SQL Interview Questions Collection

Q1. write an SQL query to find the name of the product with the highest price in each country .Return product_name, price, country

You have two tables: Product and Supplier
product table columns: product_id, product_name, supplier_id, Price
Supplier table columns: supplier_id, supplier_name, Country
-- Steps to solve
-- 1. join the tables on supplier_id COLUMN
-- 2. group by country
-- 3. Maximum price for each group

-- Solution 1 with CTE

```
WITH CTE as (
        SELECT * ,
            row_number() over(partition by s.country ORDER by price
DESC) as rn,
            Dense_rank() over(partition by s.country order by price
DESC) as d_rank
        from product as P
        JOIN supplier as s
        ON p.supplier_id = s.supplier_id
        )
SELECT product_name,price, country
FROM cte
where d_rank=1;
```

-- Solution 2 using subquery

```
SELECT product_name, price, country
```

```
FROM(
        SELECT *,
        row_number() OVER (partition by s.country ORDER by p.price DESC)as
rn
        FROM product as p
        JOIN supplier as s
        ON p.supplier_id = s.supplier_id) as subquery
WHERE rn =1;
```

Q.2. write a SQL query to calculate the total transaction amount for each customer for the current year. The output should contain customer_name and total amount.

You have two tables : customers and transactions.
-customer table : customer_id, customer_name, registration_date
- transaction table: transaction_id, customer_id,transaction_date, amount

```
-- Solution 1
SELECT c.customer_name, c.customer_id, sum(t.amount) as total_amount
FROM customers as c
JOIN transactions as t
ON c.customer_id = t.customer_id
WHERE extract(YEAR from t.transaction_date) = extract(YEAR from
current_date)
GROUP BY 1,2;

-- Solution 2
SELECT c.customer_name,  sum(t.amount) as total_amount
FROM customers as c
JOIN transactions as t
ON c.customer_id = t.customer_id
WHERE extract(YEAR from t.transaction_date) = extract(YEAR from
current_date)
GROUP BY c.customer_id;
```

Q3. Write a query to return the IDs of the Facebook pages that have zero likes. The output should be sorted in ascending order based on the page IDs.

Assume you're given two tables containing data about Facebook Pages and their respective likes (as in "Like a Facebook Page")
pages Table:page_id, page_name
page_likes Table:user_id, page_id, liked_date

```
-- Solution
SELECT p.page_id
FROM pages as p
LEFT JOIN page_likes as l
ON p.page_id = l.page_id
GROUP BY p.page_id
HAVING count(l.page_id)=0
ORDER BY p.page_id;
```

Q4. Write a query to calculate the click-through-rate CTR for the app in 2022 and round the results to 2 decimal places.

percentage of ctr = 100.0* number of clicks / number of impressions
to avoid integer devision, multiply the ctr by 100.0, not 100
events table : app_id, event_type (clicks / impression), timestamp

```
-- Solution
SELECT
        app_id,
        ROUND( 100.0 *
        SUM(CASE WHEN event_type = 'click' THEN 1 ELSE 0 END)/
        SUM(CASE WHEN event_type = 'impression' THEN 1 ELSE 0 END),2)
FROM
        events
WHERE
        EXTRACT (YEAR FROM timestamp) = '2022'
GROUP BY
        app_id;
```

Q5. Write a SQL query to calculate the difference between the highest salaries in the marketing and engineering department output the absolute difference in salaries.

Leetcode problem LeetCode SQL Premium Problem 2853: 'Highest Salary Difference'

```
-- Solution (Using case)
SELECT
        abs
        (
                MAX(CASE WHEN department= 'Marketing' Then salary END) as
Mark_highest_sal -
                MAX(CASE WHEN department= 'Engineering' Then salary END)
as Eng_highest_sal
        )
FROM salaries


-- Solution (Using CTE)
WITH Mark_highest_sal as
(
        SELECT MAX(Salary) mark_sal
        FROM salaries
        WHERE department = 'Marketing'
)
,Eng_highest_sal AS
(
        SELECT MAX(Salary) eng_sal
        FROM salaries
        WHERE department = 'Engineering'
)
SELECT ABS( SELECT mark_sal FROM Mark_highest_sal  - SELECT eng_sal
FROM Eng_highest_sal )
 as abs_diff;
```

Q6. Write a query to identify the top two power users who sent the highest number of messages ON Microsoft Teams in August 2022. Display the IDs of the 2 users along with the total number of messages they sent.

Output the result in descending order based on the count of the messages.

messages Table : message_id, sender_id, receiver_id, content, sent_date

-- Solution

```
SELECT sender_id, COUNT(1) as message_count
FROM messages
WHERE EXTRACT (MONTH FROM sent_date) = 8 AND
        EXTRACT (YEAR FROM sent_date) = 2022
-- WHERE sent_date BETWEEN '2022-08-01' AND '2022-08-31'
GROUP BY sender_id
ORDER BY message_count DESC
LIMIT 2;
```

Q7. Write a sql query to find the average order amount for the male and female customers separately. return the results with 2 decimal. Customer segmentation problem

IBM Data Analyst interview question
customers table : customer_id, customer_name, age, gender
orders table: order_id, customer_id, order_date, total_amount

-- Solution

```
SELECT c.gender AS Gender, ROUND( avg(o.total_amount), 2) as
avg_order_amount
FROM customer c
JOIN order o
ON c.customer_id = o.customer_id
GROUP BY gender;
```

Q8. Write a sql query to find out the total sales revenue generated for each month in the year 2023.

Sales table : order_id, order_date, product_id, quantity, price_per_unit

-- Solution

```
SELECT
    TO_CHAR (order_date, 'month') AS month_name,
    SUM(quantity * price_per_unit)
FROM
    sales
WHERE
    EXTRACT(YEAR FROM order_date) = '2023'
```

```
GROUP BY
        month_name
ORDER BY
        EXTRACT (MONTH FROM order_date)
```

Q9. Write a sql query to obtain the third transaction of every user from the transactions table.

UBER data analyst interview question
output the user_id, spend, transaction_date

```
-- Solution
        WITH CTE as
        (
                SELECT user_id,
                                spend,
                                transaction_date,
                                ROW_NUMBER() OVER (partition by user_id order
by transaction_date) as rn
                FROM transactions
        )
         SELECT *
         FROM CTE
         WHERE rn =3;
```

Q10. Find the top 5 products whose revenue has decreased in comparison to the previous year(both 2022 and 2023) Return the product_name, revenue for the previous year, revenue for the current year, revenue decreased and decreased ratio.

(prev_revenue -current_revenue)/prev_revenue) *100
Sales table : product_name, year, revenue
--- if table having date column with different years than first need to filter
for only 2022 and 2023 and aggregate the sale for each year group by product.

```
-- Solution
        WITH revenue AS
        (
                SELECT
                        product_name,
                        year,
```

```
                revenue as current_revenue,
                LAG(revenue) OVER (PARTITION BY product_name
ORDER BY year) as prev_revenue
        FROM sales
)
SELECT
        Product_name,
        prev_revenue,
        current_revenue,
        (prev_revenue - curent_revenue) as revenue_decreased,
        ((prev_revenue - curent_revenue)/ prev_revenue)*100
decreased_ratio%
FROM revenue
WHERE prev_revenue IS NOT NULL AND prev_revenue >
current_revenue
ORDER BY revenue_decreased DESC
LIMIT 5
;
```

Q11. Write a query to calculate the total viewership for the laptops and mobile devices,where mobile is defined as the sum of tablets and phone viewership.

output the total viewership for laptop_views and total viewership for mobile devices as mobile_views.

```
-- Solution
        SELECT
                SUM(CASE WHEN device_type IN ('Phone', 'Tablet') THEN 1 ELSE
0 END)AS mobile_views,
                SUM(CASE WHEN device_type = 'Laptop' THEN 1 ELSE 0 END)
AS laptop_views
FROM viewership
;
```

Q12. write a query to identify the top two highest grossing products within each category in the year 2022.

Assume you are given a table containing data on Amazon customers and their spending on products in different category .

The output should include category, product and total spendproduct_spend table : category, product, user_id, spend, transaction_date

-- Solution

```
WITH CTE As(
    SELECT
        category,
        product,
        SUM(spend) as total_spend,
        ROW_NUMBER() OVER(partition by category order by SUM(spend)
DESC) AS rn
   FROM
     product_spend
   WHERE
     EXTRACT(YEAR from transaction_date) = 2022
   GROUP BY
     category,product
)
SELECT
  category,
    product,
    total_spend
FROM CTE
WHERE rn <3
;
```

Q13. Write a query to obtain a histogram of tweets posted per user in 2022. Output the tweet count per user as the bucket and the number of Twitter users who fall into that bucker.

Twitter question : HISTOGRAM
In other words , group the users by the number of tweets the posted in 2022 and count the
number of users in each group.
tweets table : tweet_id, user_id, msg, tweet_date

-- Solution

```
SELECT
        tweet_bucket,
        count(*) as user_num
FROM
```

```
(
        SELECT
                user_id,
                count(*) as tweet_bucket
        FROM
                tweets
        WHERE
                EXTRACT(year from tweet_date) = 2022
        group by 1
) as x
GROUP BY
        tweet_bucket
ORDER BY
        tweet_bucket
;
```

Q14. Leetcode -185 Department top 3 salaries. A company's executives are interested in seeing who earns the most money in each of the company's departments. A high earner in a department is an employee who has a salary in the top three unique salaries for that department.

employee table: id, name, salary, departmentID
department table: id, name
id is the primary key (column with unique values) for this table.
departmentId is a foreign key (reference column) of the ID from the Department table.
Output table : department, employee, salary

```
-- Solution
        SELECT
                department,
                employee,
                salary
        FROM
        (
                SELECT
                        d.name as department,
                        e.name as employee,
                        e.salary as salary,
                        DENSE_RANK() OVER (partition by d.name order by e.salary
desc) as rnk
```

```
        FROM
                employee e
        JOIN
                department d
        ON
                e.departmentID = d.id
) as x
WHERE
        rnk <=3
ORDER BY
        department, salary DESC
;
```

Q15.Write a SQL query to find for each month and country, the number of transactions and their total amount, the number of approved transactions and their total amount.

transactions table: id, country, status, amount, trans_date
OUtput : month, contry, trans_count, approved_count, trans_total_amount, approved_total_amount

```
-- Solution
        SELECT
                to_char(trans_date, 'YYYY-MM') as month,
                country,
                count(*) as trans_count,
                SUM(CASE WHEN status= 'approved' THEN 1 ELSE 0 END) as
approved_count,
                SUM(amount) as trans_total_amount,
                SUM(CASE WHEN status= 'approved' THEN amount ELSE 0 END) as
approved_total_amount
        FROM
                transactions
        GROUP BY 1,2
        ;
```

Q16. write a query to retrieve the average star rating for each product, group by month. the output should display the month as a numerical value, product ID,and average star rating rounded to two decimal places.sort the output first by month and then by product ID

reviews table : review_id, user_id,submit_date, product_id, stars

-- Solution

```
SELECT
        to_char(submit_date, 'MM') as month,
        -- EXTRACT( month from submit_date) as month,
        product_id,
        ROUND( avg(stars),2) as avg_ratings
FROM
        reviews
GROUP BY 1,2
ORDER BY month, product_id
;
```

Q17. Identify IBM High capacity users.Identify users who have made purchases totaling more than $10,000 in the last month from the purchase table.

purchases table : user_id, purchase_date, product_id, amount

-- Solution

```
SELECT
        EXTRACT(month from purchase_date) as month
        user_id,
        SUM(amount) as total_amount
FROM
        purchases
WHERE
        Extract (month from purchase_date) = Extract(month from
current_date) - INTERVAL '1 month'
GROUP BY 1,2
HAVING SUM(amount) >10000;
```

Q18. Average duration of employee's service. Given the data on IBM employees, can you find the average duration of service across different departments?

The duration of the service is represented as end_date - start_date.
if the end_date is null, consider it as the current_date.
employees table: emp_id, emp_name, start_date, end_date, department

-- Solution

```
SELECT
        department,
        AVG(
                (CASE WHEN end_date ISNULL THEN current_date ELSE
end_date END)
        - start_date ) as avg_duration
FROM employees
GROUP BY 1
;
```

Q19. Identify the top 3 posts with the highest engagement (likes + comments) for each user on a facebook page.

Display the user_id, engagement count rank for each post.
posts table : post_id, user_id,likes, comments

-- Solution

```
SELECT
        user_id, post_id, engagement_count, rnk
FROM (
        SELECT
                user_id, post_id,
                SUM(likes + comments) as engagement_count,
                DENSE_RANK() OVER (partition by user_id
                                                order by
SUM(likes+comments) DESC) as rnk
        FROM posts
        GROUP BY user_id, post_id
        ) as x
WHERE rnk <=3
;
```

Q20. Determine the users who have posted more than 2 times in the past week and calculate the total number of likes they have received.

return user ID and number of post and number of likes.
posts table : post_id, user_id, likes, post_date

-- Solution

```
SELECT
```

```
        user_id,
        count(post_id) as number_of_post,
        SUM(likes) as number_of_likes
FROM posts
WHERE post date between current_date and current_date - interval '7
days'
GROUP BY user_id
HAVING count(post_id) > 2
;
```

Q21. write a query to retrieve the count of companies that have posted duplicate job listings.

Linkedin data analyst interview question Assume you're given a table containing job postings form various companies on the linkedin platform. Duplicate job listings are defined as two job listings within the same company that share identical titles and descriptions.
job_listings table : job_id, company_id, title, description

```
-- Solution
SELECT
    count(*) as count_companies
FROM(
    SELECT
        company_id,
        title,
        description,
        count(job_id) as count_job
    FROM job_listings
    GROUP BY company_id, title, description
    HAVING count(job_id) > 1
    ) as x
;
```

Q22. Identify the region with the lowest sales amount for the previousMonth.

Flipkart business analyst sql interview question
-- group by region and months
-- filter by last MONTH
-- order by asc and limit 1 or use min function on amount

sales table : region, sale_date, amount

-- Solution

```
WITH CTE AS(
SELECT
	Region,
	EXTRACT (month from current_date - interval '1 month') as
prev_month,
	SUM(amount) as total_amount,
	RANK() OVER (ORDER BY sum(amount) ASC) as rnk
FROM sales
WHERE
	EXTRACT (MONTH from saleDate)=EXTRACT (month from
current_date - interval '1 month')
	AND
	EXTRACT(YEAR FROM saleDate)=EXTRACT(YEAR FROM
current_date)
GROUP BY
	region,2
)

SELECT
	region,
	total_amount
FROM CTE
WHERE rnk = 1
;
```

Q23. Find the median within a series of numbers in SQL. TIK TOK Data Analyst Interview Question.

1 8 3 4 5 odd
1 3 4 5 8 9 even
tiktok table : views column

-- Solution

```
WITH CTE AS(
	SELECT
		views
		,ROW_NUMBER() OVER (ORDER BY views ASC)
asc_rank
```

```
                ,ROW_NUMBER() OVER (ORDER BY views DESC)
desc_rank
        FROM tiktok)

SELECT
        AVG(views) as median
FROM
        CTE
WHERE
        ABS(desc_rank - asc_rank) <=1
;
```

Q24. How many delayed orders do delivery partners have, considering the predicted delivery time and the actual delivery time?

Zomato Business analyst Interview Question
order_details table : order_id, del_partner, estimated_time, actual_time
ex. (111, 'Amit', '2024-11-28 07:10:23', '2024-11-28 07:29:56')

```
-- Solution
SELECT del_partner,
                count(order_id) as n_delayed_order
FROM order_details
WHERE estimated_time < actual_time
Group BY del_partner
;
```

Q25. Which metro city has the highest number of restaurant orders in September 2021?Write a query to retrieve the city name and the total count of orders , ordered by the total count of orders in descending order.

Swiggy Business Analyst Interview Question
Metro city - Mumbai, Delhi, Banglore, Hyderabad
orders table : order_id, restaurant_id, city, order_date

```
-- Solution
SELECT
        city,
        count(order_id) as total_orders
FROM orders
WHERE city in ('Mumbai', 'Delhi','Banglore','Hyderabad')
```

```
                AND date_part('Year', order_date) = 2021
                AND date_part('Month', order_date) = 9
                -- AND order_date between '2021-09-01' AND '2021-09-30'
GROUP BY 1
ORDER BY 2 DESC
LIMIT 1
;
```

Q26. Get the count of distinct student that are unique

```
-- Solution
        SELECT  count(*) count_of_distinct_unique_student
        FROM(
            SELECT name ,COUNT(name)
            FROM students
            GROUP BY name
            HAVING count(name) =1
```

Q27. Get the count of distinct student that are not unique

```
-- Solution
        SELECT count(*) count_of_students_not_unique
        FROM(
            SELECT name, count(name)
            FROM students
            GROUP BY name
            HAVING Count(name)>1
        );
```

Q28. Find the city wise customers count who have placed more than three orders in november 2024

order table : order_id, city, customer_id, order_date, amount
Zomato business anlayst interview Question
--group by city
-- count(customer)
-- filter more than 3 orders and date in november 2024

-- Solution

```
SELECT city, count(customer_id) tota_custmer_count
FROM(
        SELECT  city,
                        customer_id,
                        count(order_id) as n_orders
        FROM orders
        WHERE order_date between '2024-11-01' AND '2024-11-30'
        GROUP BY 1,2
        HAVING count(order_id) > 3
        ) as x
GROUP BY city;
);
```

Q29. Find the top-performing two months by revenue for each hotel for each year.

Booking.com Data Analyst Interview Question
Return hotel_id, year, months, revenue
Solution :
-- Group by hotel, year , months and year
-- aggregate amount
-- use DENSE_RANK() window function

-- Solution

```
SELECT hotel_id, year, month, revenue
FROM (
                SELECT
                        hotel_id,
                        date_part('YEAR', booking_date) as year,
                        date_part('Month', booking_date) as month,
                        SUM(amount) as revenue,
                        DENSE_RANK() OVER(PARTITION BY hotel_id,

date_part('YEAR', booking_date),

date_part('Month', booking_date)
                                                ORDER BY SUM(amount)
DESC)) as d_rank
                FROM bookings
```

```
                GROUP BY 1,2,3
        ) as x
WHERE d_rank <=2
;
```

Q30. Write a SQL query to retrieve the emp_id, emp_name, manager_name from a given employee table. It's important to note that managers are also employees in the table.

TCS Data Analyst Interview Question
employees table : emp_id, emp_name, manager_id

```
-- Solution
SELECT e.emp_id,
        e.emp_name,
        e.manager_id,
        m.emp_name as manager_name
FROM employees e
JOIN employees m
ON e.manager_id = m.emp_id
;
```

Q31. write a SQL query to find all salaries greater than the average salary .Given the employee table as emp_id and salary. Return emp_id and salary

```
-- Solution
SELECT emp_id, salary
FROM employee
WHERE salary > (select avg(salary) from employees)
;
```

Q32. Write a SQL query to find all the duplicate email addresses in the customer table.

Consider a table named customers : customer_id, first_name, last_name, email

```
-- Solution
SELECT customer_id, first_name, last_name, email
FROM(
        SELECT customer_id,
```

```
                first_name,
                last_name,
                email,
                ROW_NUMBER() OVER (partition by email order by email)
as rn
        FROM customers
        )
WHERE rn > 1
;
```

-- Solution

```
SELECT email, count(*) cnt_email
FROM customers
GROUP BY email
HAVING count(email)>1
;
```

Q33. Write a SQL query to calculate the running total revenue for each combination of date and product ID.

Flipkart Business Analyst Interview Question.
Table orders
Expected output : date, product_id, product_name, revenue, running_total. order by product_id, date ascending

-- Solution (using WINDOW function)

```
SELECT date, product_id, product_name, revenue,
            sum(revenue) over (partition by date, product_id order by
product_id, date) as running_total
FROM orders;
```

-- Solution (using SELF JOIN)

```
SELECT o1.date,
        o1.product_id,
        o1.product_name,
        o1.revenue
        SUM(o2.revenue)
FROM orders o1
JOIN orders o2
ON o1.product_id = o2.product_id
```

```
        and o1.date >= o2.date
GROUP BY 1,2,3,4
ORDER BY 2,1
;
```

Q34. Write a SQL query to find the top 5 customers with the highest percentage of return items out of their total purchases.

Amazon Data Analyst Interview Question (Hard Category Question)

Suppose you have given two tables - Orders and Returns. the table contain information regarding orders and returns by customers.

Return Customer_id and the percentage of return items rounded by 2 decimal places.

-- orders : order_id, customer_id, order_date, order_items, amount

-- returns : return_id, order_id, return_date, return_items, amount

```
-- Solution
SELECT x.customer_id,
        (SUM(x.return_items)/ SUM(x.order_items)) *100 as
percenatge_return,
FROM(
        SELECT
        (SELECT order_id,customer_id, sum(order_items) as order_items
        FROM orders
        GROUP BY 1,2) as o
        INNER JOIN
        (SELECT return_id, order_id,sum(return_items) as return_items
        FROM returns
        GROUP BY 1,2) r
        ON o.order_id = r.order_id
        ) x
GROUP BY x.customer_id
ORDER BY percenatge_return
LIMIT 5
;

-- Note : The solution may change depends on what input table provided in questions.
```

Q35. Write a SQL query to fetch user_ids that have only bought both 'Burger' and 'Cold drink'.

Flipkart Expected output column : user_id
-- orders table : user_id, items

-- Solution

```
SELECT user_id
FROM orders
GROUP BY user_id
HAVING COUNT(DISTINCT items) = 2
        AND SUM(CASE WHEN items IN ('Burger', 'Cold drinks' THEN 1  ELSE 0
END) =2
;
```

Q36. Write a query to find top 3 sellers with highest sales amount but lowest return quantity.

AMAZON : orders table : order_id, seller_id, sales_amount
returns table : return_id, seller_id, return_qty

-- Solution

```
SELECT o.seller_id, total_sale, return_qty FROM
((SELECT seller_id, SUM(sale_amount) as total_sale FROM orders
GROUP BY 1) o
LEFT JOIN
(SELECT seller_id, SUM(COALESCE(return_quantity,0)) as return_qty
FROM returns GROUP BY 1) r
ON o.seller_id = r.seller_id)
ORDER BY 2 DESC, 3 ASC
;
```

Q37. Write a solution to select the product id, year, quantity and price for the first year of every product sold.

Walmart Data Analyst Interview Question

-- Solution

```
SELECT product_id,
                MIN(EXTRACT (YEAR from order_date)) as first_year,
                SUM(quantity),
```

```
                SUM(price)
FROM sales
GROUP BY product_id
ORDER BY 1;
```

Q38. Spotify - Write a sql query to find the 10 most popular songs by total number of listens.

songs table : song_id, song_name, artist_name
listens table : listen_id, user_id, song_id, listen_date

```
-- Solution
WITH CTE AS(
    SELECT s.song_name
        ,n_listens
        ,DENSE_RANK()OVER(order by n_listens DESC) as
popularity_rank
    FROM songs s
    JOIN
    (SELECT song_id , count(listen_id) as n_listens FROM listens
GROUP BY song_id) l
    ON s.song_id = l.song_id
    )
SELECT *
FROM CTE
WHERE popularity_rank <=10
;
```

Q39. Write the SQL query to find the second highest salary

```
-- Solution 1
SELECT max(salary) FROM employees
    WHERE  salary <  ( SELECT  max(salary) FROM employees);

-- Solution 2
SELECT salary
FROM(
    SELECT salary,
        RANK() OVER (order by salary DESC) as rnk
    )
```

```
WHERE rnk =2;
```

Q40. write an SQL query to calculate the total numbers of returned orders for each month.

Given the Orders table with columns OrderID, OrderDate, and TotalAmount, and the Returns table with columns ReturnID and OrderID

-- Solution

```
SELECT EXTRACT(MONTH FROM o.orderDate) || '-' || EXTRACT(Year
FROM o.orderDate) as month,
        COUNT(r.returnID)
FROM orders o
JOIN returns r ON o.orderID = r.orderID
GROUP BY month;
```

Q41. Write SQL query to find the top-selling products in each category.

assuming products table has column
product_id, product_name, category, quantity_sold

-- Solution

```
SELECT category, product_name
FROM(
    SELECT category, product_name, SUM(quantity_sold) ,
        RANK() OVER( PARTITION BY category, product_name order by
    SUM(quantity_sold) DESC) rnk
    FROM products
    GROUP BY category, product_name
    )
WHERE rnk =1;
```

Q42. Find the top 2 products in the top 2 categories based on spend amount?

-- Solution

```
Select a.category, a.product, a.total_spend, b.cat_rank, a.prod_rank
from
        (select category, product, sum(spend) total_spend
```

```
        ,RANK()OVER(partition by category order by
sum(spend)desc) prod_rank
from orders_tb
group by category, product) a
JOIN
(select category, sum(spend) total_spend
        ,RANK()OVER(order by sum(spend)desc) cat_rank
from orders_tb
group by category) b
ON a.category = b.category
WHERE cat_rank <=2 and Prod_rank <=2
order by cat_rank, prod_rank;
```

Q43. Write an SQL query to find customers who haven't made any purchases in the last month, assuming today's date is April 2, 2024.

Given tables customers (columns: customer_id, name, email) and orders (columns: order_id, customer_id, order_date, amount),

```
-- Solution
SELECT DISTINCT customer_id, name FROM customers
WHERE customer_id NOT IN
    ( select customer_id from orders
        where order_date between date_trunc ('month',current_date)
    - interval '1 month'
            and date_trunc ('month',current_date) - interval '1 day'
        ;
```

Q44. Find customer who has done purchase this month and also last month

```
-- Solution
SELECT DISTINCT customer_id, name FROM customers
WHERE customer_id NOT IN
    ( select customer_id from orders
        where order_date between date_trunc ('month',current_date)
    - interval '1 month'
            and current_date' ;
```

Q45. How would you identify duplicate entries in a SQL in given table employees?

columns are emp_id, name, department, salary

-- Solution

```
SELECT emp_id, name, department, salary, count(*)
FROM employees
GROUP BY emp_id, name, department, salary
HAVING count(*) >1;
```

Q46. Write a SQL query to find all products that haven't been sold in the last six months.

Products : product_id, product_name, category, price
Sales : sales_id, product_id, sale_date, quantity
Return the product_id, product_name, category, and price of these products.

-- Solution

```
SELECT * FROM products
WHERE product_id NOT IN (
        SELECT DISTINCT product_id FROM sales
        WHERE sale_date between date_trunc('month' , current_date) -
        interval '6 months' and
                                        date_trunc('month' , current_date)-1
        as end_date);
```

Q47. write a SQL query to find customers who bought Airpods after purchasing an iPhone.

-- Solution

```
SELECT * from customers_tb
WHERE customerid in (
        select customerid FROM
        (
                SELECT customerid
                        ,MIN(CASE WHEN productname ='iPhone' THEN
purchasedate END) as iPhone_date
                        ,MAX(CASE WHEN productname ='Airpods' THEN
purchasedate END) as Airpods_date
                from purchases
```

```
            group by customerid
        )
        WHERE iPhone_date IS NOT NULL and
                airpods_date is not null and
                iPhone_date < airpods_date
);
```

Q48. what is the % of chance is there that the customer who bought MacBook will buy an Airpods

-- Solution

```
Select
(select count(*) FROM
(select * from purchases WHERE productname ='MacBook') a
JOIN purchases b ON a.customerid = b.customerid
    and b.productname = 'Airpods'
    and a.purchasedate< b.purchasedate) /
(select count(*) from purchases where
productname='MacBook')*100.0::float
||'%' pct_chance;
```

Q49. Write a SQL query to classify employees into three categories based on their salary

"High" - Salary greater than $70,000
"Medium" - Salary between $50,000 and $70,000 (inclusive)
"Low" - Salary less than $50,000
Your query should return the EmployeeID, FirstName, LastName, Department, Salary, and
a new column SalaryCategory indicating the category to which each employee belongs.

-- Solution

```
SELECT EmployeeID, FirstName, LastName, Department, Salary,
    CASE WHEN salary > 70000 THEN 'High'
        WHEN salary >= 50000 THEN 'Medium'
        WHEN salary < 50000 THEN 'Low'
    END as salarycategory
FROM employees;
```

Q50. 50 day challenge day start from 11

```
-- Solution
SELECT
```

Q51.

```
-- Solution
SELECT
```