# 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:**

● **Send** the word doc to your instructor.

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

[ernest36912@gmail.com](mailto:ernest36912@gmail.com)

● **Reflect** on the following questions.

**Questions:**

● **What did you like about this project?**

● **What did you struggle with in this project?**

● **What would make your experience with this assessment better?**

**(see final page on reflection)**

Deadline: 6 June 2022, 330pm

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

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

Name: Chua Yew Guan Desmond

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.

**Must use the alias name**

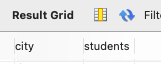
Example of the Result Grid headers:



|  |
| --- |
| *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 email\_domain  ORDER BY num\_students 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)

Example of the Result Grid headers:



|  |
| --- |
| *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 city  ORDER BY students DESC LIMIT 25 |

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

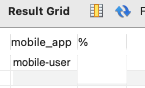
Example of the Result Grid headers:



|  |
| --- |
| *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 email\_domain LIKE '%.edu' AND city LIKE 'NEW YORK'  GROUP BY city  HAVING city = 'New York'; |

* 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 (%).

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1d.*  SELECT mobile\_app,  count(user\_id) AS mobile\_user,  count(user\_id) / (select count(\*) FROM users) \* 100 AS mobile\_app\_percentage  FROM users  GROUP BY mobile\_app; |

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

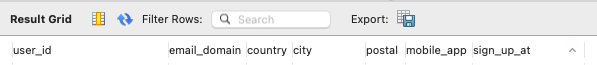
Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1e.*  SELECT learn\_sql AS students\_completed\_sql, count(user\_id) AS students 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.

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 1f.*  SELECT \* FROM users  WHERE sign\_up\_at BETWEEN '2017-03-01' AND '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? (List according to ascending order of email\_domain)

(Print Screen the top 25 records)

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2a.*  SELECT u.email\_domain, u.country, u.city, p.learn\_cpp,p.learn\_sql,p.learn\_html,p.learn\_javascript,p.learn\_java FROM users u  JOIN progress p on u.user\_id=p.user\_id  AND u.city LIKE 'NEW YORK'  ORDER BY u.email\_domain DESC 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)

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2b.*  SELECT u.email\_domain, u.country, u.city, p.learn\_sql, p.learn\_java FROM users u  JOIN progress p on u.user\_id=p.user\_id  WHERE p.learn\_sql = 'completed' AND p.learn\_java = 'completed'  ORDER BY u.email\_domain ASC 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)

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2c.*  SELECT u.user\_id, u.email\_domain, u.country, u.city, p.learn\_cpp,p.learn\_sql,p.learn\_html,p.learn\_javascript,p.learn\_java FROM users u  JOIN progress p on u.user\_id=p.user\_id  WHERE u.city LIKE 'F%' OR u.city LIKE'%d'  ORDER BY u.user\_id 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)

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2d.*  Not sure if answer correct, didn’t put in |

* 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.*

Example of the Result Grid headers:



|  |
| --- |
| *Paste your SQL scripts and the printscreen of the result outputs here Qns 2e.*  Not sure if answer correct, didn’t put in |

**Questions:**

**What did you like about this project?**

Appreciated that the assessment gives a opportunity to practise all the SQL query statements. Assessment is mainly on select statements, would have appreciated more if we are given the opportunity to build database from scratch – DDL, DML rather than being given a ready database to work with. Would like to have initial hands on with the creation of the initial database creation before jumping into writing the SQL query statements.

**What did you struggle with in this project?**

Timing given was good but had to do some self-practice on select statements, select from multiple tables using join and aggregate functions before jumping into completion of the assessment.

**What would make your experience with this assessment better?**

As mentioned, let us build a database so that we can better understand how to create the the tables and instead of do a join on only two tables, let us experience how to join data from multiple tables.

(Student: Desmond Chua)