DataLemur - Data Science Skills

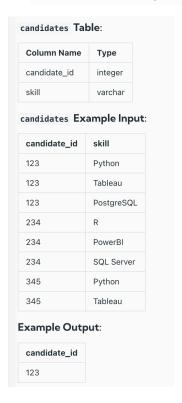
LinkedIn SQL Interview Question

Given a table of candidates and their skills, you're tasked with finding the candidates best suited for an open Data Science job. You want to find candidates who are proficient in Python, Tableau, and PostgreSQL.

Write a query to list the candidates who possess all of the required skills for the job. Sort the output by candidate ID in ascending order.

Assumption:

There are no duplicates in the candidates table.



Solution:

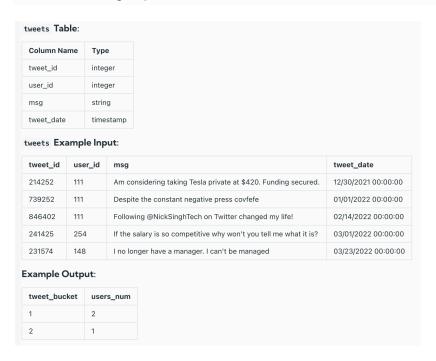
```
WITH total_skill AS
(select candidate_id , count(skill) as total_skill from candidates
where skill in ('Python', 'Tableau', 'PostgreSQL')
group by candidate_id
)

select candidate_id from total_skill
where total_skill >=3
```

DataLemur - **Histogram of Tweets**Twitter SQL Interview Question

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.



Solution:

```
WITH total_tweets AS (
SELECT
user_id,
COUNT(tweet_id) AS tweet_count_per_user
FROM tweets
WHERE tweet_date BETWEEN '2022-01-01'
AND '2022-12-31'
GROUP BY user_id)
```

SELECT

```
tweet_count_per_user AS tweet_bucket,
COUNT(user_id) AS users_num
FROM total_tweets
GROUP BY tweet_count_per_user;
```

Datalemur - Page With No Likes Facebook SQL Interview Question

Assume you're given two tables containing data about Facebook Pages and their respective likes (as in "Like a Facebook Page").

Write a query to return the IDs of the Facebook pages that have zero likes. The output should be sorted in ascending order based on the page IDs.

pages Tak	ole:					
Column Name		Тур	е			
page_id		inte	ger			
page_name		vard	char			
pages Example Input:						
page_id	page_name					
20001	SQL Solutions					
20045	Brain Exercises					
20701	Tips for Data Analysts		alysts			
page_likes Table :						
Column Name		Тур	е			
user_id	user_id in		ger			
page_id		integer				
liked_date		datetime				
page_likes Example Input:						
user_id	page	page_id		d_date		
111	20001		04/0	08/2022 00:00:00		
121	20045		03/1	2/2022 00:00:00		
156	20001		07/2	25/2022 00:00:00		

Solution:

SELECT page_id FROM pages WHERE page_id not in (select DISTINCT page_id from page_likes)

Datalemur Unfinished Parts

Tesla SQL Interview Question

Tesla is investigating production bottlenecks and they need your help to extract the relevant data. Write a query to determine which parts have begun the assembly process but are not yet finished.

Assumptions:

- parts_assembly table contains all parts currently in production, each at varying stages of the assembly process.
- An unfinished part is one that lacks a finish date.

This question is straightforward, so let's approach it with simplicity in both thinking and solution.

Effective April 11th 2023, the problem statement and assumptions were updated to enhance clarity.

Column Name		Туре	
part		string	
finish_date		datetime	
assembly_step		integer	
part	embly Example Inpe		assembly_step
battery	01/22/2022 00:00:0		1
	02/22/2022 00:00:		
oattery	02/22	2/2022 00:00:00	2
		2/2022 00:00:00	3
battery	03/22		<u> </u>
battery	03/22	2/2022 00:00:00	3
battery battery bumper bumper bumper	03/22	2/2022 00:00:00	3

Solution:

SELECT part, assembly_step FROM parts_assembly where finish_date is null

Datalemur: Laptop vs. Mobile Viewership New York Times SQL Interview Question

Assume you're given the table on user viewership categorised by device type where the three types are laptop, tablet, and phone.

Write a query that calculates the total viewership for laptops and mobile devices where mobile is defined as the sum of tablet and phone viewership. Output the total viewership for laptops as laptop reviews and the total viewership for mobile devices as mobile views.

viewership Table

Column Name	Туре
user_id	integer
device_type	string ('laptop', 'tablet', 'phone')
view_time	timestamp

viewership Example Input

user_id	device_type	view_time
123	tablet	01/02/2022 00:00:00
125	laptop	01/07/2022 00:00:00
128	laptop	02/09/2022 00:00:00
129	phone	02/09/2022 00:00:00
145	tablet	02/24/2022 00:00:00

Solution:

select

COUNT(CASE

WHEN device_type = 'laptop' THEN 1

ELSE NULL

END) AS laptop_views,

COUNT(CASE

WHEN device_type = 'tablet' or device_type = 'phone' THEN 1

ELSE NULL

END) AS mobile views

from viewership