**Week 5 Homework**

The following practice may need to be completed using the databases hosted on wmc3317-2 using WB. Use WB whether you are running queries or creating a relational schema (a data model, essentially). Submit homework as a SQL script file. To mark the answers I will copy+paste the SQL script into my WB and execute the code. You could even write essay answers in an MS Word file or the SQL script itself – or in a separate MS Word/PDF file. If there is any dirty data, clean it by making reasonable assumptions.

1. **Consider the Sakila data model used in class.** (you cannot test these queries since either the schema does not match or there is no data)

Customer(custid (PK), firstname)

Rental (rentalid (PK), invid (FK), customerid (FK), dtMonth, dtYear, dtDay)

Inventory (invid (PK), filmid (FK))

Film (filmid (PK), length, name, cost, releaseMth, releaseYr)

Film\_actor(filmid (PK FK), actorid (PK FK))

Actor (actorid (PK), actorname)

Note: Film.Cost refers to the film studio’s cost of making the film

1. We need to analyze whether the cost of making a movie does indeed predict the rental volume (number of rentals) in the first three months after its release month, create a query that lists each movie, its name, alongside its cost, and the rental volume in the three months after its release. *For simplicity, assume no film is released from Oct to December of any year.*
2. Now consider additional data acquired by mining tweets (from Twitter). numMentions is the number of times the actor has been mentioned in all the tweets for each month of each year. Make any other assumptions you need to.

ActorTweet(actorid (PK FK), numMentions, monthyear (PK FK))

Create the query that lists the actor, the total volume of rentals (number) during a month for all the movies that actor has acted in, the number of mentions for that actor on Twitter the previous month, and the total cost of making all those films he/she has acted in. *Assume no film is released in January.*

1. Consider the following data model for the Employees database.

emp(empid (PK), name, bach, yrbach, mstr, yrmstr, phd, yrphd)

dept(deptid (PK), depname, empMgrid, dtstart, dtend)

deptEmp(deptid (PK, FK), empid (PK, FK), startDt (PK), endDt)

salary(empid (PK FK), salary, fromdt (PK), todt)

The columns - bach, mstr and phd - refer to the number of such degrees. The years (e.g., yrbach) pertain to the most recent ones (most recent bachelors, masters and Ph.D.). Write a query to list the employees, the number of degrees each has, the number of years since the most recent degree (use curryear()) to refer to the current year, and the most recent salary (use max(todt) to filter out the most recent salary for each employee).

1. Create a data model for the Volunteer Club running case. Write the following queries based on that model.
2. List all events, their location, and the total value of the donations received for each event, in descending order of the total value received.
3. How many donors participated in each event?
4. Which events received no donations?
5. Which club members participated in more than three events?
6. List each event that has not taken place yet and the names and phone numbers of each member who is working on it.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_