

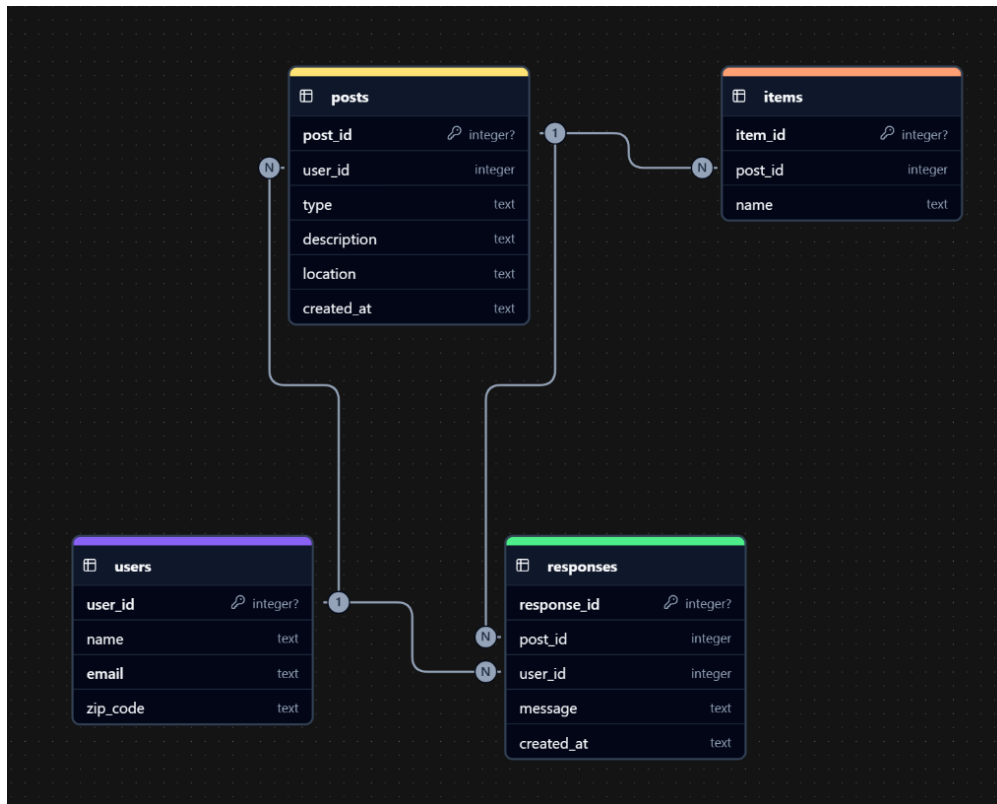
## Introduction to Applied Data Science

### Homework 3

Provide answers to the following questions as a single PDF file or Word document file.

**Q1:** Your task is to design an ER diagram for a community website that helps give away free items in your garage to others who might find them useful. Here is what you need to consider:

- Each **user** signs up with an **email address**, their **name**, and their **zip code**.
- The **zip code** is used to show the user **items** in their area and similarly to determine the audience for their own postings.
- A user can post **items** that they want to give away. They describe the **items** with some kind of **textual description**: pots and pans, collection of children's books, set of Pokemon cards or whatever.
- Each **post** can consist of **one item** or **multiple items**. The three examples above could be in separate posts or in one big post.
- Each **post** specifies a **general area** or **location** where they can be picked up. (Not the exact address, for privacy, but a neighborhood, nearby intersection or other landmark.)
- Other users can **respond** to a **posting**, indicating interest.
- **Users** can also **post wanted** or **seeking** notices, such as I'm looking for a children's size 12 party dress. As with offering items, the post includes the general location.
- Other users can respond to a wanted post, indicating that they have such an item.



### Primary Keys:

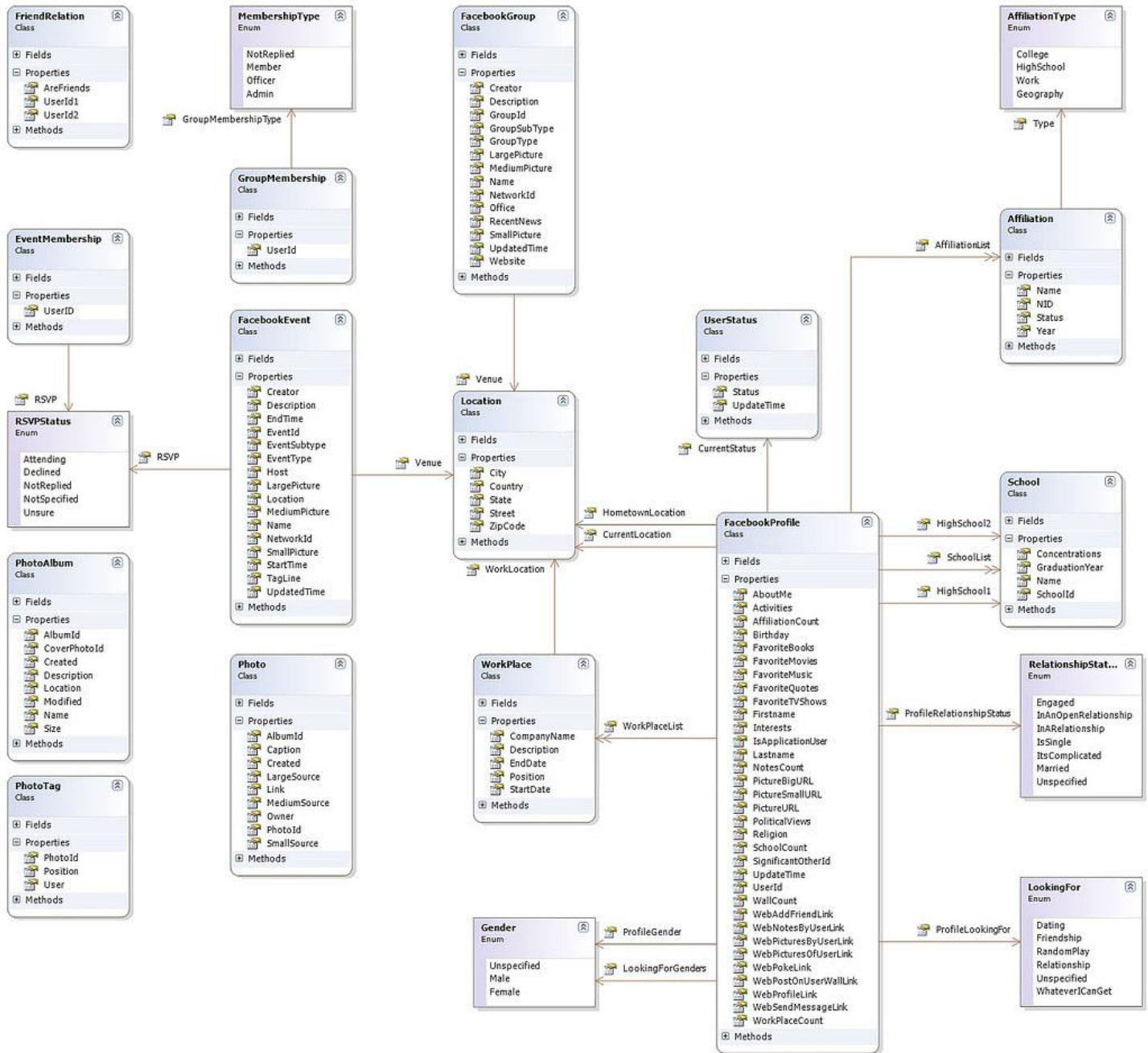
- `user_id`
- `post_id`
- `item_id`
- `response_id`

Posts are attached to user which contains a **zip\_code**, posts contain a location so locations become attached to a **zip\_code**.

Q2: Given the ER Diagram in the following web page:

<https://www.flickr.com/photos/ikhnaton2/533233247>

Please write down the names of all entities and relationships. Note that there is no relationships explicitly shown. But you can find possible relationships by looking at the attribute names. Also write down the type of entities (strong, weak, specialization) and the cardinalities of the relationships. For each entity write down possible candidate keys.



Please note - **this is not object relational diagram (entity diagram)**. Connected on this diagram means some type has a field of some other type. So photos are not related with tag and album. Each of them can be independent – (From the Blog Site)

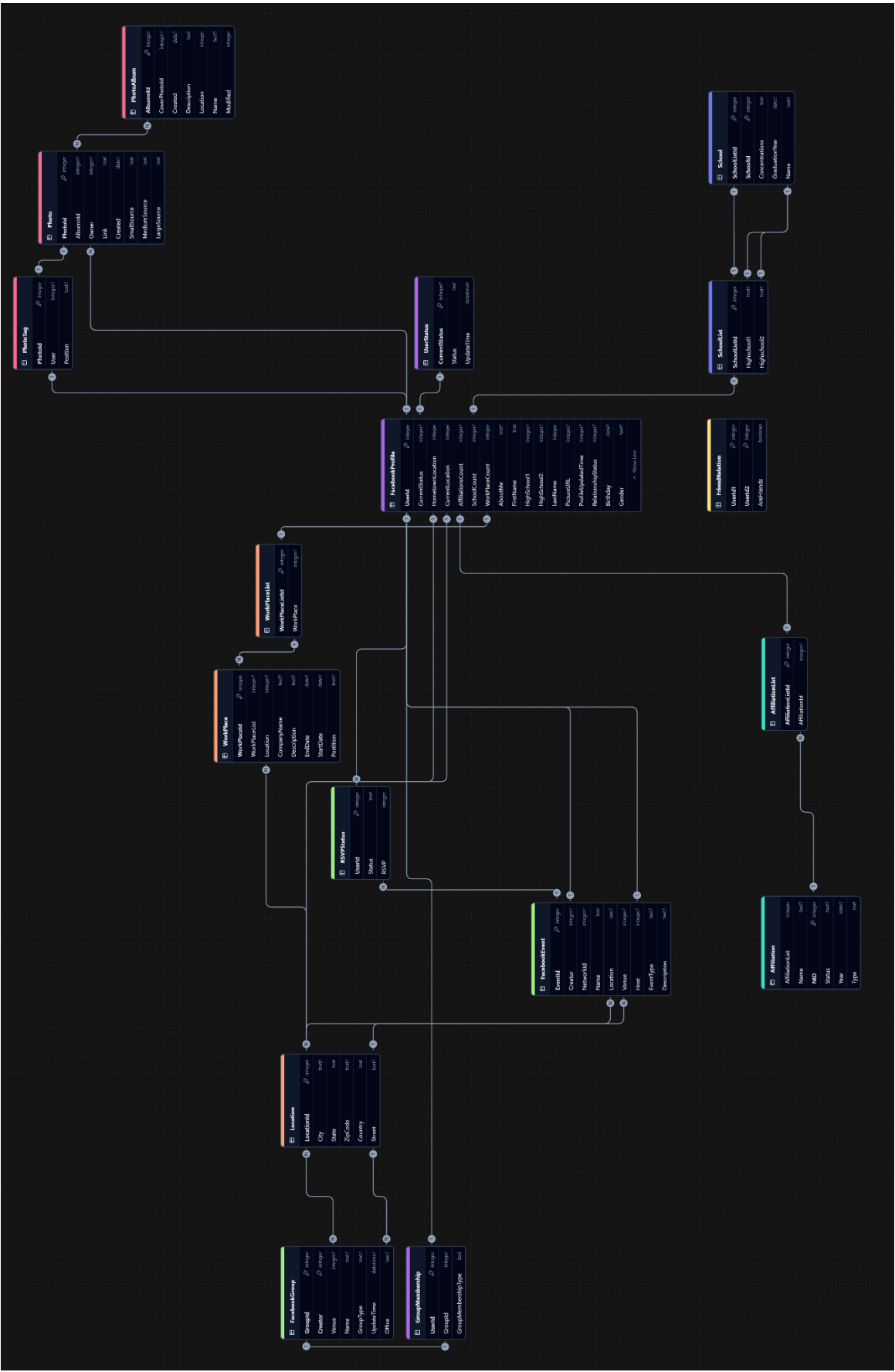
Entities with relationships in main schema	FacebookProfile, Gender, RelationshipStatus, LookingFor, UserStatus, Workplace, School, Affiliation, AffiliationType, Location, FacebookGroup, FacebookEvent, RSVPStatus, EventMembership
Entities independent of main schema (no relationship)	FriendRelation, PhotoAlbum, PhotoTag, Photo
Entities (other)	GroupMembership, MembershipType
<i>Relationship Sets (Different Class color?) – light purple</i>	<i>AffiliationType, Gender, RelationshipStatus, LookingFor, RSVPStatus, MembershipType</i>

While we cannot see the Fields (only Class properties) we can see the Cardinality which likely specifies the primary component/key for the relationships

Origin Entity	Related Entity	Cardinality (Relation)	Related Field/Key
FacebookProfile	Gender	One-to-One	ProfileGender LookingForGenders
FacebookProfile	RelationshipStatus	One-to-One	ProfileRelationshipStatus
FacebookProfile	LookingFor	One-to-One	ProfileLookingFor
FacebookProfile	UserStatus	One-to-One	CurrentStatus
FacebookProfile	School	One-to-Many One-to-One	SchoolList (Many) HighSchool1, HighSchool2 (One)
FacebookProfile	Affiliation	One-to-Many	AffiliationList
Affiliation	AffiliationType	One-to-One	Type
FacebookProfile	Workplace	One-to-Many	WorkPlaceList
Workplace	Location	One-to-One	WorkLocation
FacebookProfile	Location	One-to-One	HometownLocation CurrentLocation
FacebookGroup	Location	One-to-One	Venue
FacebookEvent	Location	One-to-One	Venue
FacebookEvent	RSVPStatus	One-to-One	RSVP
EventMembership	RSVPStatus	One-to-One	RSVP
GroupMembership	MembershipType	One-to-One	GroupMembershipType

This table is based solely off the arrows present, where One-to-Many indicates the multi-headed arrows (and a *VariableList* List relationship). It is hard to garner more information without looking directly at the fields/methods, however it would be safe to assume the expected primary keys for each entity.

For example, **FacebookProfile** -> **UserId** / **Photo** -> **Photoid** / **FacebookGroup** -> **NetworkId**, so on and so forth. The information present under the inserted photo lists why there are entities without relationships and what they are. (Even though it would make sense for **PhotoTag#User** -> **FacebookProfile#UserId** to be related)



Upon working on Q2, I realized that everything I put down for Q1 does not make any sense, so rather than rewrite up a whole page, I thought it be easier to just copy and paste the create statements in SQLite, then export that empty db into ChartDB, and define the relationships there.

I can't tell if we we're supposed to come up with our OWN ER Diagram based on the web-endpoint structure, and frankly it's my fault for starting this assignment late and not being able to ask in time.

**Q3: Convert the ER Diagram at the following web page:**

<https://www.flickr.com/photos/ikhnaton2/533233247>

**into a set of tables using the ER2DB methods we saw in class. Use the create table syntax we saw in class. Make any necessary assumptions and note them in your answers.**

Entity	Type	CREATE Statement
FacebookProfile	Strong	<pre>CREATE TABLE FacebookProfile (   UserId INTEGER PRIMARY KEY NOT NULL,   CurrentStatus INTEGER NOT NULL,   HometownLocation INTEGER NOT NULL,   CurrentLocation INTEGER NOT NULL,   AffiliationsCount INTEGER,   SchoolCount INTEGER,   WorkPlaceCount INTEGER NOT NULL,   AboutMe TEXT,   FirstName TEXT NOT NULL,   HighSchool1 INTEGER,   HighSchool2 INTEGER,   LastName INTEGER NOT NULL,   PictureURL TEXT,   ProfileUpdateTime DATETIME,   Birthday DATE,   Gender TEXT CHECK(Gender IN ('Male', 'Female', 'Unspecified')),   RelationshipStatus TEXT CHECK(RelationshipStatus IN ('Single', 'In a   Relationship', 'Engaged', 'Married', 'Its Complicated', 'Separated', 'Divorced',   'Widowed', 'Unspecified')),   FOREIGN KEY (CurrentLocation) REFERENCES Location(LocationId),   FOREIGN KEY (HometownLocation) REFERENCES Location(LocationId),   FOREIGN KEY (WorkPlaceCount) REFERENCES   WorkPlaceList(WorkPlaceListId),   FOREIGN KEY (AffiliationsCount) REFERENCES AffiliationList(AffiliationListId) );</pre>
FriendRelation	Weak	<pre>CREATE TABLE FriendRelation (   UserID1 INTEGER,   UserID2 INTEGER,   AreFriends BOOLEAN NOT NULL,   FOREIGN KEY (UserID1) REFERENCES FacebookProfile(UserId),   FOREIGN KEY (UserID2) REFERENCES FacebookProfile(UserId),   PRIMARY KEY (UserID1, UserID2) );</pre>
FacebookGroup	Strong	<pre>CREATE TABLE FacebookGroup (   GroupId INTEGER NOT NULL,   Creator INTEGER NOT NULL,   Venue INTEGER,</pre>

		Name TEXT, GroupType TEXT, UpdateTime TEXT, Office TEXT, PRIMARY KEY (GroupId, Creator), FOREIGN KEY (Venue) REFERENCES Location(LocationId), FOREIGN KEY (Office) REFERENCES Location(Street) );
GroupMembership	Weak	CREATE TABLE GroupMembership ( UserId INTEGER PRIMARY KEY NOT NULL, GroupId INTEGER NOT NULL, GroupMembershipType TEXT NOT NULL, FOREIGN KEY (GroupId) REFERENCES FacebookGroup(GroupId), FOREIGN KEY (UserId) REFERENCES FacebookProfile(UserId) );
FacebookEvent	Strong	CREATE TABLE FacebookEvent ( EventId INTEGER PRIMARY KEY NOT NULL, Creator INTEGER, NetworkId INTEGER, Name TEXT NOT NULL, Location TEXT, Venue INTEGER, Host INTEGER, EventType TEXT, Description TEXT, FOREIGN KEY (Venue) REFERENCES Location(LocationId), FOREIGN KEY (Location) REFERENCES Location(Street), FOREIGN KEY (Creator) REFERENCES FacebookProfile(UserId), FOREIGN KEY (Host) REFERENCES FacebookProfile(UserId) );
RSVPStatus	Weak	CREATE TABLE RSVPStatus ( UserId INTEGER PRIMARY KEY NOT NULL, Status TEXT CHECK(Status IN ('Attending', 'Declined', 'NotSpecified', 'NotReplied', 'Unsure')), RSVP INTEGER NOT NULL, FOREIGN KEY (UserId) REFERENCES FacebookProfile(UserId), FOREIGN KEY (RSVP) REFERENCES FacebookEvent(EventId) );
PhotoAlbum	Strong	CREATE TABLE PhotoAlbum ( AlbumId INTEGER PRIMARY KEY NOT NULL, CoverPhotoId INTEGER, Created DATE NOT NULL, Description TEXT NOT NULL, Location INTEGER, Name TEXT,



		Modified DATE NOT NULL, FOREIGN KEY (Location) REFERENCES Location(LocationId), FOREIGN KEY (CoverPhotoId) REFERENCES Photo(PhotoId) );
Photo	Weak	CREATE TABLE Photo ( PhotoId INTEGER PRIMARY KEY NOT NULL, AlbumId INTEGER, Owner INTEGER NOT NULL, Link TEXT NOT NULL, Created DATE NOT NULL, SmallSource TEXT NOT NULL, MediumSource TEXT NOT NULL, LargeSource TEXT NOT NULL, FOREIGN KEY (AlbumId) REFERENCES PhotoAlbum(AlbumId), FOREIGN KEY (Owner) REFERENCES FacebookProfile(UserId) );
PhotoTag	Weak	CREATE TABLE PhotoTag ( PhotoId INTEGER PRIMARY KEY NOT NULL, User INTEGER NOT NULL, Position TEXT, FOREIGN KEY (PhotoId) REFERENCES Photo(PhotoId), FOREIGN KEY (User) REFERENCES FacebookProfile(UserId) );
WorkPlaceList	Weak	CREATE TABLE WorkPlaceList ( WorkPlaceListId INTEGER PRIMARY KEY NOT NULL, WorkPlace INTEGER, FOREIGN KEY (WorkPlace) REFERENCES WorkPlace(WorkPlaceId) );
WorkPlace	Strong	CREATE TABLE WorkPlace ( WorkPlaceId INTEGER PRIMARY KEY NOT NULL, WorkPlaceList INTEGER NOT NULL, Location INTEGER, CompanyName TEXT NOT NULL, Description TEXT, EndDate DATE, StartDate DATE, Postition TEXT, FOREIGN KEY (Location) REFERENCES Location(LocationId), FOREIGN KEY (WorkPlaceList) REFERENCES WorkPlaceList (WorkPlaceListId) );
AffiliationList	Weak	CREATE TABLE AffiliationList ( AffiliationListId INTEGER PRIMARY KEY NOT NULL, AffiliationId INTEGER, FOREIGN KEY (AffiliationId) REFERENCES Affiliation(NID),

		FOREIGN KEY (AffiliationListId) REFERENCES FacebookProfile(AffiliationCount) );
Affiliation	Strong	CREATE TABLE IF Affiliation ( AffiliationList INTEGER, Name TEXT, NID INTEGER PRIMARY KEY NOT NULL, Status TEXT, Year DATE, Type TEXT CHECK(Type IN(('College', 'Highschool', 'Work', 'Geography'))), FOREIGN KEY (AffiliationList) REFERENCES AffiliationList(AffiliationListId) );
SchoolList	Weak	CREATE TABLE SchoolList ( SchoolListId INTEGER PRIMARY KEY NOT NULL, Highschool1 TEXT, Highschool2 TEXT, FOREIGN KEY (Highschool1) REFERENCES School(Name), FOREIGN KEY (Highschool2) REFERENCES School(Name), FOREIGN KEY (SchoolListId) REFERENCES FacebookProfile(SchoolCount) );
School	Strong	CREATE TABLE School ( SchoolListId INTEGER NOT NULL, SchoolId INTEGER NOT NULL, Concentrations TEXT NOT NULL, GraduationYear TEXT, Name TEXT, PRIMARY KEY (SchoolListId, SchoolId), FOREIGN KEY (SchoolListId) REFERENCES SchoolList(SchoolListId) );
Location	Strong	CREATE TABLE Location ( LocationId INTEGER PRIMARY KEY NOT NULL, City TEXT, State TEXT, Country TEXT NOT NULL, ZipCode TEXT, Street TEXT, );
UserStatus	Weak	CREATE TABLE UserStatus ( CurrentStatus INTEGER PRIMARY KEY, Status TEXT NOT NULL, UpdateTime TEXT, FOREIGN KEY (CurrentStatus) REFERENCES FacebookProfile(CurrentStatus) );

## Assumptions:

A lot of liberty was taken when trying to take this image and turn it into an actual database schema, as frankly, no way using everything from the original worked neatly (debatably functionally).

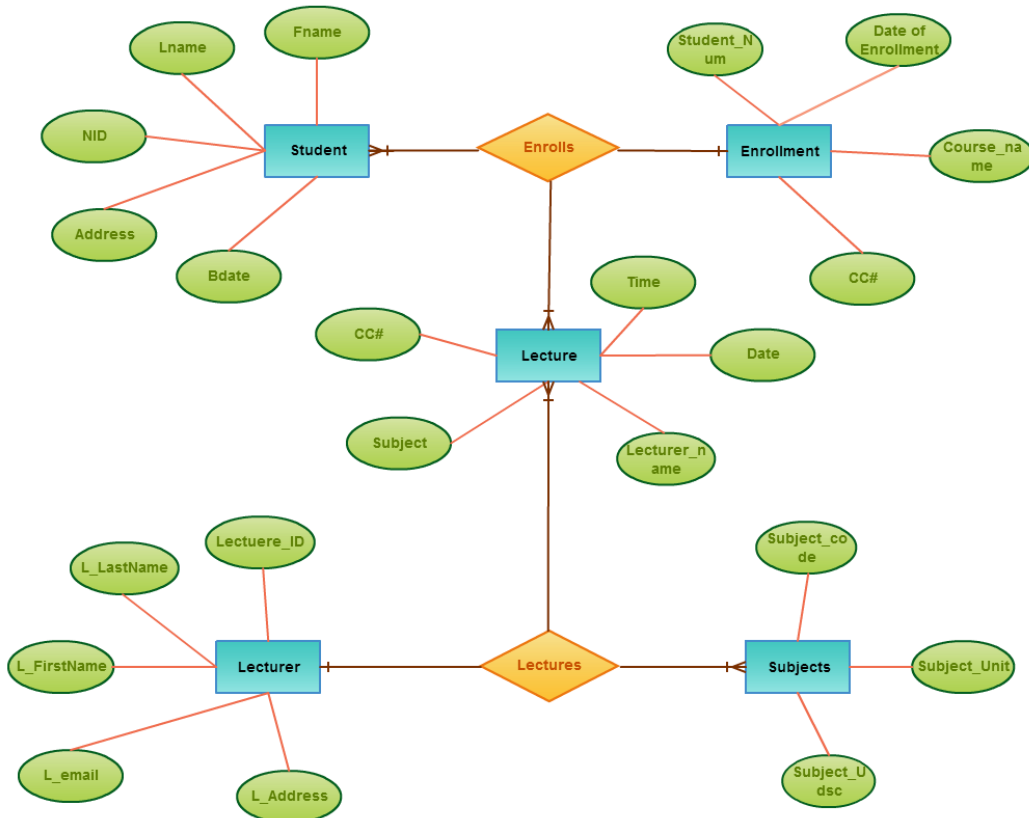
Instead, I took the image and re-built it from the ground-up and included attributes where appropriate to properly link entities that were disjointed but contained properties/attributes that were visibly present in each class.

Consider this more an 'artist's interpretation', as this frankly gave me a headache. I'm forgetting some attributes, but I included all of the key components as well most subcomponents.

- **AttributeName\_Id** has been added to entities that needed one in order to link properly given the original diagram's intended function
- Photo, PhotoTag, PhotoAlbum are all actually linked, and are linked to the profile as well.
- Venue is attribute that holds the actual LocationId, while the other attributes—such as Office and Location—are the holders for the Name (I chose street name for this).
- RSVP in RSVPStatus is the EventId, otherwise it becomes difficult to link
- The **purple colored entities** from the original diagram (weak) can be described inserted in other entities by adding a CHECK statement
- In **FacebookProfile**, all of the **EntityCount** attributes are the **Id variables** that link to lists (affiliation, work, etc.)

**Q4: Convert the following ER Diagram into a set of tables using the ER2DB methods we saw in class. Use the create table syntax we saw in class. Make any necessary assumptions and note them in your answers.**

**ER DIAGRAM FOR STUDENT ENROLLMENT SYSTEM**



Entity	CREATE Statement
Lecture	<pre> CREATE TABLE Lecture (   Lecture_ID INT AUTO_INCREMENT PRIMARY KEY,   CC_Num INTEGER NOT NULL,   Subject TEXT NOT NULL,   Lecturer_name TEXT NOT NULL,   Date DATE NOT NULL,   Time TIME NOT NULL,   FOREIGN KEY (Lecturer_name) REFERENCES Lecturer(Lecturer_ID),   FOREIGN KEY (CC_num) REFERENCES Subjects(Subject_code),   FOREIGN KEY (Subject) REFERENCES Subjects(Subject_Unit) ); </pre>
Lecturer	<pre> CREATE TABLE Lecturer (   Lecturer_ID INTEGER PRIMARY KEY,   L_FirstName TEXT NOT NULL,   L_LastName TEXT NOT NULL,   L_email TEXT UNIQUE NOT NULL,   L_Address TEXT </pre>

	);
Enrollment	CREATE TABLE Enrollment ( Student_Num INTEGER, Course_name TEXT NOT NULL, Date_of_Enrollment DATE NOT NULL, CC_Num INTEGER, PRIMARY KEY (Student_Num, CC_Num), FOREIGN KEY (Student_Num) REFERENCES Student(NID), FOREIGN KEY (CC_Num) REFERENCES Subjects(Subject_code), );
Subjects	CREATE TABLE Subjects ( Subject_code INTEGER PRIMARY KEY, Subject_Unit TEXT NOT NULL, Subject_U_desc TEXT );
Student	CREATE TABLE Student ( NID INTEGER PRIMARY KEY, Lname TEXT NOT NULL, Fname TEXT NOT NULL, Address TEXT, Bdate DATE );
<b>Relationship</b>	<b>CREATE Statement</b>
Lectures	CREATE TABLE Lectures ( Lecture_ID INTEGER PRIMARY KEY, Lecturer_ID INTEGER NOT NULL, Subject_code INTEGER NOT NULL, FOREIGN KEY (Lecture_ID) REFERENCES Lecture(Lecturer_ID), FOREIGN KEY (Lecturer_ID) REFERENCES Lecturer(Lecturer_ID), FOREIGN KEY (Subject_code) REFERENCES Subjects(Subject_code) );
Enrolls	CREATE TABLE Enrolls ( Student_ID INTEGER NOT NULL, Enrollment_ID INTEGER NOT NULL, PRIMARY KEY (Student_ID, Enrollment_ID), FOREIGN KEY (Student_ID) REFERENCES Student(NID), FOREIGN KEY (Enrollment_ID) REFERENCES Enrollment(Student_Num) );

### Assumptions:

- All values specified as primary key's (CC\_num/NID/Subject\_code) are primary keys
- There is no id/attribute specified for the Lecture\_ID unless it's CC\_num, which can either be course-code or the lecture id. As CC\_num is referenced in Enrollment, it makes more sense for it to be the Subject\_code
  - o As such, to make lectures unique, it make's sense to make the **primary key a composite of Subject\_code and date** as this will be unique, otherwise, we can create Lecture\_ID as:

Lecture\_ID INT AUTO\_INCREMENT PRIMARY KEY

To create an auto-incrementing identifier that is unique and not present in the ER Diagram. As this begins to enter the area of SQL beyond my experience, I'm not sure how it would handle the circular logic if we used a composite key for the Lectures table as well.

As such, I'll just add Lecture\_ID to Lecture through the method above, but it should be noted a composite key would likely be more efficient.

Something like:

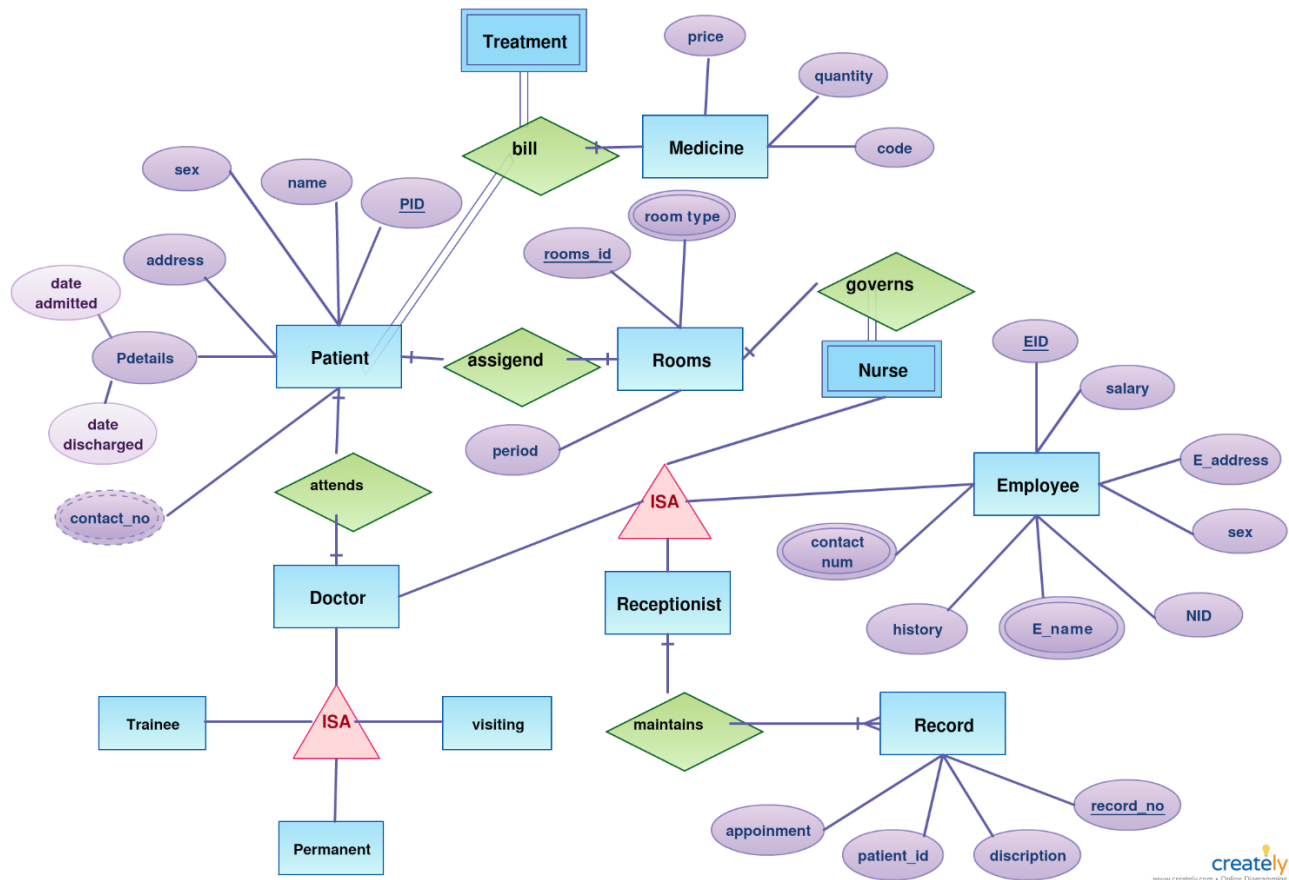
```
CREATE TABLE Lecture (  
  CC_Num INTEGER NOT NULL,  
  Subject TEXT NOT NULL,  
  Lecturer_name TEXT NOT NULL,  
  Date DATE NOT NULL,  
  Time TIME NOT NULL,  
  PRIMARY KEY (CC_Num, Date),  
  FOREIGN KEY (Lecturer_name) REFERENCES Lecturer(Lecturer_ID),  
  FOREIGN KEY (CC_num) REFERENCES Subjects (Subject_code),  
  FOREIGN KEY (Subject) REFERENCES Subjects(Subject_Unit)  
);
```

```
CREATE TABLE Lectures (  
  CC_Num INTEGER NOT NULL  
  Date DATE NOT NULL,  
  Lecturer_ID INTEGER NOT NULL,  
  Subject_code INTEGER NOT NULL,  
  FOREIGN KEY (CC_Num, Date) REFERENCES Lecture(CC_Num, Date),  
  FOREIGN KEY (Lecturer_ID) REFERENCES Lecturer(Lecturer_ID),  
  FOREIGN KEY (Subject_code) REFERENCES Subjects(Subject_code)  
);
```

But this is just a *really* messy way of linking Lecturer\_ID to all the lectures and subjects they taught.

**Q5: Convert the following ER Diagram into a set of tables using the ER2DB methods we saw in class. Use the create table syntax we saw in class. Make any necessary assumptions and note them in your answers.**

**E-R Diagram for Hospital Management System**



Entity	CREATE Statement
Patient	<pre>CREATE TABLE Patient (   PID INTEGER PRIMARY KEY,   name TEXT NOT NULL,   sex TEXT CHECK (sex IN ('Male', 'Female', 'Other')),   address TEXT,   date_admitted DATE NOT NULL,   date_discharged DATE NOT NULL,   contact_no TEXT );</pre>
Doctor	<pre>CREATE TABLE Doctor (   DID INTEGER PRIMARY KEY,   name TEXT NOT NULL,   specialization TEXT,   FOREIGN KEY (DID) REFERENCES Employee(EID)</pre>

	FOREIGN KEY (name) REFERENCES Employee(E_name) );
Doctor ISA	CREATE TABLE Doctor_Type ( DID INTEGER PRIMARY KEY, type TEXT CHECK (type IN ('Trainee', 'Permanent', 'Visiting')), FOREIGN KEY (DID) REFERENCES Doctor(DID) );
Rooms	CREATE TABLE Rooms ( rooms_id INTEGER PRIMARY KEY, room_type TEXT );
Medicine	CREATE TABLE Medicine ( code INTEGER PRIMARY KEY, price REAL, quantity INTEGER );
Treatment	CREATE TABLE Treatment ( PID INTEGER, code INTEGER, bill REAL, PRIMARY KEY (PID, code), FOREIGN KEY (PID) REFERENCES Patient(PID), FOREIGN KEY (code) REFERENCES Medicine(code) );
Employee	CREATE TABLE Employee ( EID INTEGER PRIMARY KEY, E_name TEXT NOT NULL, sex TEXT CHECK (sex IN ('Male', 'Female', 'Other')), salary REAL, E_address TEXT, NID INTEGER, contact_num TEXT, type TEXT CHECK (type IN ('Nurse', 'Receptionist')) );
Receptionist	CREATE TABLE Receptionist ( EID INTEGER PRIMARY KEY, visiting TEXT, FOREIGN KEY (EID) REFERENCES Employee(EID) );
Record	CREATE TABLE Record ( record_no INTEGER PRIMARY KEY, patient_id INTEGER, appointment TEXT, history TEXT, description TEXT,



	FOREIGN KEY (patient_id) REFERENCES Patient(PID) );
Relationship	CREATE Statement
Attends	CREATE TABLE Attends ( PID INTEGER, DID INTEGER, PRIMARY KEY (PID, DID), FOREIGN KEY (PID) REFERENCES Patient(PID), FOREIGN KEY (DID) REFERENCES Doctor(DID) );
Assigned	CREATE TABLE Assigned ( PID INTEGER, rooms_id INTEGER, period TEXT, PRIMARY KEY (PID, rooms_id), FOREIGN KEY (PID) REFERENCES Patient(PID), FOREIGN KEY (rooms_id) REFERENCES Rooms(rooms_id) );
Governs	CREATE TABLE Governs ( EID INTEGER, rooms_id INTEGER, PRIMARY KEY (EID, rooms_id), FOREIGN KEY (EID) REFERENCES Employee(EID), FOREIGN KEY (rooms_id) REFERENCES Rooms(rooms_id) );
Maintains	CREATE TABLE Maintains ( EID INTEGER, record_no INTEGER, PRIMARY KEY (EID, record_no), FOREIGN KEY (EID) REFERENCES Receptionist(EID), FOREIGN KEY (record_no) REFERENCES Record(record_no) );

### Assumptions:

This one was a bit more straight forward, so not many assumptions were made regarding what is what. I believe everything should be properly linked and that you should be able to find the information by query accordingly.

- I have absolutely no idea what NID is
- Nurse -> governs IS linked to Rooms
- Bill is a component of treatment and is a function-based parameter (sum of all medicine costs) as such the only way to implement it is with **triggers** otherwise you would need to use queries:

```

CREATE TRIGGER update_bill
AFTER INSERT OR UPDATE ON Treatment
FOR EACH ROW
BEGIN
    UPDATE Treatment
    SET bill = (SELECT SUM(price)
                FROM Medicine
                WHERE code = NEW.code)
    WHERE PID = NEW.PID AND code = NEW.code;
END;

CREATE TRIGGER update_bill_on_medicine_change
AFTER UPDATE OF price ON Medicine
FOR EACH ROW
BEGIN
    UPDATE Treatment
    SET bill = (SELECT SUM(price)
                FROM Medicine
                WHERE code = Treatment.code)
    WHERE code = OLD.code;
END;

```

Should fix the issue while keeping the pricing updated as well, treatment is still linked to PID.

**Q6. For this question, you need to install the SQLite database management system on your laptop or PC and then download the chinook database ([SQLite Sample Database And Its Diagram \(in PDF format\)](#) ([sqliitetutorial.net](#)) and run it in SQLite DBMS. The following two links will help:**

[How to Install SQLite on Windows, Mac or Linux \(servermania.com\)](https://servermania.com/how-to-install-sqlite-on-windows-mac-or-linux/)

[SQLite Sample Database And Its Diagram \(in PDF format\) \(sqllitetutorial.net\)](http://sqllitetutorial.net)

**Do the following and answer the questions:**

1. Perform the command: `.tables`

## How many tables are there?

There are 11 tables (albums, employees, invoices, playlists, artists, genres, media\_types, tracks, customers, invoice\_items, playlist\_track)

2. Perform the query: `select * from genres;`

### What is the Genre Id for Jazz?

The GenreId for Jazz is 2.

### 3. Perform the query:

```
select Name from tracks where GenreId = 2;
```

## How many rows are there?

Using `select COUNT(Name) from tracks where GenreId = 2;`  
Returns **130 rows**.

**4. Write a note on what do you think you did in the above three tasks?**

- Loaded a database
- Explored the number of tables
- Explored the information in the tables
- Queried the amount of data present for a key topic of the database

This would likely qualify as rudimentary data exploration in concept relating to initial exposure to a new language/data structure.