### --> Question 2.1 What are the Top 25 schools (.edu domains)?

#### Ans:

**SELECT** 

email\_domain,

COUNT(\*)

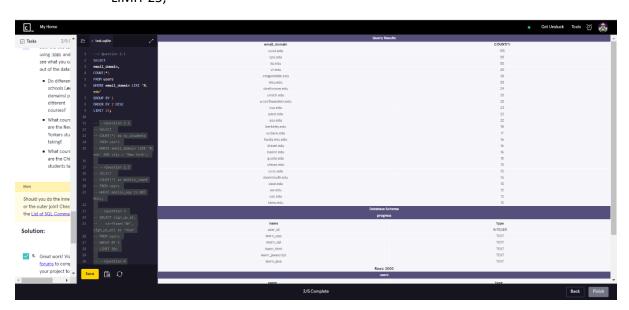
FROM users

WHERE email\_domain LIKE '%.edu'

GROUP BY 1

**ORDER BY 2 DESC** 

LIMIT 25;



### -->Question 2.2 How many .edu learners are located in New York?

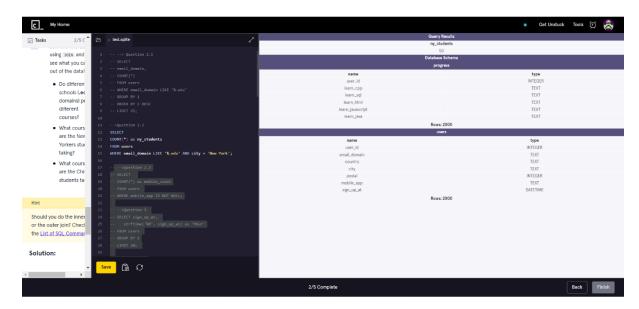
### Ans:

**SELECT** 

COUNT(\*) as ny\_students

FROM users

WHERE email\_domain LIKE '%.edu' AND city = 'New York';



-->Question 2.3 The mobile\_app column contains either mobile-user or NULL. How many of these Codecademy learners are using the mobile app?

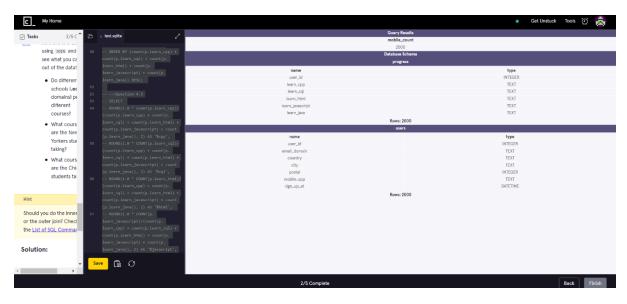
### Ans:

**SELECT** 

COUNT(\*) as mobile\_count

FROM users

WHERE mobile\_app IS NOT NULL;



-->Question 3

The data type of the sign\_up\_at column is DATETIME. It is for storing a date/time value in the database.

Notice that the values are formatted like:

2015-01-01 18:33:52

So the format is:

YYYY-MM-DD HH:MM:SS

SQLite comes with a strftime() function - a very powerful function that allows you to return a formatted date.

It takes two arguments:strftime(format, column)

Let's test this function out:

SELECT sign\_up\_at,

strftime('%S', sign\_up\_at)

**FROM** users

**GROUP BY 1** 

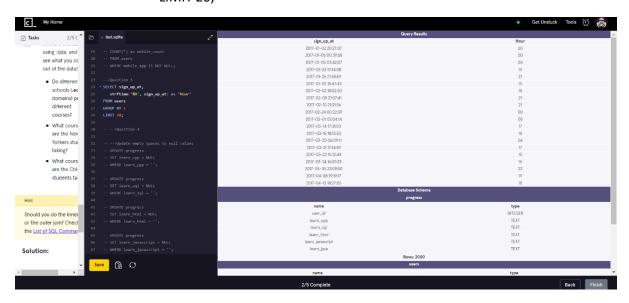
LIMIT 20; Now, using this function, query for the sign up counts for each hour.

Ans:

SELECT sign\_up\_at,
strftime('%H', sign\_up\_at) as 'Hour'
FROM users

LIMIT 20;

**GROUP BY 1** 



-->Question 4 Join the two tables using JOIN and then see what you can dig out of the data!

Ans:

-->Update empty spaces to null values

Ans:

**SELECT** 

email\_domain,

```
UPDATE progress
                       SET learn_cpp = NULL
                       WHERE learn_cpp = ";
                       UPDATE progress
                       SET learn_sql = NULL
                       WHERE learn_sql = ";
                       UPDATE progress
                       SET learn_html = NULL
                       WHERE learn_html = ";
                       UPDATE progress
                       SET learn_javascript = NULL
                       WHERE learn_javascript = ";
                       UPDATE progress
                       SET learn_java = NULL
                       WHERE learn_java = ";
-->Question 4.1 Do different schools (.edu domains) prefer different courses?
ROUND(1.0 * count(p.learn_cpp)/(count(p.learn_cpp) + count(p.learn_sql) + count(p.learn_html) +
count(p.learn_javascript) + count(p.learn_java)), 2) AS '%cpp',
ROUND(1.0 * COUNT(p.learn_sql)/(count(p.learn_cpp) + count(p.learn_sql) + count(p.learn_html) +
count(p.learn_javascript) + count(p.learn_java)), 2) AS '%sql',
```

ROUND(1.0 \* COUNT(p.learn\_html)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html)

+ count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%html',

ROUND(1.0 \* COUNT(p.learn\_javascript)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%javasript',

ROUND(1.0 \* COUNT(p.learn\_java)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%java'

FROM users u

JOIN progress p

ON u.user\_id = p.user\_id

WHERE email\_domain LIKE '%.edu'

**GROUP BY 1** 

ORDER BY (count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)) DESC

#### LIMIT 10;



#### -- Question 4.2 What courses are the New Yorkers students taking?

#### Ans:

**SELECT** 

ROUND(1.0 \* count(p.learn\_cpp)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%cpp',

ROUND(1.0 \* COUNT(p.learn\_sql)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%sql',

ROUND(1.0 \* COUNT(p.learn\_html)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%html',

ROUND(1.0 \* COUNT(p.learn\_javascript)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_javas)), 2) AS '%javasript',

ROUND(1.0 \* COUNT(p.learn\_java)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%java'

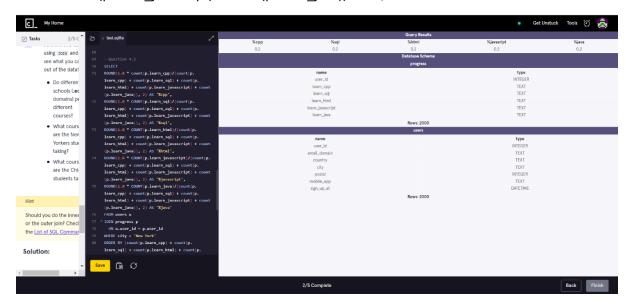
FROM users u

JOIN progress p

ON u.user\_id = p.user\_id

WHERE city = 'New York'

ORDER BY (count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)) DESC;



#### -->Question 4.3 What courses are the Chicago students taking?

#### Ans:

#### **SELECT**

ROUND(1.0 \* count(p.learn\_cpp)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%cpp',

ROUND(1.0 \* COUNT(p.learn\_sql)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%sql',

ROUND(1.0 \* COUNT(p.learn\_html)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%html',

ROUND(1.0 \* COUNT(p.learn\_javascript)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%javasript',

ROUND(1.0 \* COUNT(p.learn\_java)/(count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)), 2) AS '%java'

FROM users u

JOIN progress p

ON u.user\_id = p.user\_id

WHERE city = 'New York'

ORDER BY (count(p.learn\_cpp) + count(p.learn\_sql) + count(p.learn\_html) + count(p.learn\_javascript) + count(p.learn\_java)) DESC;

