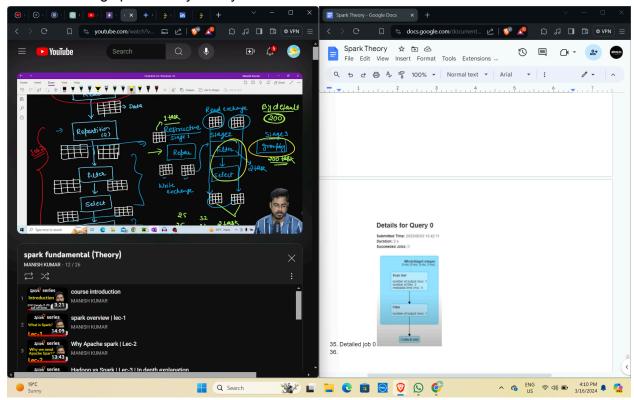Day40 - March 16th 2024

1. Started my day as usual
2. Marketed my profile for data engineering on linkedin and dice
3. Solved one leetcode medium problem
4. Started learning spark theory from youtube

5. Please find my learning doc :
   https://docs.google.com/document/d/1vtgtCNflkiyHrPocS1pdTSaT2FpQFQBiKt_0IjkNISc/edit?usp=sharing
6. Ended my day by solving HARD SQL question from datalemur

datalemur.com/questions/yoy-growth-rate

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

‹ Back to questions

Y-on-Y Growth Rate [Wayfair SQL Interview Question]

| Description | Solution | Discussion | Submissions |

product_id	integer
spend	decimal
transaction_date	datetime

user_transactions Example Input:

transaction_id	product_id	spend	transaction_date
1341	123424	1500.60	12/31/2019 12:00:00
1423	123424	1000.20	12/31/2020 12:00:00
1623	123424	1246.44	12/31/2021 12:00:00
1322	123424	2145.32	12/31/2022 12:00:00

Example Output:

year	product_id	curr_year_spend	prev_year_spend	yoy_rate
2019	123424	1500.60	NULL	NULL
2020	123424	1000.20	1500.60	-33.35
2021	123424	1246.44	1000.20	24.62
2022	123424	2145.32	1246.44	72.12

Explanation:

Product ID 123424 is analyzed for multiple years: 2019, 2020, 2021, and 2022.

```
1  --SELECT *,EXTRACT(year from transaction_date) FROM user_transactions;
2
3  -- select Extract(year from transaction_date) as year,product_id,spend as curr_year_spend
4  -- lag(spend,1) over(order by Extract(year from transaction_date)) as prev_year_spend
5  -- from user_transactions
6
7
8  select *,sum(spend) over(PARTITION BY Extract(year from transaction_date) order by year)
9  from user_transactions
```

PostgreSQL 14   Run Code   Submit

Error   ⌄

column "year" does not exist. FYI: in PostgreSQL all identifiers (including column names)
that are not double-quoted are folded to lowercase. So if you're using any uppercase letters
in the names and want to specify them in SQL-query, please double-quote them first. (LINE: 8)

10°C Clear   Q Search   ENG US   10:15 PM 3/16/2024

7.

---

datalemur.com/questions/yoy-growth-rate

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

‹ Back to questions

Y-on-Y Growth Rate [Wayfair SQL Interview Question]

| Description | Solution | Discussion | Submissions |

Accepted

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

🐦 Share on Twitter   in 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

year	product_id	curr_year_spend	prev_year_spend	yoy
2019	123424	1500.60	NULL	NULL
2020	123424	1000.20	1500.60	-33.35
2021	123424	1246.44	1000.20	24.62
2022	123424	2145.32	1246.44	72.12
2019	234412	1800.00	NULL	NULL
2020	234412	1234.00	1800.00	-31.44

```
1  --SELECT *,EXTRACT(year from transaction_date) FROM user_transactions;
2
3  -- select Extract(year from transaction_date) as year,product_id,spend as curr_year_s
4  -- lag(spend,1) over(order by Extract(year from transaction_date)) as prev_year_spend
5  -- from user_transactions
6
7  -- with cte as(
8  -- select *,sum(spend) over(PARTITION BY Extract(year from transaction_date)) as curr
9  -- row_number() over(partition by Extract(year from transaction_date) order by produc
10 -- from user_transactions
11 -- select *,lag(curr_year_spend,1) over(order by Extract(year from transaction_date))
12 -- where rn = 1
13
14 with cte as(
15 select Extract(year from transaction_date) as year,product_id,spend as curr_year_spend
16 select *, round((curr_year_spend-prev_year_spend)*100/prev_year_spend,2) as yoy from
```

PostgreSQL 14   Run Code   Submit

Output   ⌄

year	product_id	curr_year_spend	prev_year_spend	yoy
2019	123424	1500.60	NULL	NULL
2020	123424	1000.20	1500.60	-33.35
2021	123424	1246.44	1000.20	24.62
2022	123424	2145.32	1246.44	72.12

9°C Partly cloudy   Q Search   ENG US   10:39 PM 3/16/2024