Assignment 3: Foreign Keys and Joins

Question 1 (10 marks)

For this question, you will need to recall your knowledge of relational design, and familiarize yourself with MySQL's `CREATE TABLE` syntax.

Below is a schema (this time, expressed as a series of create statements) for a firm's employee data:

```
CREATE TABLE Salary (
  title varchar(12),
  years_experience int,
  salary decimal(15,2),
  PRIMARY KEY (title, years_experience)
);

CREATE TABLE Employee (
  sin int(9),
  title varchar(10),
  years_in_role int,
  full_name varchar(30) NOT NULL,
  manager varchar(30),
  PRIMARY KEY (sin),
  FOREIGN KEY (manager) REFERENCES Employee (full_name),
  FOREIGN KEY (title) REFERENCES Salary (title),
  FOREIGN KEY (years_in_role) REFERENCES Salary (years_experience)
);

CREATE TABLE OfficeBuilding (
  name_abbrev varchar(4), -- unique short abbreviation for office
  name varchar(20),
  -- address
  street_number int,
  street_name varchar(20),
  city varchar(20),
  province_state varchar(20),
  country varchar(20),
  pcode varchar(6),
  -- end address
  employee_sin int(9),
  PRIMARY KEY (name_abbrev),
  FOREIGN KEY (employee_sin) REFERENCES Employee (sin)
);
```

Here is some semantic information regarding the relationships above:

- An employee's salary is calculated based on their present job title, and the number of years of experience they have accumulated.
- Every employee has one manager (Except for the CEO, who is managed by themselves)
- Many employees work at one office building

Part 1 (4 marks)

For each of the identified FOREIGN KEYS in this schema, **identify any apparent design issues (if any). If there are no apparent design issues with a FOREIGN KEY, explicitly state there are no issues with that KEY.** (*Hint: you may want to sketch out a relational schema diagram for your own reference. This should not be handed in.*)

Part 2 (6 marks)

Create a new schema (you may do this in the form of CREATE TABLE commands or a relational schema drawn with draw.io or another diagramming tool) which resolves these issues and explain how your solution solves the problems.

Question 2 (14 marks)

Included with this assignment is a MySQL database backup file "data604_a3_bak.sql".

Use mysqlsh to run the commands in this file. First, connect to your chosen database.

```
<username>@<hostname or IP>:33060
```

> For example:
>
> `root@localhost:33060` (if MySQL is installed on your own machine)
>
> `lewu@162.246.156.87:33060` (if you are the user lewu and are accessing the remote instance available for DATA 604). **REMEMBER that you must edit your SQL file** so that you are using an appropriate database name containing your username as part of the name, on each of lines 22, 24, and 26 if you are using this remote instance).

Then copy the path to your script and run it by entering in your mysqlsh window in SQL mode.

```
\sql
```

```
\source <path to your script>
```

Once you have the database created in your environment, complete the following tasks using only SQL, and provide us with the resulting queries in order, in a text file named assignment3_responses.sql. Use SQL comments to indicate which part your queries are answering.

A) Return the name and address of each attraction in the city. (1 mark)
B) Return a list of the names of all community centres in SW Calgary. (1 mark)
C) Return a list of all communities in the NE containing 'ridge' in their name. (1 mark)
D) Return a list of all communities which are classified as 'Industrial' in the city. (1 mark)

E) Write a SQL statement to perform an inner join between `*community_services*` and `*community*` on the attribute *comm_code,* returning only the attributes Community.sector, comm_code and Community_Services.name. (2 marks)

F) Select the ward_number of every ward which with at least one community that had a median assessed value > 700000 in any year of assessment. (2 marks)

G) Return a list of all attractions in communities not classified as residential or industrial. (2 marks)

H) Return a list of PHS Clinics in Ward 2. (2 marks)

I) Generate a set of all community names, and for each community service, list it with its accompanying community name. If the community has no services available, it should still be listed in the table. (2 marks)

You may use a Jupyter notebook to display your code if this is preferable, but all code must run inside the notebook (i.e., you must connect to a MySQL instance using the mysql connector for Python)

Reminder: Here are some commands you may find helpful in the MySQL console:

```
SHOW DATABASES;

USE [database_name];

SHOW TABLES;

DESCRIBE [table_name];
```