CS145 Written Assignment #4

Due Thursday May 6*

NOTE: *We will post the sample solution for selected parts of Problem #1 by 7:00 PM on Monday April 26. You do not need to turn anything in for these parts; nevertheless, you should complete these parts yourself since they will help you to prepare for the midterm exam.*

1. Once again, here is our movie database with three tables:

```
MovieDirector(title, director, year)
MovieCast(title, actor, salary)
MovieReview(title, reviewer, score)
```

All attributes are strings except `year`, `salary`, and `score`, which are integers. Write the following queries in SQL. To save you some writing, you may use the abbreviated schema `MD(T,D,Y)`, `MC(T,A,S)`, and `MR(T,R,S)`. Write your queries so that the final answers will never include duplicates, but add the keyword `DISTINCT` *only* in cases where duplicates would otherwise be produced.

(a) † Using subqueries in `WHERE`, find the directors who have also acted. Sort their names alphabetically.

(b) Without using subqueries in `WHERE`, find the actors who have never directed.

(c) † Find the titles of all movies in the 1960's starring both Dustin Hoffman and Anne Bancroft.

(d) Without using `EXCEPT`, find the actors who starred in every movie directed by Kurosawa.

(e) † Using SQL aggregates, find the titles of the movies with at least two reviews.

(f) Same as (e), but use subqueries in `WHERE` instead of aggregates.

(g) Same as (e), but do not use subqueries in `WHERE` or aggregates.

(h) † Using SQL aggregates, find the titles of the movies with exactly one review.

(i) Same as (h), but do not use aggregates.

(j) † For each reviewer, find the average scores of all his/her reviews. However, do not list reviewers who have given out scores less than 1.

(k) Let's define the rating of a movie to be the average score given to this movie by reviewers. Find title, rating, and year of production for each movie. Movies with higher ratings should be listed first, and among the movies with the same rating, newer movies should be listed first.

(l) Let's define the rating of a director to be the average rating of all his/her movies. Find the names and ratings of all directors who have directed movies starring Al Pacino.

(m) † Is the following a correct solution for (l)? Why?

```
SELECT director, AVG(score)
FROM MovieDirector d, MovieReview r
WHERE d.title IN (SELECT title FROM MovieCast WHERE actor = 'Al Pacino')
AND d.title = r.title;
```

*Please refer to CS145 Course Information Page (`http://www.stanford.class/cs145/info.html`) for submission instructions and late policy.

†Sample solution for these parts will be posted on the course Web page on April 26.

(n) For each actor, find the total salary he/she earned each year during 1995–1999.

2. Consider the same movie database as described in Problem #1. Write SQL data modification statements to complete the following tasks.

   (a) If a movie spends a total of more than $10,000,000 on its cast, put in a review by its director with a score of 10.

   (b) If a director or an actor reviews his/her own movie, set the score to the average score given to this movie.

   (c) For each movie, remove one of its reviews with the highest score. If this movie has more than one reviews with the same highest score, remove only one of them. If this movie has only one review, do not remove the review.