

به نام خدا

تکلیف پنجم درس پایگاه داده‌ها ۱

ترم مهر ۹۹

فرمت ارسال : کل پاسخ ها را در قالب یک فایل zip روی سامانه قرار دهید که شامل دو فایل باشد، یکی با فرمت pdf حاوی پاسخ شما به سوالات، و دیگری یک اسکریپت، که باید داخل فایل پیوست با نام `YourFamily-HW5.sql`، طبق آنچه برای سوالات توضیح داده شده قرار گیرد. نام فایل zip با شماره دانشجویی شروع شود و سپس فامیل (انگلیسی). مثال: `9511113-Irani.zip`

تحويل سوال آخر (گروهی): طبق توضیحات فقط توسط یکی از اعضای گروه با همین تکلیف آپلود شود.

نحوه ارسال روی سامانه : وارد سامانه yekta.iut.ac.ir شوید و فایل پاسخنامه خود را در قسمت تکلیف پنجم قرار دهید.

توجه: ارسال مستقیم با ایمیل و پیام‌رسان به استاد یا همکاران درس نادیده گرفته خواهد شد.

مهلت ارسال پاسخ: جمعه ۲۱ آذر ۹۹، ساعت ۱۳:۰۰

(تحويل صرفاً از طریق سامانه انجام می‌شود و در پایان ترم نیز نمرات از روی سامانه به گلستان منتقل می‌شود و ارسال دستی قابل ورود نیست)

توجه: برای سوالات این تکلیف از پایگاه `adventureworks` روی `PostgreSQL` که بکاپ آن روی `iutbox` در پوشه `dataset` قرار داده شده است و همچنین پایگاه دانشگاه مربوط به کتاب (نسخه large) استفاده کنید.

هدف از این تکلیف آشنایی بهتر شما با امکانات `Ranking/Windowing/OLAP` است. همچنین در سوال آخر یک کار

تحقیقی-اجرایی را در قالب گروه‌های ۳ الی ۵ نفره انجام خواهید داد.

نحوه تهیه خروجی: (الف) برای سوالات ۱ تا ۳ که مانند قبل بصورت انفرادی است، برای هر کدام، متن اسکریپت دستورات را با نامهایی به صورت نام خانوادگی شما به اضافه `Q1a`, `Q1b` و ... در فایل `sql` کلی و تصویر نتایج را با نامهای `R1a`, `R1b` و ... و تصویر پیامهای موفقیت یا خطا را با نامهای `M1a`, `M1b` و ... در فایل‌های pdf کلی پاسخ خود قرار دهید.

(ب) پاسخ سوال آخر فقط توسط یکی از اعضای گروه و با فایل‌هایی به نام `Q4xxx` آپلود شود. داخل فایلها نام اعضای گروه را بنویسید.

نکته مهم در نامگذاری (در صورت عدم رعایت از آن سوال نمره نمی‌گیرید): هر فایل و اسکریپت باید با نام خانوادگی شما شروع شود، یا داخل آن نام خود را در بالای فایل (مثلاً بصورت کامنت در کد) نوشته باشید.

۱) با استفاده از پایگاه داده `university` به سوالات زیر پاسخ دهید.

a. دستوری بنویسید که برای هر استاد، اولین حقوق کمتر از حقوق خودش را در دپارتمان خودش پیدا کند. (خروجی دارای پنج ستون شامل `id`، نام، نام دپارتمان، حقوق و اولین حقوق کمتر در دپارتمان است) راهنمایی: برای حل این سوال کاربرد تابع `lag` را مطالعه فرمایید.

b. دستوری بنویسید که برای هر دانشجو، رتبه دانشجو در کل دانشگاه را بر اساس تعداد واحد گذرانده، برگرداند. (دانشجویی که تعداد واحد بیشتر گذرانده رتبه بهتری می‌گیرد)

۲) جدولی به نام turn_over را در نظر بگیرید که تراکنش های حساب های بانک را در خود ذخیره میکند. دستوری بنویسید که جدولی به شکل factdeptn را برگرداند که حاوی فیلد مانده بعد از تراکنش است (فرض کنید مانده حساب قبل از این تراکنشها صفر بوده است).

Turn_Over			Factdeptn			
Dep_Id	Trn_Time	Trn_over	Dep_Id	Trn_Time	Trn_over	Balance
1022	2018-06-15 14:00	100	1022	2018-06-15 14:00	100	100
1022	2018-06-15 14:28	-50	1022	2018-06-15 14:28	-50	50
1022	2018-06-16 14:58	25	1022	2018-06-16 14:58	25	75
1067	2019-07-18 23:32	300	1067	2019-07-18 23:32	300	300

۳) با استفاده از دیتابیس dvdrental به سوالات زیر پاسخ دهید:

- دستوری بنویسید که برای هر پرداخت مشتری، میانگین کل پرداختی ها و جمع پرداختی کل پس از آن پرداخت را محاسبه کند.
- مشتریان را براساس مجموع پرداختی هایشان به ۴ گروه تقسیم کنید و لیست نام مشتریان برای برترین گروه را برگردانید.
- دستور نوشته شده در قسمت (a) را فقط برای جمع پرداختی کل پس از آن پرداخت بدون استفاده از توابع window بنویسید (راهنمایی: با استفاده از join)
- دستوری طراحی کنید که تعداد مشتریان و تعداد سفارشات آنها را به تفکیک هر شهر و کشور نمایش دهد.
- دستوری بنویسید که تعداد کل فیلم ها و تعداد فیلم ها به تفکیک ژانر و امتیاز و به تفکیک امتیاز برگرداند.
- دستوری بنویسید که جمع کل پرداختی ها را به تفکیک همه جایگشت های شهر و تاریخ پرداخت برگرداند.

هدف از سوال آخر، آشنایی بهتر شما با جزئیات بیشتری از شاخص TPC-H است که توسط این موسسه برای ارزیابی DBMS های مختلف روی بسترهای مختلف از نظر سناریوهای OLAP/Decision Making طراحی شده است. همچنین شما با یک مهارت مهم در حوزه پایگاه که نگاشت از یک ساختار به ساختار دیگر است (Schema Mapping) آشنا می شوید. نحوه تهیه خروجی: برای هر کدام، متن اسکریپت دستورات را با نامهایی به صورت Q1a, Q1b و ... در فایل sql کلی و تصویر نتایج را با نامهای R1a, R1b و ... و تصویر پیامهای موفقیت یا خطا را با نامهای M1a, M1b و ... در فایل های pdf کلی پاسخ خود قرار دهید.

- ۴) مشخصات آزمون TPC-H روی سایت آن در دسترس است که در فایل tpc-h-specification برای شما قرار داده شده است.
- اطلاعات کلی این آزمون را در صفحات ۸ الی ۱۲ مطالعه نمایید و هدف از این آزمون و همچنین تفاوت OLTP با OLAP را بنویسید. پاسخ این قسمت در حدود نصف یک صفحه A4 کافی است.
 - ساختار پایگاه TPC-H را در شکل صفحه ۱۳ بررسی کنید و با توجه به شباهتهای آن با جداول مربوطه از پایگاه AdventureWorks، فایل Word پیوست با نام Schema-mapping را برای کلیه جدولهای TPC-H کامل کنید.

c. آزمون TPC-H شامل ۲۲ مورد Query است که از Q1 تا Q22 در فایل پیوست توضیح داده شده اند. با همفکری تیم خودتان، سه مورد از Query ها که روی AW قابل اجرا باشند را انتخاب نموده و آنها را روی AW بازنویسی و اجرا نمایید.