

به نام خدا

تکلیف دوم درس پایگاه داده‌ها ۱

ترم اول ۹۸-۹۹

توجه مهم: این فایل را یک بار دیگر در لحظه آخر و هنگام ارسال بخوانید. عدم رعایت شرایط ذکر شده در نام گذاری یا ارسال فایل‌های پراکنده یا هر چیزی غیر از یک فایل PDF، یک فایل SQL باعث کاهش تمام یا بخشی از نمره شما می شود.

فرمت ارسال: کل پاسخ ها را در قالب یک فایل Zip روی سامانه قرار دهید. پاسخ بخش دوم سوالات باید فقط شامل قسمت هایی باشد که تحت عنوان «**خروجی**» در زیر هر سوال مشخص شده است. نام فایل با شماره دانشجویی شروع شود و سپس فامیل (انگلیسی). مثال: 9511113-Irani.pdf

نحوه ارسال روی سامانه: وارد سامانه lms.iut.ac.ir شوید و فایل پاسخنامه خود را در قسمت تکلیف دوم قرار دهید. ارسال مستقیم با ایمیل و ... به هیچ وجه پذیرفته نمی شود. لطفاً چند روز قبل از پایان مهلت، از عدم وجود مشکل در قسمت آپلود تکالیف مربوط به خودتان روی سامانه مطمئن شوید و اگر مشکلی هست آن را رفع کنید.

توجه: برای انجام این تمرین باید PostgreSQL را دریافت و نصب نموده باشید.

در صورت وجود هرگونه اشکال در نصب PostgreSQL روی سیستم عامل های ویندوز یا لینوکس و سوال در مورد IDE ها میتوانید با TA در ارتباط باشید.

مهلت ارسال پاسخ: سه شنبه ۲۱ آبان ۹۸ ساعت ۲۲:۳۰

(تحويل صرفاً از طريق سامانه انجام خواهد شد و هیچ روش دیگری پذیرفته شده نیست)

بخش اول

۱- با در نظر گرفتن پایگاه داده University که در کلاس درس مطرح شد، به سوالات زیر پاسخ دهید:

الف) چرا در پرس و جوی زیر با JOIN کردن جدول Section در قسمت FROM، تغییری در نتیجه حاصل نمیشود؟

```
SELECT year, semester, sec_id, course_id, AVG(tot_cred)
FROM takes NATURAL JOIN student
WHERE year = ۲۰۰۹
GROUP BY year, semester, sec_id, course_id
HAVING COUNT(ID) >= ۲;
```

ب) آیا خروجی دو پرس و جوی زیر تفاوت دارد؟ چرا؟

```
SELECT name, title
FROM instructor NATURAL JOIN teaches, course
WHERE teaches.course_id= course.course_id;
```

```
SELECT name, title
FROM instructor NATURAL JOIN teaches
NATURAL JOIN course;
```

ج) تحت چه شرایطی خروجی پرس و جوی زیر رکورد هایی با مقدار NULL به ازای فیلد title خواهد داشت؟

```
SELECT * FROM student NATURAL FULL OUTER JOIN takes
NATURAL FULL OUTER JOIN course
```

۲- الف) تعریف توابع تجمعی sum ، count ، min و max برای دو رابطه S_1 و S_2 مشخص است. نحوه محاسبه این توابع را برای رابطه $S_1 \cup S_2$ بیان کنید.

ب) فرض میکنیم رابطه $r(A, B, C, D, E)$ را داشته باشیم؛ همچنین فرض میکنیم T یک خاصیت از این رابطه باشد. (یعنی برای مثال داشته باشیم $T = (A, B, C, D)$) به این ترتیب که حاصل تمامی توابع تجمعی بیان شده در قسمت قبل و همچنین میانگین و انحراف معیار را برای T در جدولی دیگر با عنوان aggregate-on- T داشته باشیم. با استفاده از عبارت group by روی این جدول، دستورهای SQL بنویسید که دو مقدار میانگین و انحراف معیار را برای زیر رابطه S که $S \subseteq T$ باشد را محاسبه کرده و نمایش دهند.

* راهنمایی: برای محاسبه انحراف معیار روی رابطه S از فرمول زیر استفاده کنید:

$$(S) \text{ انحراف معیار} = \sum_{s \in S} \frac{s^2}{|S|} - avg(S)^2$$

۳- الف) عبارت group by cube(a, b, c, d) را یک بار با استفاده از rollup و یکبار بدون rollup باز نویسی کنید. (در

هر دو مورد نباید از عبارت cube استفاده کنید)

ب) در مورد بافت نگار یا Histogram در علم آمار تحقیق کنید. سپس با در نظر گرفتن رابطه $r(A, B, C, D)$ ، پرس و جویی طراحی کنید که یک بافت نگار از D نسبت به A که مقادیر A را به ۲۰ قسمت مساوی تقسیم میکند نمایش دهد.

(*راهنمایی: هر قسمت باید شامل ۵ درصد از رکورد ها باشد که نسبت به A مرتب شده اند)

ج) دو مثال از عملیات Grouping روی داده ها ذکر کنید که تنها با یک عبارت Group by به همراه Cube و Rollup

قابل انجام نباشد.

د) رابطه r در جدول زیر داده شده است؛ با توجه به آن خروجی پرس و جو را مشخص کنید.

Building	Room_number	Time_slot_id	Course_id	Section_id
ECE_Dept	32	A	DS-17	1
ECE_Dept	30	B	DS-17	2
MEC_Dept	25	A	AD-102	2
MEC_Dept	20	C	ED-74	1
IND_Dept	17	D	OR-41	1
IND_Dept	34	D	MI-38	1

```
SELECT building, room number, time_slot_id, count (*)  
FROM r  
GROUP BY rollup (building, room number, time_slot_id)
```

۴- فرض کنید می‌خواهیم یک سامانه نرم افزاری برای مدیریت کارمندان یک سازمان طراحی کنیم که شعبه‌هایی در شهرهای

مختلف دارد. روابط داده‌های این پایگاه داده به صورت زیر می‌باشد؛

- کارمند (کد پرسنلی، نام، نام خانوادگی، کد شهر محل زندگی)
- شهرها (کد شهر، نام شهر)
- شعبه‌ها (کد شعبه، نام شعبه، کد شهر محل شعبه)
- کارمندان شعبه (کد پرسنلی، کد شعبه، میزان حقوق)
- مدیرگروه‌ها (کد پرسنلی کارمند، کد پرسنلی مدیر)

فرض کنید این شرکت شعبه‌ای با نام «Main Branch» دارد، در همه شعبه‌ها کارمندان به صورت گروهی فعالیت می‌کنند و هر گروه یک مدیر دارد و در سیستم مالی این شرکت، این کارمندان ممکن است در چند شعبه مشغول به کار باشند و از

هر شعبه به صورت جداگانه حقوق بگیرند؛ با توجه به این موضوع و جداول فوق به سوالات زیر پاسخ دهید (در صورتی که برای

حالتی خاص پیشفرضی را در نظر می‌گیرید آن را در پاسخنامه تشریحی خود ذکر کنید):

A) Candidate Key و Primary Key مناسب را برای هر جدول مشخص کنید. سپس کد DDL ساخت جداول فوق

را با توجه به دستورات PostgreSQL با ذکر مفروضات و رعایت تمامی Constraint ها و انواع داده‌ای مناسب و

لازم بنویسید.

(B) پرس و جویی طراحی کنید که نام و نام خانوادگی کارمندانی که با مدیر گروهشان در یک شهر زندگی می کنند را نمایش دهد.

(C) پرس و جویی طراحی کنید که تعداد کارکنان و مجموع حقوق های آن ها را به تفکیک هر شهر نمایش دهد.

(D) پرس و جویی طراحی کنید که نام و نام خانوادگی همه کارمندانی که در شعبه Main Branch مشغول به فعالیت نیستند را نمایش دهد.

(E) پرس و جویی طراحی کنید که نام و نام خانوادگی همه کارمندانی را که از تک تک کارمندان شعبه Main Branch کمتر حقوق میگیرند را مشخص کند.

(F) پرس و جویی طراحی کنید که نام و نام خانوادگی مدیران مشغول در شعبه هایی را که میانگین حقوق کارمندان از میانگین حقوق کارمندان شعبه Main Branch بیشتر است را به همراه نام شعبه محل فعالیت آن ها نمایش دهد.

(G) دستور DML بنویسید که به حقوق مدیر گروه های شعبه Main Branch به میزان ۵ درصد از حقوق فعلی شان اضافه شود. اگر با این عمل میزان حقوق آن ها بیشتر از ۳ میلیون تومان می شود، به جای ۵ درصد، ۲ درصد به حقوق آن ها اضافه شود.

(H) آیا می توان جدول مدیر گروه ها را حذف نمود ولی همچنان مدیرگروه ها را در پایگاه داده مشخص کرد؟ چه روشی پیشنهاد می کنید؟

خروجی: کد SQL پرس و جو ها را به ترتیب در قسمت های Q4-A تا Q4-G در فایل HW2.sql اضافه کنید. توضیح قسمت A و H را در پاسخنامه تشریحی خود وارد کنید.

بخش دوم

برای حل سوال ۵، فایل dvdrental.zip را که در تکالیف قبل ارائه شد مورد استفاده قرار دهید .
در صورتی که از pgAdmin 4 استفاده می کنید در صورت بروز هر گونه مشکل در restore کردن فایل ها می توانید با TA در تماس باشید.

جهت حل تمرین های شماره ۵ از تکالیف باید کد های SQL خود را درون فایل HW2.sql در قسمت مشخص شده برای سوال که به همراه تکالیف و فایل پایگاه داده dvdrental در سامانه بارگذاری شده است وارد کنید و screenshot واضح و خوانا از خروجی کوئری های خود (یا بخشی از خروجی ها) را در پاسخنامه متنی خود قرار دهید. پاسخ به این سوالات به هر نحو دیگری منجر به از دست دادن نمره خواهد شد.

۵- برای هریک از موارد زیر اسکریپت های SQL مناسب بنویسید. ممکن است برای هر مورد بیش از یک پاسخ درست موجود باشد.

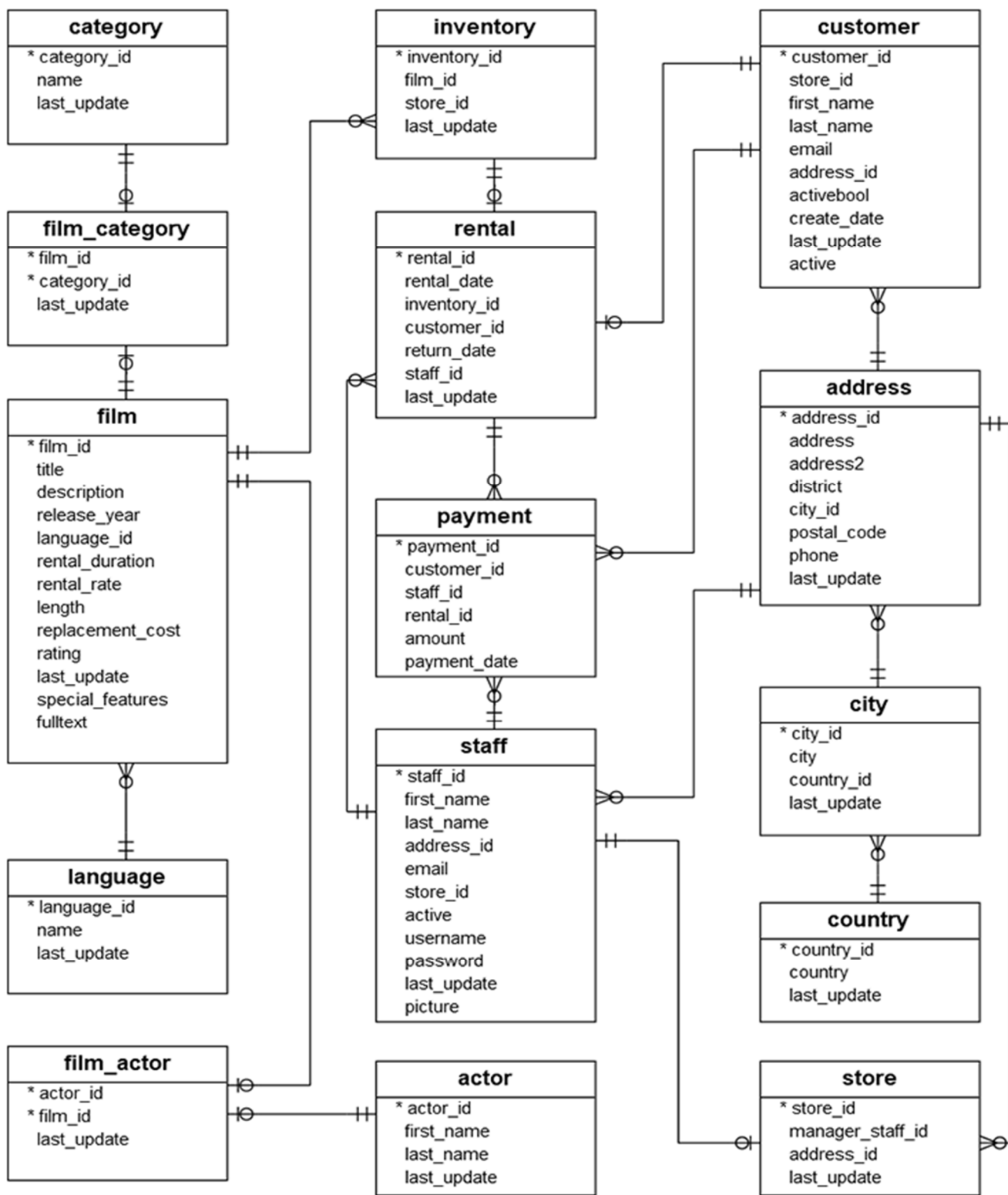
(A) پرس و جویی بنویسید که مشخص کند در هر فروشگاه کدام بازیگر است که بیشترین تعداد فیلم را در آن فروشگاه به خود اختصاص داده است.

(B) پرس و جویی بنویسید که تعداد اجاره فیلم ها و مجموع هزینه پرداختی را برای هر شهر همراه با کشوری که شهر به آن تعلق دارد نمایش دهد ضمن آنکه مجموع تعداد و بهای پرداختی کل اجاره ها را برای تمام شهر ها در یک کشور و در پایان گزارش برای همه ی کشور ها نشان دهد.

(C) پرس و جویی بنویسید که فیلم ها را بر اساس طولشان به ۶ دسته تقسیم کرده سپس در هر دسته بر اساس امتیاز مرتب نموده و مبلغ تجمعی (Cumulative Sum) هزینه جایگزینی فیلم در هر دسته را برای هر رکورد نسبت به ابتدای دسته محاسبه کند و نمایش دهد.

(D) پرس و جویی بنویسید که تعداد فیلم های اجاره شده در هر گروه سنی (rating) را به ازای هر کشور به صورت طراحی زیر نمایش دهد؛

Country	G	PG	PG-13	R	NC-17
---------	---	----	-------	---	-------



شکل ۱ - نمودار ER پایگاه داده dvdrental