# RDB Assessment

Welcome to the course assessment for the RDB module!

In this session, you will complete the RDB project.

By the end of this session, you will be able to:

● Apply SQL techniques for querying, aggregating and joining data.

● Solve the given challenges using SQL.

Put your Relational Databases development skills to the test! Use your knowledge of SQL to analyse learners data.

After completing it, make sure you come back to complete a reflection.

**Directions:**

● **Download** the word doc and **send** the word doc (attach to the email) to your instructors. (Do not require to upload to Github)

[genfsd2021@gmail.com](mailto:genfsd2021@gmail.com)

[awbingxian@gmail.com](mailto:awbingxian@gmail.com)

● **Reflect** on the following questions.

**Questions:**

● **What did you like about this project?** *Nothing.*

● **What did you struggle with in this project?** *Time and multiple subqueries.*

● **What would make your experience with this assessment better?** *Not having to print and paste in the result would be great. No time limitation would be perfect.*

Deadline: 28 April 2023, 330pm

### Task: Use your knowledge of SQL and analyze some mockup learners data.

### Download this document and place your SQL solution (in text) and the printscreen of the result outputs in the respective question area.

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_-

There are two tables provided:

**users table: (There should be 2000 records)**

* user\_id
* email\_domain
* Country
* City
* postal
* Mobile\_app
* sign\_up\_at

**progress table: (There should be 2000 records)**

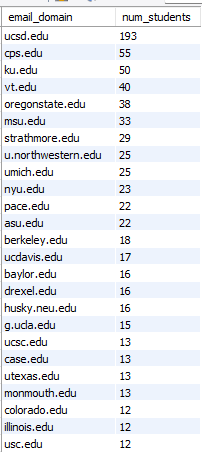
* user\_id
* learn\_cpp
* learn\_sql
* learn\_html
* learn\_javascript
* learn\_java

### 

### **Answer the following questions. You are required to paste your SQL scripts and the printscreen of the result outputs to the respective questions.**

1. Use your knowledge of queries and aggregate functions to get to know the data:
   1. What are the Top 25 schools with the most number of students in descending order on **.edu** domains? Please filter by school email domain and the number of students. (*Hint: COUNT, GROUP BY, LIMIT*)

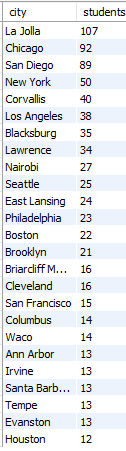
**Result:**

****

|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here for Qns 1a.*  SELECT email\_domain, COUNT(user\_id) AS num\_students  FROM users  GROUP BY 1  ORDER BY 2 DESC  LIMIT 25; |

* 1. List out all the cities with the number of students from the respective cities in descending order of the number of students. (Print Screen the top 25 records)

**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here for Qns 1b.*  SELECT city, COUNT(user\_id) AS students  FROM users  GROUP BY 1  ORDER BY 2 DESC  LIMIT 25; |

* 1. How many students are located in New York?

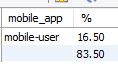
**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here for Qns 1c.*  SELECT city, COUNT(user\_id) AS students  FROM users  WHERE city = "New York"  GROUP BY 1; |

* 1. The mobile\_app column contains either mobile-user or empty. How many of these students are using the mobile app and how many are not? Please show the results in percentages (%). (*Hint:* [*ROUND*](https://www.w3schools.com/sql/func_mysql_round.asp))

**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1d.*  SELECT mobile\_app,  ROUND(COUNT(user\_id) / (SELECT COUNT(user\_id) FROM users) \* 100, 2) AS "%"  FROM users  GROUP BY 1; |

* 1. How many students have completed sql from ALL Schools?

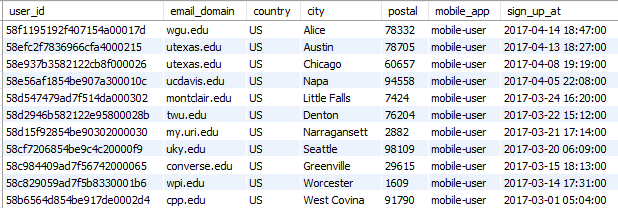
**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1e.*  SELECT COUNT(user\_id) AS "students completed sql"  FROM progress  WHERE learn\_sql = "completed"; |

* 1. List out all students’ details with the sign up date from 1st of March 2017 to 15th April 2017.

**Result:**

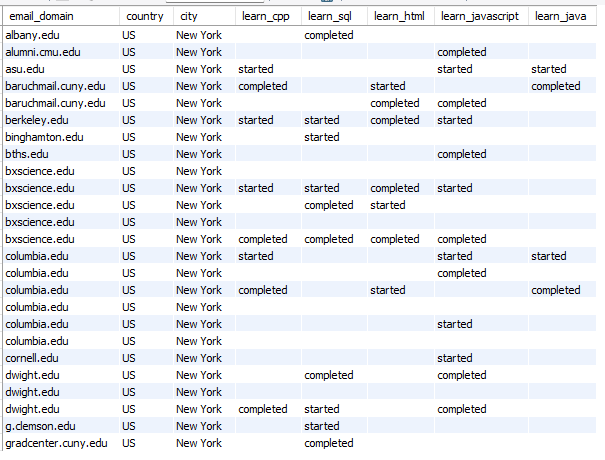


|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1f.*  -- modify 'sign\_up\_at' column to DATETIME format first  ALTER TABLE users  MODIFY sign\_up\_at DATETIME;  -- query based on required conditions  SELECT \*  FROM users  WHERE sign\_up\_at > '2017-03-01'  AND sign\_up\_at < '2017-04-15'; |

1. Join the two tables using JOIN and then see what you can dig out of the data!
   1. What courses are the New Yorkers students taking and have taken (both started and completed)? (List according to ascending order of email\_domain)

(Print Screen the top 25 records)

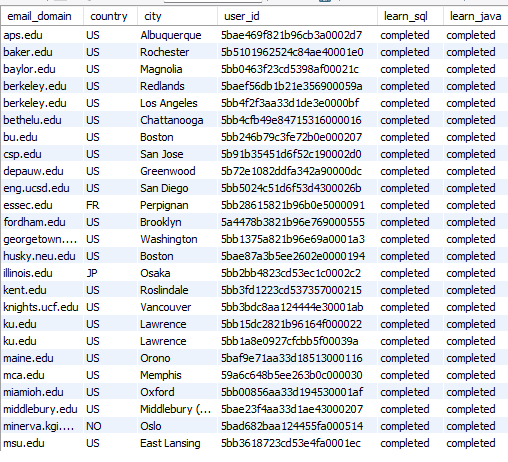
**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2a.*  SELECT users.email\_domain,  users.country,  users.city,  progress.learn\_cpp,  progress.learn\_sql,  progress.learn\_html,  progress.learn\_javascript,  progress.learn\_java  FROM users, progress  WHERE users.user\_id = progress.user\_id  AND users.city = "New York"  ORDER BY 1  LIMIT 25; |

* 1. List the details of the students completed sql and java from their respective Schools (**.edu** domains) (List according to ascending order of email\_domain) (Print Screen the top 25 records)

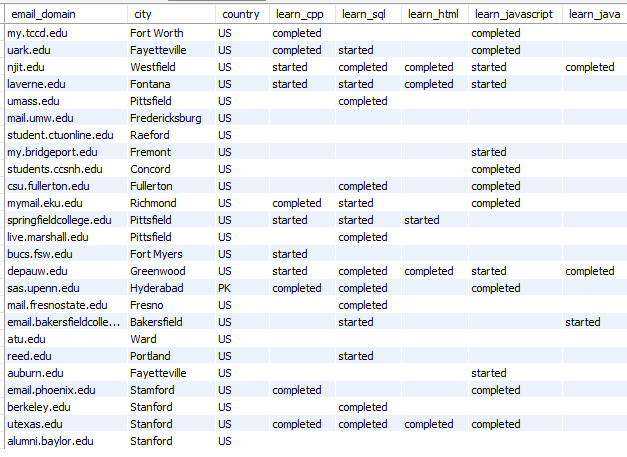
**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2b.*  SELECT users.email\_domain,  users.country,  users.city,  users.user\_id,  progress.learn\_sql,  progress.learn\_java  FROM users, progress  WHERE users.user\_id = progress.user\_id  AND progress.learn\_sql = "completed"  AND progress.learn\_java = "completed"  ORDER BY 1  LIMIT 25; |

* 1. List the details of the students with their modules progress in the City that starts with ‘F’ or the City that ends with ‘D’. (Print Screen the top 25 records)

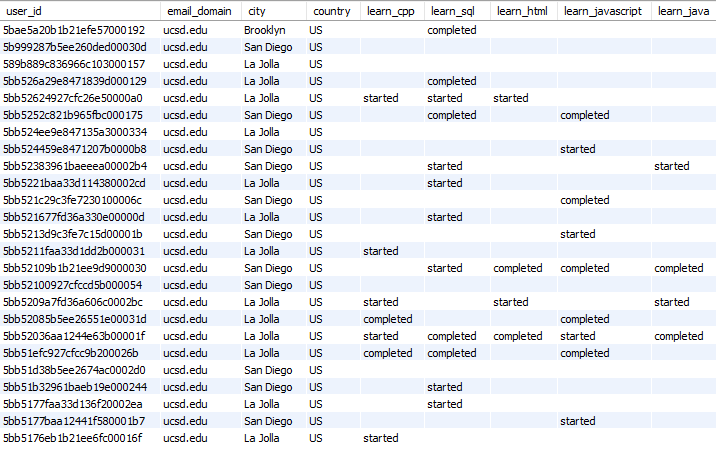
**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2c.*  SELECT users.email\_domain,  users.city,  users.country,  progress.learn\_cpp,  progress.learn\_sql,  progress.learn\_html,  progress.learn\_javascript,  progress.learn\_java  FROM users  INNER JOIN progress  ON users.user\_id = progress.user\_id  WHERE users.city LIKE "F%" || users.city LIKE "%D"  LIMIT 25; |

* 1. List the details of the students taking different courses from the School with the most number of students. (Note: You are not supposed to use the answer derived from Question 1a) (Print Screen the top 25 records) *(Hint:* [*Subquery*](https://www.mysqltutorial.org/mysql-subquery/%20%20%20%20%20payments%20WHERE%20%20%20%20%20amount%20=%20(SELECT%20MAX(amount)%20FROM%20payments);))

**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2d.*  SELECT users.user\_id,  users.email\_domain,  users.city,  users.country,  progress.learn\_cpp,  progress.learn\_sql,  progress.learn\_html,  progress.learn\_javascript,  progress.learn\_java  FROM users, progress  WHERE users.user\_id = progress.user\_id  AND users.email\_domain = (SELECT email\_domain  FROM users  GROUP BY email\_domain  HAVING COUNT(user\_id) = (SELECT MAX(num\_students)  FROM (  SELECT COUNT(user\_id) AS num\_students  FROM users  GROUP BY email\_domain) AS t))  LIMIT 25; |

* 1. Which module is most popular among the students from the School with the most number of students (Top School)? And which module is the least popular among the students? (Note: You are not supposed to use the answer derived from Question 1a)

(*Hint:* [*Count(If)*](https://thispointer.com/count-with-if-condition-in-mysql-query/))

*Please show the number of students (with “started” and “completed”) of the 5 modules from the Top School in the Result Grid, and write in the answer box which module is the most popular and which module is the least popular.*

**Result:**



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2e.*  SELECT COUNT(IF(progress.learn\_cpp != '', progress.user\_id, NULL)) AS learn\_cpp,  COUNT(IF(progress.learn\_sql != '', progress.user\_id, NULL)) AS learn\_sql,  COUNT(IF(progress.learn\_html != '', progress.user\_id, NULL)) AS learn\_html,  COUNT(IF(progress.learn\_javascript != '', progress.user\_id, NULL)) AS learn\_javascript,  COUNT(IF(progress.learn\_java != '', progress.user\_id, NULL)) AS learn\_java  FROM users, progress  WHERE users.user\_id = progress.user\_id  AND users.email\_domain = (SELECT email\_domain  FROM users  GROUP BY email\_domain  HAVING COUNT(user\_id) = (SELECT MAX(num\_students)  FROM (  SELECT COUNT(user\_id) AS num\_students  FROM users  GROUP BY email\_domain) AS t));    -- Most popular: learn\_sql  -- Least popular: learn\_java |