SQL Interview Question #1

Let’s say you have two SQL tables: `authors` and `books`.

The `authors` dataset has 1M+ rows; here’s the first six rows:

author_name	**book_name**
author_1	book_1
author_1	book_2
author_2	book_3
author_2	book_4
author_2	book_5
author_3	book_6
…	…

The `books` dataset also has 1M+ rows and here’s the first six:

book_name	**sold_copies**
book_1	1000
book_2	1500
book_3	34000
book_4	29000
book_5	40000
book_6	4400
…	…

Create an SQL query that shows the TOP 3 authors who sold the most books in total!

Answer:

```
SELECT author_name
FROM authors, books
WHERE authors.book_name = books.book_name
GROUP BY author_name
ORDER BY sum(sold_copies) DESC
LIMIT 3
```

SQL Interview Question #2

You work for a startup that makes an online presentation software. You have an event log that records every time a user inserted an image into a presentation. (One user can insert many images.) The `event_log` SQL table looks like this:

user_id	event_date_time
7494212	1535308430
7494212	1535308433
1475185	1535308444
6946725	1535308475
6946725	1535308476
6946725	1535308477
…	…

…and it has over one billion rows.

Note: The `event_date_time` *column's format is an epoch timestamp*

Write an SQL query to find out how many users inserted more than 1000 but less than 2000 images in their presentations!

Answer:

```
SELECT count(user_id)
FROM event_log
GROUP BY user_id
HAVING count(event_data_time) >1000 and
count(event_data_time) < 2000
```

SQL Interview Question #3

You have two SQL tables! The first one is called `employees` and it contains the employee names, the unique employee ids and the department names of a company. Sample:

department_name	employee_id	employee_name
Sales	123	John Doe
Sales	211	Jane Smith
HR	556	Billy Bob
Sales	711	Robert Hayek
Marketing	235	Edward Jorgson
Marketing	236	Christine Packard
…	…	…

The second one is named `salaries`. It holds the same employee names and the same employee ids – and the salaries for each employee. Sample:

salary	employee_id	employee_name
500	123	John Doe
600	211	Jane Smith
1000	556	Billy Bob
400	711	Robert Hayek
1200	235	Edward Jorgson
200	236	Christine Packard
…	…	…

The company has 546 employees, so both tables have 546 rows.

Print every department where the average salary per employee is lower than $500!

Answer:

```
SELECT department_name
FROM employees, salaries
WHERE employees.employee_id = salaries.employee_id
and employees.employee_name = salaries. employee_name
GROUP BY department_name
HAVING avg(salary) < 500
```