

به نام خدا

تکلیف اول 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 و 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 اول 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 از لینک زیر دریافت و نصب نمایید. در بین گزینه‌های نصب بسته pgAdmin 4 را نیز انتخاب نمایید. (نسخه ویندوز ۶۴ بیت)

<https://iutbox.iut.ac.ir/index.php/s/EBY5L3pgaNyaoLD>

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: شنبه ۲۷ مهر ۹۸، ساعت ۲۳:۳۰

(تحويل صرفاً از طریق سامانه انجام می شود و هیچ روش دیگری پذیرفته نخواهد شد)

بخش اول

- ۱- از میان DBMS های رابطه ای مورد استفاده توسعه دهندگان، یک DBMS متن باز (Open-Source) را انتخاب کنید و ضمن بیان نام این DBMS در باره پرسش های زیر برای این DBMS تحقیق کنید:
(الف) این DBMS اولین بار در چه سالی و برای چه سیستم عامل هایی ارائه شد؟
(ب) آیا این DBMS یک رابط کاربری گرافیکی برخوردار است؟ مختصراً شرح دهید. (ارائه یک تصویر مناسب نیز کافیهست)
(ج) تاکنون چه پروژه های معروفی از این DBMS برای مدیریت داده های خود استفاده کرده اند؟

د) ترجیح توسعه دهندگان برای استفاده از این DBMS با چه زبان های تولید و توسعه نرم افزاری بیشتر بوده است؟ آیا این استفاده محدود به چارچوب کاری (Framework) خاصی در این زبان ها است؟

۲- الف) ایجاد سطوح مختلف انتزاع در سیستم توسط توسعه دهنده به چه منظور است؟ با ذکر مثال توضیح دهید.
ب) تفاوت میان استقلال داده ها در سطوح فیزیکی و منطقی انتزاع را بیان کنید.
ج) دو روش طراحی دیتابیس را نام برده و هر کدام را با توجه به تفاوت دیدگاه در بکار گیری آن ها مختصراً شرح دهید.

۳- الف) چهار مورد از معایب استفاده مستقیم از فایل ها جهت ذخیره سازی داده ها را بیان کنید.
ب) آیا برای استفاده از پایگاه داده های رابطه ای نیز عیبی متصور است؟ اگر جوابتان مثبت است مختصراً شرح دهید.

۴- تفاوت بین Super key، Candidate key و Primary key را به طور خلاصه بیان کنید.

۵- دو دلیل برای اینکه چرا مقادیر NULL باید در پایگاه داده معرفی شوند بیان کنید .

۶- نشان دهید در SQL دو عبارت ALL > و NOT IN عملکرد یکسانی دارند.

۷- با توجه به پایگاه داده University که در کلاس درس مورد بررسی قرار گرفت، به سوالات زیر پاسخ دهید؛

الف) توضیح دهید کوئری زیر در چه حالتی با خطا مواجه می شود؟ نوع خطا را نیز مشخص کنید.

```
SELECT name FROM instructor
```

```
WHERE salary * ۱۰ > (SELECT budget FROM department
```

```
WHERE department.dept_name = instructor.dept_name)
```

ب) پرس و جوی زیر را در نظر بگیرید:

```
SELECT S.ID, S.name
```

```
FROM student AS S
```

```
WHERE NOT EXISTS ((SELECT course id
```

```
FROM course
```

```
WHERE dept_name = 'Biology')
```

```
EXCEPT
```

```
(SELECT T.course id
```

```
FROM takes AS T
```

```
WHERE S.ID = T.ID));
```

بیان کنید این پرس و جو به چه منظور نوشته شده است؟ آیا میتوان این پرس و جو را با استفاده از عبارت ALL بازنویسی

کرد؟ اگر جوابتان مثبت است، پرس و جو را با این عبارت بازنویسی کنید و اگر منفی است علت را مختصراً شرح دهید.

بخش دوم

برای حل سوالات ۸، ۹ و ۱۰، فایل dvdrental.zip را از صفحه همین تکلیف در سامانه دریافت نمایید و دیتابیس را روی PostgreSQL ایجاد کنید:

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 ۹ و ۱۰ از تکلیف باید کد های SQL خود را **درون فایل hw1.sql** در قسمت مشخص شده برای سوال که به همراه تکلیف 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.

۸- در مورد مدل داده ای پایگاه داده dvdrental و نمودار ER آن (شکل ۱) به سوالات زیر پاسخ دهید:

الف) کدام جدول(ها) شامل اطلاعات پایه هستند که معمولاً در ابتدای کار موسسه وارد می شوند؟ علت انتخاب (های) خود را توضیح دهید.

ب) داده های کدام جدول(ها) در طول کار روزانه موسسه مرتب رشد می کنند؟ علت انتخاب (های) خود را بیان کنید.

۹- دستور زیر را در PostgreSQL اجرا کرده و تصویر خروجی آن را دقیقاً در پاسخنامه خود کپی کنید؛

```
SELECT pg_postmaster_start_time(), version(), pg_conf_load_time(), current_database();
```

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

A) کوئری برای پایگاه داده dvdrental طراحی کنید که لیستی شامل نام، نام خانوادگی و شهر محل سکونت تمامی مشتریان ایرانی سیستم را در خروجی نمایش دهد.

B) لیست تمام بازیگرانی که در فیلم های مشاهده شده توسط ایرانیان نقش آفرینی کرده اند را استخراج کنید.

(C) مدیر فروشگاه می‌خواهد لیستی از نام، نام خانوادگی مشتریان ایرانی که حداقل یکی از فیلم‌های اجاره شده شان را همان روز که اجاره کرده‌اند به فروشگاه برگردانده‌اند، مشاهده کند. کوئری بنویسید که مشخصات خواسته شده به اضافه زمان اجاره و زمان تحویل فیلم‌ها به فروشگاه توسط هر فرد را به نمایش بگذارد.

(D) لیستی از نام و نام خانوادگی بازیگرانی که در فیلم‌هایی با امتیاز بیشتر از ۴ بازی کرده‌اند یا طول فیلم‌هایی که در آن بازی کرده‌اند بیشتر از ۱۰۰ دقیقه بوده است تهیه کنید.

(E) کدام ژانر ها (Category ها) هستند که شامل فیلم‌هایی میشوند که در لیست موجودی (inventory) فروشگاه شماره ۱ وجود ندارند ولی در فروشگاه شماره ۲ به مدت بیش از ۹ روز اجاره شده‌اند؟

(F) لیستی از فیلم‌هایی تهیه کنید که از همه فیلم‌هایی که عنوان آن‌ها به حرف g ختم میشود یا حداقل دو حرف s در عنوان آن‌ها وجود دارد، طولانی‌ترند. (توجه: ممکن است این حروف به صورت Uppercase یا Lowercase در نام فیلم‌ها به کار رفته باشد که طبیعتاً در هر دو صورت نام فیلم باید در لیست قرا گیرد)

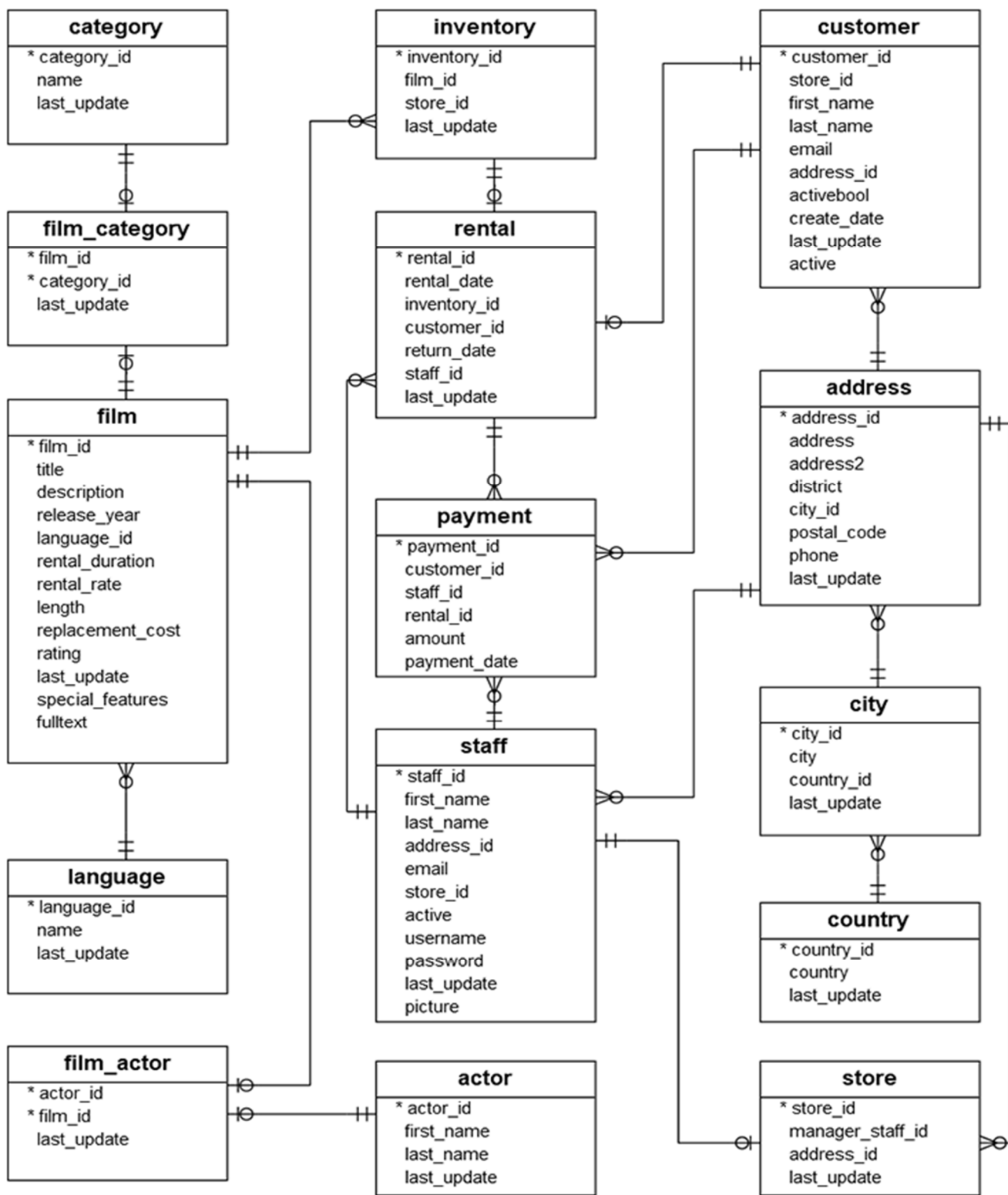
(G) مشتریان را به ترتیب تعداد خریدی که از هر کدام از فروشندگان های شماره ۱ و ۲ داشته‌اند به ازای هر فروشنده لیست کنید؛ (این لیست باید شامل تعداد خرید هر مشتری از هر کدام از فروشندگان‌ها نیز باشد)

(H) یک جدول به نام category_rating ایجاد کنید و در آن لیست ژانر‌ها را به همراه میانگین امتیاز مشتریان به فیلم‌های هر یک از ژانر‌ها و طول بلندترین فیلم آن ژانر ذخیره کنید.

(I) به جدول category_rating یک ستون با عنوان age_group اضافه کنید که مقادیر آن برای هر ژانر به این صورت باشد:

- اگر بیشترین تعداد فیلم‌های ژانر از رده سنی G بود در ستون age_group مقدار ۱ ذخیره شود
- اگر بیشترین تعداد فیلم‌های ژانر از رده سنی PG بود در ستون age_group مقدار ۲ ذخیره شود
- اگر بیشترین تعداد فیلم‌های ژانر از رده سنی PG-13 بود در ستون age_group مقدار ۳ ذخیره شود
- اگر بیشترین تعداد فیلم‌های ژانر غیر از موارد فوق بود در ستون age_group مقدار ۴ ذخیره شود

J) تمامی ژانر هایی را که میانگین طول فیلم هایشان از میانگین طول فیلم های ۳ ژانری که بیشترین میانگین امتیاز فیلم ها را دارند، بیشتر است از جدول category_rating حذف کنید. **توجه:** برای این سوال نیاز است تصویر نهایی از تمامی رکوردهای موجود در جدول category_rating را در پاسخنامه متنی خود وارد کنید.



شکل ۱ - نمودار ER پایگاه داده dvdrental