

Day44 - March 20th 2024

1. Solved one leetcode problem on stack concepts : [DSA\\_03/20/2024\\_Stacks](#)
2. PySpark theory learning is in progress

The screenshot shows a desktop environment with two windows open. The left window is a YouTube video titled "spark fundamental (Theory)" by MANISH KUMAR. The video content shows a blackboard with the following text:

```
df = spark.read.format("csv")\
    .option(...)\
    .load("Path")
```

Below the code, it says "Same more transformation". The video player shows a progress bar at 5:26 / 32:43. The right window is a Google Docs document titled "SparkTH Day7 03/20/24". The document contains a diagram of a 3x3 grid of boxes. The top row has three empty boxes. The middle row has three boxes, with the first two containing diagonal lines and the third containing a single vertical line. The bottom row has three empty boxes. Arrows point from the text "SOI" to the first and third boxes of the middle row. Below the diagram, the document contains the following text:

6. Now we know where the caching is happening
7. Actually caching is an optimization technique..which stores an intermediate result
8. What is an intermediate result?

The desktop taskbar at the bottom shows the system clock as 8:31 PM on 3/20/2024, with a temperature of 9°C and a clear sky.

### 3. Ended my day by solving 2 interview questions from datalemur

**DataLemur** Practice Interview Questions Learn SQL Get 1:1 Coaching kaushik varma KV

[Back to questions](#)

## User's Third Transaction [Uber SQL Interview Question]

Description Solution Discussion Submissions

Medium Uber [Share on Twitter](#) [Share on LinkedIn](#)

This is the same question as problem #11 in the SQL Chapter of [Ace the Data Science Interview](#).

Assume you are given the table below on Uber transactions made by users. Write a query to obtain the third transaction of every user. Output the user id, spend and transaction date.

**transactions Table:**

Column Name	Type
user_id	integer
spend	decimal
transaction_date	timestamp

**transactions Example Input:**

user_id	spend	transaction_date
111	100.50	01/08/2022 12:00:00
111	55.00	01/10/2022 12:00:00
121	36.00	01/18/2022 12:00:00
145	24.99	01/26/2022 12:00:00
111	89.60	02/05/2022 12:00:00

```
1 --SELECT * FROM transactions;
2
3 with cte as(
4 select *,row_number() over(partition by user_id order by transaction_date) as rn
5 from transactions)
6 select user_id,spend,transaction_date
7 from cte
8 where rn = 3
9
10 /* Explanation :
11 Step1 : To find the 3rd transaction of a user, used row_number partitioned by user
12         id over transaction date.
13 Step2 : Just select rows with rn = 3 ..which gives us the 3rd transaction of a user*/
```

PostgreSQL 14 [Run Code](#) [Submit](#)

**Output**

user_id	spend	transaction_date
111	89.60	02/05/2022 12:00:00
121	67.90	04/03/2022 12:00:00
263	100.00	07/12/2022 12:00:00

**DataLemur** Practice Interview Questions Learn SQL Get 1:1 Coaching kaushik varma KV

[Back to questions](#)

## User's Third Transaction [Uber SQL Interview Question]

Description Solution Discussion Submissions

Accepted

Congrats 🎉 - Share this problem, and your solution, on LinkedIn or Twitter!

[Share on Twitter](#) [Share on LinkedIn](#)

In your post, don't forget to tag Nick Singh, so that he can comment on and share your post with his audience of 150k+ followers on [LinkedIn](#) and 25k+ followers on [Twitter](#) (which will give your post and profile more visibility)!

**Output**

user_id	spend	transaction_date
111	89.60	02/05/2022 12:00:00
121	67.90	04/03/2022 12:00:00
263	100.00	07/12/2022 12:00:00

**Expected**

user_id	spend	transaction_date
111	89.60	02/05/2022 12:00:00

```
1 --SELECT * FROM transactions;
2
3 with cte as(
4 select *,row_number() over(partition by user_id order by transaction_date) as rn
5 from transactions)
6 select user_id,spend,transaction_date
7 from cte
8 where rn = 3
9
10 /* Explanation :
11 Step1 : To find the 3rd transaction of a user, used row_number partitioned by user
12         id over transaction date.
13 Step2 : Just select rows with rn = 3 ..which gives us the 3rd transaction of a user*/
```

PostgreSQL 14 [Run Code](#) [Submit](#)

**Output**

user_id	spend	transaction_date
111	89.60	02/05/2022 12:00:00
121	67.90	04/03/2022 12:00:00
263	100.00	07/12/2022 12:00:00

