**Diploma in Information Technology**

**Data Management & Security**

**CA3 Group Assignment**

**Apr 2023 Semester**

|  |  |  |
| --- | --- | --- |
| **No.** | **Name** | **SIM Student No.** |
| L1. | Angeles Micah Josephine Flores | 10240945 |
| M2. | Sumcad Bryce Nicolas Fernandez | 10244185 |
| M3. | Ho Ka Yan Jeslyn | 10241485 |
| M4. | Eng Wei Quan Bill | 10243810 |
| M5. | Liam Israel | 10241506 |

**Date of Submission: 15 May 2023**

**Table of Content**

1 Contribution 3

2 Part 1: User data setup 4

2.1 ER Diagram 4

2.2 Create Table 5

2.3 INSERT DATA 6

3 Part 2: Migrating old post messages 9

3.1 ER Diagram 9

3.2 Create Table 9

3.3 INSERT CSV DATA 10

3.4 Loading CSV data 13

3.5 Usage report 14

3.5.1 Display the amount of topics and posts made by the person. 14

3.5.2 Displays the amount of time each user has spent playing the game. 14

4 Part 3: New features 15

4.1 ER Diagram 15

4.2 SQL query showing the number of post 16

4.3 SQL Query 17

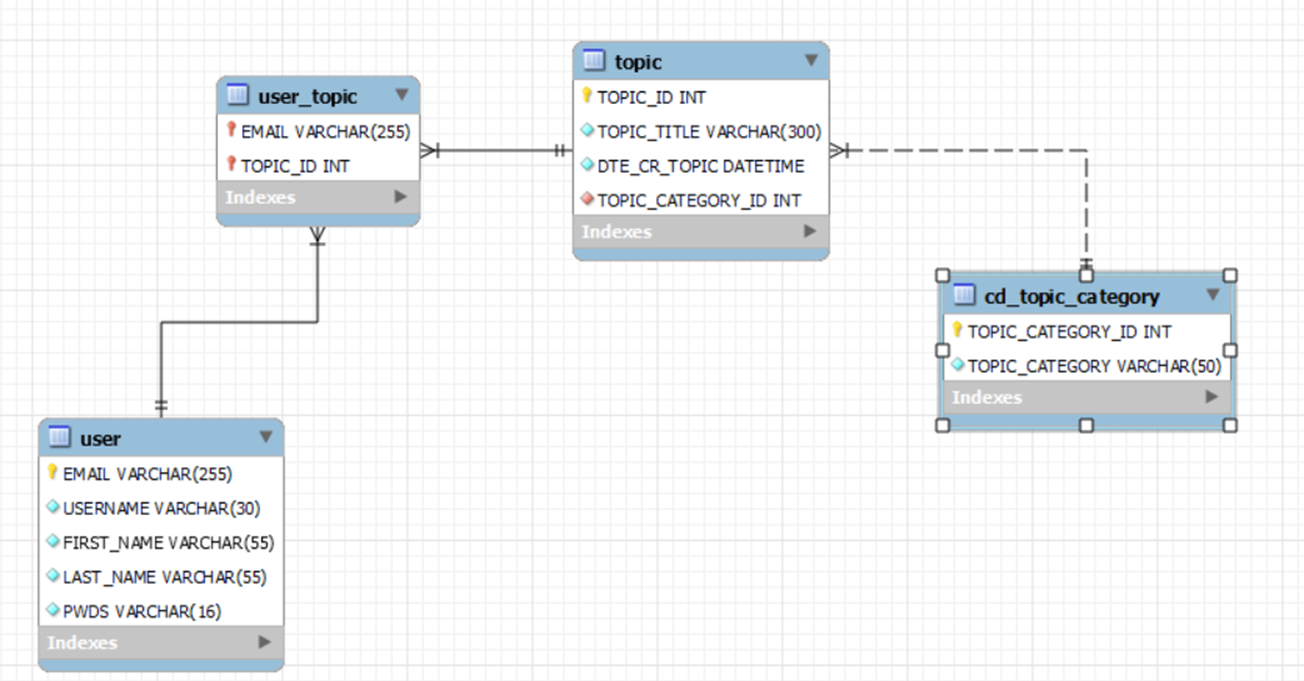
5 Summary 18

# Contribution

|  |  |
| --- | --- |
| Contribution | Members |
| Design database schema | All |
| Create table scripts | Xxx – Jeslyn  Xxx - |
| Insert data scripts | Xxx – Jeslyn  Xxx - |
| Design logic to display the amount of topics and posts made by the person | All |
| Script to display the amount of topics and posts made by the person |  |
| Design logic to ??? |  |
| Script to ??? |  |
| Documentation | Part 1 –  Part 2 –  Part 3 –  Summary -  Overall - |
|  |  |

# Part 1: User data setup

## ER Diagram



## Create Table

### Create SFVI.USER

|  |
| --- |
| CREATE TABLE SFVI.USER (  EMAIL VARCHAR(255) PRIMARY KEY,  USERNAME VARCHAR(30) NOT NULL,  FIRST\_NAME VARCHAR(55) NOT NULL,  LAST\_NAME VARCHAR(55) NOT NULL,  PWDS VARCHAR(16) NOT NULL  ); |

### Create SFVI.CD\_TOPIC\_CATEGORY

|  |
| --- |
| CREATE TABLE SFVI.CD\_TOPIC\_CATEGORY (  TOPIC\_CATEGORY\_ID INT AUTO\_INCREMENT PRIMARY KEY,  TOPIC\_CATEGORY VARCHAR(50) NOT NULL  ); |

### Create SFVI.TOPIC

|  |
| --- |
| CREATE TABLE SFVI.TOPIC (  TOPIC\_ID INT AUTO\_INCREMENT PRIMARY KEY,  TOPIC\_TITLE VARCHAR(300) NOT NULL,  DTE\_CR\_TOPIC DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP,  TOPIC\_CATEGORY\_ID INT(9) NOT NULL,    FOREIGN KEY (TOPIC\_CATEGORY\_ID) REFERENCES SFVI.CD\_TOPIC\_CATEGORY(TOPIC\_CATEGORY\_ID)  ); |

### Create SFVI.USER\_TOPIC

|  |
| --- |
| CREATE TABLE SFVI.USER\_TOPIC (  EMAIL VARCHAR(255),  TOPIC\_ID INT,  PRIMARY KEY (EMAIL,TOPIC\_ID),  FOREIGN KEY (EMAIL) REFERENCES SFVI.USER(EMAIL),  FOREIGN KEY (TOPIC\_ID) REFERENCES SFVI.TOPIC(TOPIC\_ID)  ); |

## INSERT DATA

**Insert sfvi.user**

|  |
| --- |
| INSERT sfvi.user (email, username, first\_name, last\_name, pwds) VALUES ('takanaka@capcom.com','takanaka','Takayuki','Nakayama','dzm6mw185bgibmj3');  INSERT sfvi.user (email, username, first\_name, last\_name, pwds) VALUES ('tsuchiya@capcom.com','tsuchiya','Kazuhiro','Tsuchiya','ivni4d6q0ddomzhr');  INSERT sfvi.user (email, username, first\_name, last\_name, pwds) VALUES ('yoshiono@capcom.com','theoneono','Yoshinori','Ono','3vpdn1kxarb25dn9');  INSERT sfvi.user (email, username, first\_name, last\_name, pwds) VALUES ('yusuke@capcom.com','yusuke','Yusuke','Hashimoto','2z24frcnq41rvw4k');  INSERT sfvi.user (email, username, first\_name, last\_name, pwds) VALUES ('theterayama@capcom.com','yoshitera','Yoshiya','Terayama','pxtknpr3kxqxxupb'); |

**Insert sfvi.USER\_TOPIC**

|  |
| --- |
| INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('[takanaka@capcom.com](mailto:takanaka@capcom.com)', 1);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('[takanaka@capcom.com](mailto:takanaka@capcom.com)', 2);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('[tsuchiya@capcom.com](mailto:tsuchiya@capcom.com)', 3);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('[yoshiono@capcom.com](mailto:yoshiono@capcom.com)', 4);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('[takanaka@capcom.com](mailto:takanaka@capcom.com)', 5);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('yusuke@capcom.com', 6);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('theterayama@capcom.com', 7);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('yusuke@capcom.com', 8);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('theterayama@capcom.com', 9);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('yoshiono@capcom.com', 10);  INSERT sfvi.USER\_TOPIC  (EMAIL, TOPIC\_ID) values ('tsuchiya@capcom.com', 11);  SELECT \* FROM sfvi.user\_topic; |

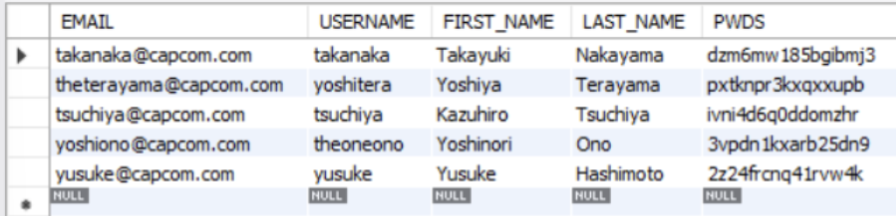
**Insert sfvi.topic**

|  |
| --- |
| INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('Drive impact is overpowered, how do we rebalance it?', STR\_TO\_DATE("29/10/2021 11:22:23", "%d/%m/%Y %H:%i:%s"), 1);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('Launch date is 2 June 2023, get ready!', STR\_TO\_DATE("12/11/2022 14:12:43", "%d/%m/%Y %H:%i:%s"), 2);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('New year''s cohesion dinner', STR\_TO\_DATE("01/01/2023 00:00:00", "%d/%m/%Y %H:%i:%s"), 2);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('I''ve been fired. Goodbye!', STR\_TO\_DATE("06/05/2021 07:08:34", "%d/%m/%Y %H:%i:%s"), 2);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('New character idea: Ballet dancer', STR\_TO\_DATE("22/06/2022 04:19:33", "%d/%m/%Y %H:%i:%s"), 3);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('New character idea: Female roman warrior', STR\_TO\_DATE("21/06/2022 03:23:13", "%d/%m/%Y %H:%i:%s"), 3);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('New BGM compositions', STR\_TO\_DATE("21/04/2021 11:11:12", "%d/%m/%Y %H:%i:%s"), 4);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('Should we bring parry back?', STR\_TO\_DATE("07/06/2021 05:45:45", "%d/%m/%Y %H:%i:%s"), 1);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('More new BGM compositions', STR\_TO\_DATE("28/02/2021 02:03:34", "%d/%m/%Y %H:%i:%s"), 4);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('Blanka doll giveaway', STR\_TO\_DATE("06/12/2020 08:18:24", "%d/%m/%Y %H:%i:%s"), 2);  INSERT sfvi.topic (topic\_title, dte\_cr\_topic, topic\_category\_id) values ('What do you guys think of having a Drive Gauge?', STR\_TO\_DATE("01/01/2023 03:19:45", "%d/%m/%Y %H:%i:%s"), 3);  SELECT \* FROM sfvi.topic; |

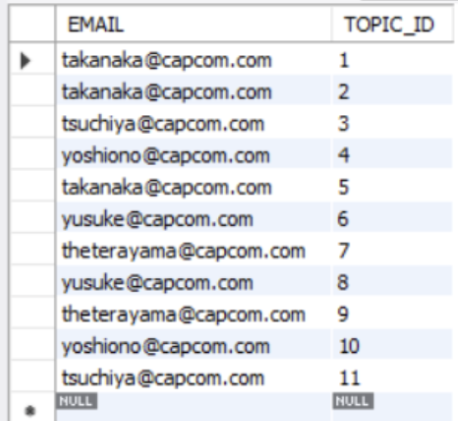
**Insert sfvi.cd\_topic\_category**

|  |
| --- |
| INSERT sfvi.cd\_topic\_category (topic\_category) values ('Mechanics & Balancing');  INSERT sfvi.cd\_topic\_category (topic\_category) VALUES ('General');  INSERT sfvi.cd\_topic\_category (topic\_category) VALUES ('New Features');  INSERT sfvi.cd\_topic\_category (topic\_category) VALUES ('Asset Creation');  SELECT \* FROM sfvi.cd\_topic\_category; |

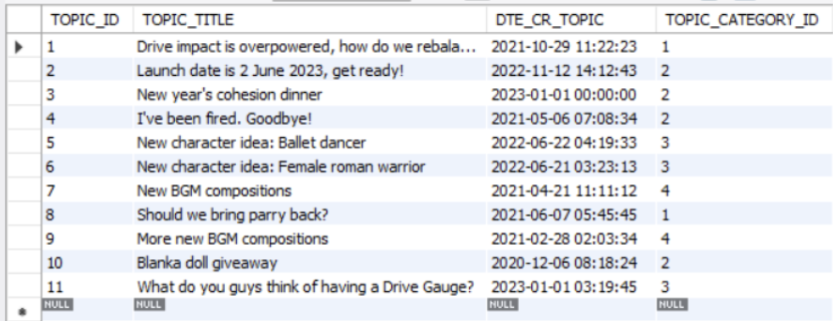
**sfvi.user**



**sfvi.user\_topic**

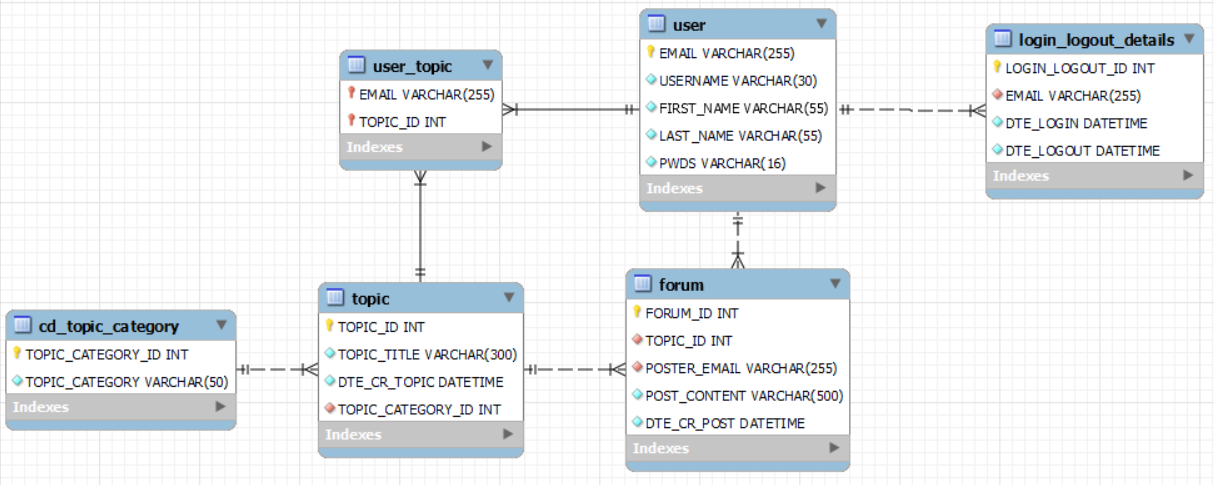


**sfvi.topic**



# Part 2: Migrating old post messages

## ER Diagram



## Create Table

**Create forum table**

CREATE TABLE sfvi.FORUM (

FORUM\_ID INT AUTO\_INCREMENT PRIMARY KEY,

TOPIC\_ID INT NOT NULL,

POSTER\_EMAIL VARCHAR(255) NOT NULL,

POST\_CONTENT VARCHAR(500) NOT NULL,

DTE\_CR\_POST DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP,

FOREIGN KEY (TOPIC\_ID) REFERENCES sfvi.TOPIC(TOPIC\_ID),

FOREIGN KEY (POSTER\_EMAIL) REFERENCES sfvi.USER(EMAIL)

);

**Create login logout detail table**

CREATE TABLE SFVI.LOGIN\_LOGOUT\_DETAILS (

LOGIN\_LOGOUT\_ID INT AUTO\_INCREMENT PRIMARY KEY,

EMAIL VARCHAR(255) NOT NULL,

DTE\_LOGIN DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP,

DTE\_LOGOUT DATETIME NOT NULL,

FOREIGN KEY (EMAIL) REFERENCES sfvi.USER(EMAIL)

);

## INSERT CSV DATA

|  |
| --- |
| INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (1, 'takanaka@capcom.com', 'There are too many hits of armor.', STR\_TO\_DATE("29/10/2021 11:22:23", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (1, 'yoshiono@capcom.com', 'I think it is fine. I like it.', STR\_TO\_DATE("30/10/2021 19:15:11", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (1, 'yusuke@capcom.com', 'Same here.', STR\_TO\_DATE("31/10/2021 14:35:31", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (2, 'takanaka@capcom.com', 'We will make lots of money.', STR\_TO\_DATE("12/11/2022 14:12:43", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (2, 'theterayama@capcom.com', 'Lets go!', STR\_TO\_DATE("12/11/2022 16:32:53", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (3, 'tsuchiya@capcom.com', 'Let''s have a good meal together.', STR\_TO\_DATE("01/01/2023 00:00:00", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (3, 'yoshiono@capcom.com', 'See you there!', STR\_TO\_DATE("01/01/2023 00:13:34", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (4, 'yoshiono@capcom.com', 'No one likes me.', STR\_TO\_DATE("06/05/2021 07:08:34", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (5, 'takanaka@capcom.com', 'Let''s call her MANon.', STR\_TO\_DATE("22/06/2022 04:19:33", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (5, 'theterayama@capcom.com', 'Sick idea bro.', STR\_TO\_DATE("23/06/2022 06:28:53", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (6, 'yusuke@capcom.com', 'Let''s all her Marisa.', STR\_TO\_DATE("24/06/2022 03:23:13", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (6, 'tsuchiya@capcom.com', 'Why give her a girly name?', STR\_TO\_DATE("21/06/2022 01:21:32", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (7, 'theterayama@capcom.com', 'Check it out.', STR\_TO\_DATE("22/04/2021 11:11:12", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (7, 'takanaka@capcom.com', 'Nobody cares.', STR\_TO\_DATE("23/04/2021 13:28:31", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (8, 'yusuke@capcom.com', 'Is it too overpowered?', STR\_TO\_DATE("07/06/2021 05:45:45", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (8, 'yoshiono@capcom.com', 'I think it''s fine.', STR\_TO\_DATE("08/06/2021 01:23:15", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (9, 'theterayama@capcom.com', 'Please care guys.', STR\_TO\_DATE("28/02/2021 02:03:34", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (9, 'takanaka@capcom.com', 'Nobody cares.', STR\_TO\_DATE("02/03/2021 01:15:33", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (9, 'yoshiono@capcom.com', 'I care.', STR\_TO\_DATE("02/03/2021 04:12:14", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (10, 'yoshiono@capcom.com', 'Before I leave.', STR\_TO\_DATE("06/12/2020 08:18:24", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (10, 'yoshiono@capcom.com', 'Don''t ignore me.', STR\_TO\_DATE("09/12/2020 12:38:11", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (10, 'yoshiono@capcom.com', 'I''m gonna cry.', STR\_TO\_DATE("12/12/2020 23:48:41", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (11, 'tsuchiya@capcom.com', 'Will it be fun?', STR\_TO\_DATE("01/01/2023 03:19:45", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (11, 'yusuke@capcom.com', 'Yes I think it will be.', STR\_TO\_DATE("02/01/2023 17:34:23", "%d/%m/%Y %H:%i:%s"));  INSERT sfvi.forum (topic\_id, poster\_email, post\_content, dte\_cr\_post) values (11, 'theterayama@capcom.com', 'I think so too.', STR\_TO\_DATE("04/01/2023 20:13:33", "%d/%m/%Y %H:%i:%s"));  SELECT \* FROM sfvi.forum; |

## Loading CSV data

LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/LOGIN\_LOGOUT\_DETAILS.csv'

INTO TABLE SFVI.LOGIN\_LOGOUT\_DETAILS

FIELDS TERMINATED BY ','

ENCLOSED BY '"'

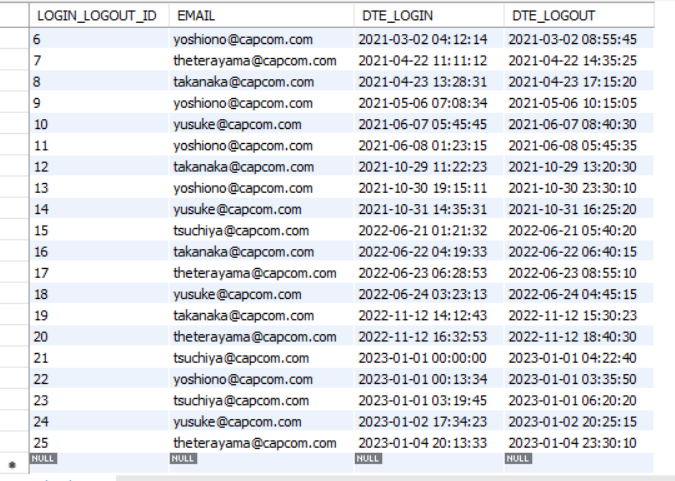
LINES TERMINATED BY '\n'

IGNORE 1 ROWS

(LOGIN\_LOGOUT\_ID, EMAIL, @DTE\_LOGIN, @DTE\_LOGOUT)

SET DTE\_LOGIN = STR\_TO\_DATE( @DTE\_LOGIN, '%d/%m/%Y %H:%i:%s'),

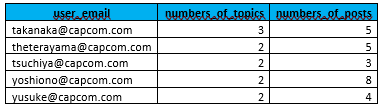
DTE\_LOGOUT = STR\_TO\_DATE( @DTE\_LOGOUT, '%d/%m/%Y %H:%i:%s');



## Usage report

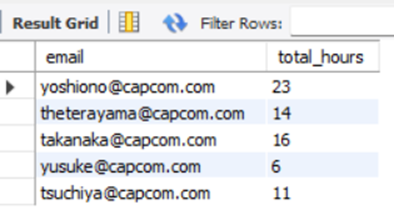
### Display the amount of topics and posts made by the person.

|  |
| --- |
| u.email AS user\_email,  (SELECT count(topic\_id) FROM sfvi.user\_topic WHERE email = u.email) AS numbers\_of\_topics,  (SELECT count(forum\_id) FROM sfvi.forum WHERE poster\_email = u.email) AS numbers\_of\_posts  FROM  sfvi.user u  ORDER BY  email; |



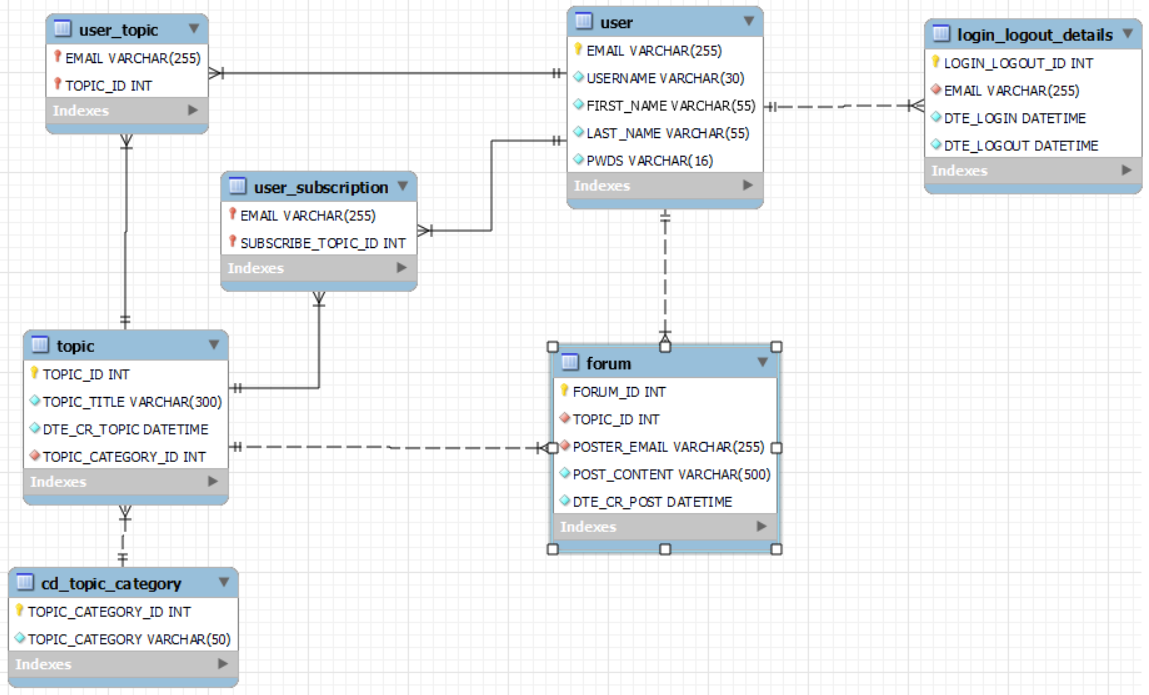
### Displays the amount of time each user has spent playing the game.

|  |
| --- |
| SELECT email, sum(hours) total\_hours FROM  (  SELECT  email, TIMESTAMPDIFF(hour, dte\_login, dte\_logout)Hours  FROM  SFVI.LOGIN\_LOGOUT\_DETAILS  ) a  GROUP BY email;      Output: |



# Part 3: New features

## ER Diagram



## SQL query showing the number of post

This shows the number of post a user has made between October 2022 to Deccember 2022.

CREATE TABLE SFVI.USER\_SUBSCRIPTION (

EMAIL VARCHAR(255),

SUBSCRIBE\_TOPIC\_ID INT,

PRIMARY KEY (EMAIL, SUBSCRIBE\_TOPIC\_ID),

FOREIGN KEY (EMAIL) REFERENCES SFVI.USER(EMAIL),

FOREIGN KEY (SUBSCRIBE\_TOPIC\_ID) REFERENCES SFVI.TOPIC(TOPIC\_ID)

);



SELECT

poster\_email, count(\*) AS number\_of\_posts

FROM

sfvi.forum

WHERE

dte\_cr\_post  BETWEEN '2022-10-01' AND '2022-12-31'

GROUP BY

poster\_email;



## SQL Query

showing each quarter year have how many comments from each indvidual

SELECT

poster\_email, concat(YEAR(dte\_cr\_post) ,'-Q', QUARTER(dte\_cr\_post)) as quarter\_year, count(\*)

FROM

sfvi.forum

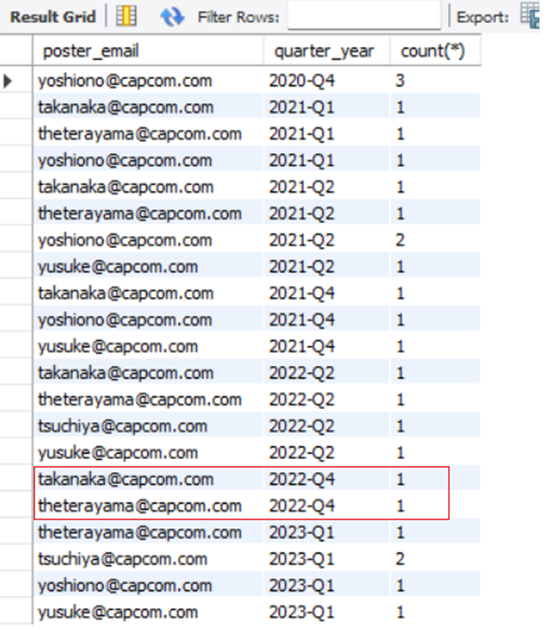
GROUP BY

poster\_email, quarter\_year

ORDER BY

quarter\_year, poster\_email;

Output:



# Summary

**Learning points and Challenges**

**Database schema design**

1. It is important to apply 3NF in database schema design to reduce redundancy. In our project, we have denormalized the XXXX table into the following tables:

|  |  |
| --- | --- |
| Table | Description/ Our thinking process |
| User | To store the detail related to users |
| Topic category | To store the detail related to the category of the topics so that the category description does not need to be replicated throughout the Topic table |
| Topic | To store the detail and title related to the topic. We created a unique primary key called Topic ID for each topic so that the topic title would not be replicated in the forum table. |
| User Topic Relationship | We decouple the user details from the topics content and only keep the bare minimum user information (i.e. the user email) and topic ID. This helps to maintain the relationship between the users and topics. If we want to know the detail of the user or the detail of the topic content, we can perform a join using the user email or topic ID to the user table or the topic table respectively. |
| Login-Logout table | This table keeps track of the login and logout timestamp of the user so that we can use this to compute the time each user spends on the game.  Initially, we wanted to include an additional field in the Forum table to keep track of the user logout time together with the user posting time. However, with this design, the duration between the posting time and logout time will overlap if there is multiple posting records inserted. This will make the computation of duration spent difficult. Hence, we decided to use a separate table to solely keep track of the login and logout activities. |
| User forum table | This table keeps the content of the posts posted by the users.  It maintains a relationship between the user email, the topic ID and the post content.  We do not replicate the topic title in this table. Instead, we just the topic ID as the foreign key. If we want to know the topic detail such as the title, we can do a join to the topic table |
| User subscription |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Sequence of table creation**

After we have designed the database schema, each of us split the work to come out with the CREATE table statements. We have learnt that we must run the scripts in proper sequence. Otherwise, we will hit into error such as Unable to create table due to missing fields referred in the foreign key.

Hence, the rule of thumb is to create the table before any foreign key referring to the field being created.

We create our table in this sequence:

1. User
2. Topic Category
3. Topic
4. Forum
5. Login Logout detail
6. User subscription

**Sequence of data insertion**

Similar to table creation, the sequence to execute the INSERT statements to the various tables is also important. We must make sure that the data being referenced by another table as a foreign key need to be first created in the main table. E.g. We cannot insert a record with a user email XXXXXX into the forum table before a record of this email is inserted in the USER table.

**Insertion of Date format**

The format of date provided in the Excel file is “XXXXX”

This is not the standard format of the data time in MySQL. Hence, we are required to convert the data in Excel using to\_str () function to format the data into the system datetime format before insertion.

**Generation of unique Topic ID**

We designed to use a system auto-generated sequence number for the Topic ID. This topic ID is referred in Forum table and User Topic Relationship table as a foreign keys. When we are preparing the insert statements for the Forum and User Topic Relationship tables, we need to make sure that we refer to the correct Topic ID that is auto generated in the Topic table. Otherwise, we will hit “TOPIC\_ID xxx does not exists” error.

**Computation of posts for a quarter**

Since the post datetime is in XXXXX format, the counting of the posts using GROUP BY statement is not straightforward. Hence, we need to convert the post datetime into different quarters using the QUARTER function. However, this function only returns the number related to the quarter without the year component. Hence, we also need to get the YEAR from the post datetime. After we have obtained the YEAR and QUARTER, we use CONCAT function to concatenate the year and quarter to form this YYYY-‘Q’Q format (e.g. 2023-Q1) and we doing the GROUP BY using this new field created in the SELECT statement so that we can count by the different quarters.