

## DataLemur - Average Post Hiatus (Part 1)

### Facebook SQL Interview Question

Given a table of Facebook posts, for each user who posted at least twice in 2021, write a query to find the number of days between each user's first post of the year and last post of the year in the year 2021. Output the user and number of the days between each user's first and last post.

**posts Table:**

Column Name	Type
user_id	integer
post_id	integer
post_content	text
post_date	timestamp

**posts Example Input:**

user_id	post_id	post_content	post_date
151652	599415	Need a hug	07/10/2021 12:00:00
661093	624356	Bed. Class 8-12. Work 12-3. Gym 3-5 or 6. Then class 6-10. Another day that's gonna fly by. I miss my girlfriend	07/29/2021 13:00:00
004239	784254	Happy 4th of July!	07/04/2021 11:00:00
661093	442560	Just going to cry myself to sleep after watching Marley and Me.	07/08/2021 14:00:00
151652	111766	I'm so done with covid - need travelling ASAP!	07/12/2021 19:00:00

### Solution:

```
WITH count_user AS (  
  SELECT user_id, COUNT(post_id) AS post_count  
  FROM posts  
  WHERE post_date < '2022-01-01'  
  GROUP BY user_id  
  HAVING COUNT(post_id) > 1  
)  
  
SELECT p.user_id,  
       EXTRACT(DAY FROM (MAX(p.post_date) - MIN(p.post_date))) AS days_between  
FROM posts p  
INNER JOIN count_user cu ON p.user_id = cu.user_id  
GROUP BY p.user_id  
ORDER BY p.user_id;
```

## DataLemur - Teams Power Users

### Microsoft SQL Interview Question

Write a query to identify the top 2 Power Users who sent the highest number of messages on Microsoft Teams in August 2022. Display the IDs of these 2 users along with the total number of messages they sent. Output the results in descending order based on the count of the messages.

**messages Table:**

Column Name	Type
message_id	integer
sender_id	integer
receiver_id	integer
content	varchar
sent_date	datetime

**messages Example Input:**

message_id	sender_id	receiver_id	content	sent_date
901	3601	4500	You up?	08/03/2022 00:00:00
902	4500	3601	Only if you're buying	08/03/2022 00:00:00
743	3601	8752	Let's take this offline	06/14/2022 00:00:00
922	3601	4500	Get on the call	08/10/2022 00:00:00

### Solution:

```
SELECT sender_id, COUNT(content) AS message_count
FROM messages
WHERE sent_date > '2022-07-31' and sent_date < '2022-09-01'
GROUP BY sender_id
order by message_count desc
LIMIT 2
```

## DataLemur - Duplicate Job Listings

### Linkedin SQL Interview Question

Assume you're given a table containing job postings from various companies on the LinkedIn platform. Write a query to retrieve the count of companies that have posted duplicate job listings.

Definition:

- Duplicate job listings are defined as two job listings within the same company that share identical titles and descriptions.

**job\_listings Table:**

Column Name	Type
job_id	integer
company_id	integer
title	string
description	string

**job\_listings Example Input:**

job_id	company_id	title	description
248	827	Business Analyst	Business analyst evaluates past and current business data with the primary goal of improving decision-making processes within organizations.
149	845	Business Analyst	Business analyst evaluates past and current business data with the primary goal of improving decision-making processes within organizations.
945	345	Data Analyst	Data analyst reviews data to identify key insights into a business's customers and ways the data can be used to solve problems.
164	345	Data Analyst	Data analyst reviews data to identify key insights into a business's customers and ways the data can be used to solve problems.
172	244	Data Engineer	Data engineer works in a variety of settings to build systems that collect, manage, and convert raw data into usable information for data scientists and business analysts to interpret.

### Solution:

```
WITH counted_num AS
(SELECT
  company_id,
  description,
  COUNT(job_id) AS job_count
FROM job_listings
GROUP BY company_id, description)
select count(job_count)
from counted_num
where job_count>1
```