Twitter SQL Interview Question

Datalemur

By Joshua kab

Description

Assume you're given a table Twitter tweet data, 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 bucket.

In other words, group the users by the number of tweets they posted in 2022 and count the number of users in each group.

TABLE OVERVIEW

tweets **Example Input:**

tweet_id	user_id	msg	tweet_date
214252	111	Am considering taking Tesla private at $420. Funding secured.	12/30/2021 00:00:00
739252	111	Despite the constant negative press covfefe	01/01/2022 00:00:00
846402	111	Following @NickSinghTech on Twitter changed my life!	02/14/2022 00:00:00
241425	254	If the salary is so competitive why won't you tell me what it is?	03/01/2022 00:00:00
231574	148	I no longer have a manager. I can't be managed	03/23/2022 00:00:00

SOLUTION

```
with cte as (SELECT user_id as tweet_bucket,
    EXTRACT(Year from tweet_date) as year,
    COUNT(user_id) as user_num
FROM tweets
GROUP BY user_id, year)

SELECT tweet_bucket,user_num
FROM cte
WHERE year = 2022
```

RESULT

Output

tweet_bucket	user_num
111	2
148	1
254	1

THANKS…