COSC 3P32 – Introduction to Database Systems
Winter 2020
Assignment #2

Due Date: 3rd March, 2020, noon Late Date: 6th March, noon
This assignment accounts for 5% of your final grade and is worth a total of 50 marks.
This assignment is to be completed *individually*.

All of the questions in this assignment use the following database schema, in which the keys of each relation are underlined:

Movie(*title*, *year*, *director*, *budget*, *earnings*)
Actor(*stagename*, *realname*, *birthyear*)
ActedIn(*stagename*, *title*, *year*, *pay*)
CanWorkWith(*stagename*, *director*)

Note: the CanWorkWith relation stores information on which actors and directors are able to work with one another – this is important as occasionally an actor will absolutely refuse to work with a given director (or vice-versa).

Question 1 [8 marks]
For each of the following pairs of queries, determine whether or not the queries are equivalent. You **must** explain your answer. Think carefully – to be equivalent, the queries must provide exactly the same answer for **every** possible set of valid data.

a) [2 marks]

(i) $\pi_{\text{stagename}}(\sigma_{\text{pay} < \text{earnings - budget}}(\text{ActedIn} \bowtie \text{Movie}))$

(ii)
```
SELECT   A.stagename
FROM     ActedIn A, Movie M
WHERE    A.title = M.title AND A.year = M.year AND
         A.pay < M.earnings – M.budget
```

b) [3 marks]

(i) ρ (M1(title $\rightarrow$ t1, year $\rightarrow$ y1, director $\rightarrow$ d1, budget $\rightarrow$ b1, earnings $\rightarrow$ e1), Movie)
ρ (M2(title $\rightarrow$ t2, year $\rightarrow$ y2, director $\rightarrow$ d2, budget $\rightarrow$ b2, earnings $\rightarrow$ e2), Movie)
$\pi_{t1,y1}$ (M1 $\bowtie_{e1 \geq e2}$ M2)

(ii) ρ (M1(title $\rightarrow$ t1, year $\rightarrow$ y1, director $\rightarrow$ d1, budget $\rightarrow$ b1, earnings $\rightarrow$ e1), Movie)
ρ (M2(title $\rightarrow$ t2, year $\rightarrow$ y2, director $\rightarrow$ d2, budget $\rightarrow$ b2, earnings $\rightarrow$ e2), Movie)
$\pi_{t1,y1}$ M1 – $\pi_{t1,y1}$ (M1 $\bowtie_{e1 < e2}$ M2)

c) [3 marks]

(i)
```
SELECT   C.stagename, COUNT(C.director)
FROM     CanWorkWith C
WHERE    C.stagename IN (  SELECT      A.stagename
                           FROM        Actor A     )
GROUP BY C.stagename
```

(ii)
```
SELECT   A.stagename, COUNT(C.director)
FROM     Actor A, CanWorkWith C
WHERE    A.stagename = C.stagename
GROUP BY A.stagename
```

Question 2 [14x3 marks]

Express each of queries (a)-(f) in both (i) Relational Algebra, and (ii) SQL:

a) Find the titles and years of movies in which actor(s) with real name "Issur Demsky" acted.
b) Find the stage names of actors with whom every director (who has made a movie) can work.
c) Find the stage names of the actors with the highest pay for acting in a movie.
d) Find the stage names and real names of actors who have never acted in a movie that has made a profit (i.e. in which earnings > budget).
e) Find the titles of movies that have been used more than once (e.g. there are many movies titled "A Christmas Carol" made in different years).
f) Find all pairs of stage names (renamed *stagename1* and *stagename2*) such that the actor with *stagename1* acted in the same movie as the actor with *stagename2*, but earned more for acting in that movie.

Express queries (g) and (h) in SQL:

g) For each actor who acted in a movie in 2019, find their stage name, their year of birth and their total pay in all movies in which they have acted (i.e. including those *not* in 2019).
h) For each director who has made at least 10 movies, find the total number of actors who have acted in a movie directed by that director.

Submission Requirements:

1. Your assignment must be placed in an envelope in the COSC 3P32 assignment box.
2. You must attach a cover sheet, completely filled out, to the envelope. This cover sheet is available from http://www.cosc.brocku.ca/forms/cover. Your assignment will not be marked unless one is submitted *with the assignment*.