Stratascratch SQL Questions

You're tasked with analyzing a Spotify-like dataset that captures user listening habits.
For each user, calculate the total listening time and the count of unique songs they've listened to. In the database duration values are displayed in seconds. Round the total listening duration to the nearest whole minute.

The output should contain three columns: 'user_id', 'total_listen_duration', and 'unique_song_count'.

```
SELECT
   user_id,
   ROUND(SUM(listen_duration) / 60) AS total_listen_duration,
   COUNT(DISTINCT song_id) AS unique_song_count
FROM
   listening_habits
GROUP BY
   user_id;
```

Write a query that calculates the difference between the highest salaries found in the marketing and engineering departments. Output just the absolute difference in salaries.

```
with cte as (
select dd.department, max(de.salary) as salary
from db_employee de
inner join db_dept dd
on de.department_id = dd.id
where dd.department in ('marketing',  'engineering' )
group by dd.department
)
select max(salary) - min(salary) as salary_difference
from cte;
```

We have a table with employees and their salaries, however, some of the records are old and contain outdated salary information. Find the current salary of each employee assuming that salaries increase each year. Output their id, first name, last name, department ID, and current salary. Order your list by employee ID in ascending order.

```
select id, first_name, last_name,department_id, max(salary)
from ms_employee_salary
group by id, first_name, last_name, department_id
order by id ASC
```

Find the last time each bike was in use. Output both the bike number and the date-timestamp of the bike's last use (i.e., the date-time the bike was returned). Order the results by bikes that were most recently used.

```
SELECT bike_number,
      MAX(end_time) AS last_used
FROM dc_bikeshare_q1_2012
GROUP BY bike_number
ORDER BY last_used DESC;
```

Find the number of rows for each review score earned by 'Hotel Arena'. Output the hotel name (which should be 'Hotel Arena'), review score along with the corresponding number of rows with that score for the specified hotel.

```
select hotel_name, reviewer_score, count(*)
from hotel_reviews
where hotel_name = 'Hotel Arena'
group by hotel_name, reviewer_score;
```

Count the number of movies that Abigail Breslin was nominated for an oscar.

```
select count(nominee) from oscar_nominees
where nominee like 'Abigail%' ;
```

Find all posts which were reacted to with a heart. For such posts output all columns from facebook_posts table.

```
select distinct fp.* from
facebook_reactions fr
inner join facebook_posts fp
on fr.post_id = fp.post_id
where fr.reaction = 'heart'  ;
```

Based on the above, find the average popularity of the Hack per office location.
Output the location along with the average popularity.

```
select avg(fb.popularity), emp.location
from
facebook_employees as emp
inner join
facebook_hack_survey as fb
```

```
on emp.id = fb.employee_id
group by  emp.location;
```

Find all Lyft drivers who earn either equal to or less than 30k USD or equal to or more than 70k USD.

```
select * from lyft_drivers
where yearly_salary <= 30000 or  yearly_salary >= 70000;
```

Find how many times each artist appeared on the Spotify ranking list
Output the artist name along with the corresponding number of occurrences.
Order records by the number of occurrences in descending order.

```
select artist, count(*) as count_number
from spotify_worldwide_daily_song_ranking
group by artist
order by count_number desc;
```

Find the base pay for Police Captains.
Output the employee name along with the corresponding base pay.

```
select employeename, basepay
 from sf_public_salaries
 where jobtitle like '%Police%' ;
```

Find libraries who haven't provided the email address in circulation year 2016 but their notice preference definition is set to email.

```
select distinct home_library_code
from library_usage
where circulation_active_year = 2016
and notice_preference_definition = 'email'
and provided_email_address  = False ;
```

Compare each employee's salary with the average salary of the corresponding department.
Output the department, first name, and salary of employees along with the average salary of that department.

```
select department,
first_name,
salary,
avg(salary) over ( partition by department) as sal_dept
from employee
group by department,first_name, salary ;
```

Consider Jill and Eva as first names of customers.
Output the order date, details and cost along with the first name.

```
select order_date, order_details, total_order_cost, first_name
from customers
join orders
on customers.id = orders.cust_id
where first_name in ('Jill','Eva')  ;
```

Find the details of each customer regardless of whether the customer made an order. Output the customer's first name, last name, and the city along with the order details.
Sort records based on the customer's first name and the order details in ascending order.

```
select  first_name, last_name, city, order_details
from customers
left join orders
on customers.id = orders.cust_id
order by first_name, order_details  asc;
```

find the number of workers by department who joined in or after April.

```
select department, count(worker_id) number_of_workers
from worker
where month(joining_date) >=4
group by department
order by number_of_workers desc ;
```

Find the number of employees working in the Admin department that joined in April or later.

```
select count(worker_id) number_of_workers
from worker
where month(joining_date) >=4
and department = 'Admin'
group by department
order by number_of_workers desc ;
```

Find the activity date and the pe_description of facilities with the name 'STREET CHURROS' and with a score of less than 95 points.

```
select activity_date, pe_description
from los_angeles_restaurant_health_inspections
where facility_name = 'STREET CHURROS'
and score <95;
```

Find the most profitable company from the financial sector. Output the result along with the continent.

```
SELECT company, continent
FROM forbes_global_2010_2014
WHERE sector = 'Financials'
    AND profits = (SELECT MAX(profits)
                FROM forbes_global_2010_2014
                WHERE sector = 'Financials');
```

Write a query that returns the number of unique users per client per month

```
select count(distinct user_id) as users_num,
month(time_id),
client_id
from fact_events
group by month(time_id),
client_id;
```

Count the number of user events performed by MacBookPro users.

```
select event_name, count(*) as no_of_users
from playbook_events
where device = 'macbook pro'
group by event_name ;
```

Write a query that will calculate the number of shipments per month. The unique key for one shipment is a combination of shipment_id and sub_id. Output the year_month in format YYYY-MM and the number of shipments in that month.

```
select count(concat(shipment_id, sub_id)),
DATE_FORMAT(shipment_date, "%Y-%m") as date_ym
from amazon_shipment
group by date_ym;
```

You have been asked to find the 5 most lucrative products in terms of total revenue for the first half of 2022 (from January to June inclusive).

```
with cte as (
select sum(units_sold * cost_in_dollars ) as revenue,
product_id,
rank() over (order by sum(units_sold * cost_in_dollars ) desc) as rnk
from  online_orders
where month(date) between 1 and 6
group by product_id
)
select product_id, revenue
from cte
where rnk<=5 ;
```

Find the average number of bathrooms and bedrooms for each city's property types. Output the result along with the city name and the property type.

```
select avg(bathrooms) as bathrooms,
      avg(bedrooms) as bedrooms,
      city,
      property_type
from airbnb_search_details
group by   city,
      property_type;
```