of gold members to non-gold members can help us identify any Sguide our marketing and sales strategies.

QUERY: SELECT user\_id,

CASE

WHEN membership = 'GOLD' THEN CAST(ROW\_NUMBER() OVER (PARTITION BY user\_id ORDER BY signup\_date) AS TEXT)

ELSE

'NA'

END AS signup\_rank FROM gold\_user\_signup ORDER BY user\_id, signup\_date;