



اصول طراحی پایگاه داده

مدرس: دکتر عین الله خنجری - بهار ۱۴۰۲-۱۴۰۳ - کمک مدرس: آرمان زرکلام

فاز چهارم: ساخت جداول، وارد کردن و تغییر داده ها+کوئری

مهلت تحویل: ۱۸ خرداد ۱۴۰۳

نحوه تحویل: بارگزاری در کوئرا (توسط هر دو عضو تیم)

انجام در گروه های دو نفره + مشورت با منتور

گام اول: جمع آوری و وارد کردن داده ها

برای تمام جداولی که در فاز قبل (جداول رابطه ای) داشتید، یک فایل `TableName.csv` بسازید و حداقل ۱۵ ردیف داده برای هر کدام وارد کنید. سعی کنید حتی الامکان داده هایی که وارد می کنید به واقعیت نزدیک باشند و ضمناً تمایز هایی با هم داشته باشند که در گام های بعدی بتوان روی ستون های مختلف شان کوئری اعمال کرد و نتایج را دید.

به جهت وارد کردن داده های فرضی، برای صفت های کلی مثل اسم، سن، نام کاربری، ایمیل، تاریخ ها، اعداد و ... از سایت هایی که داده های رندوم تولید می کنند (سرچ کنید `Random data generator`) می توانید استفاده کنید. همچنین برای موارد تخصصی تر که به موضوع پروژه تان مربوط می شود، می توانید دیتاست های واقعی که در اینترنت وجود دارد (`Kaggle.com` و `Data.world` را نگاه کنید) را بررسی کرده و از آنها داده هایی که نیاز دارید را پیدا و وارد کنید.

دیتاست هایی که در این گام ساخته اید را در پوشه `Datasets` قرار دهید.

گام دوم: ساختن جداول با Create Table :

در هر برنامه ای که کار می کنید، یک دیتابیس جدید برای پروژه ایجاد کنید. قبل از اینکه دیتاست های ساخته شده در گام قبلی را ایمپورت کنید، با استفاده از دستور Create table ، هر کدام از جدول هایتان را به طور جداگانه بسازید. دیتاتایپ ستون ها را خودتان به درستی مشخص کنید و Primary Key و Foreign Key های جدول را هم تعریف نمایید. همچنین برای هر Foreign Key مشخص کنید که برای حالات On Delete و On Update چه رویکردی را پیش می گیرید. ضمناً اگر نیاز است که صفتی به صورت Not Null تعریف شود، این کار را انجام دهید. (دقت کنید اگر جدولی به جدول دیگر Foreign Key دارد، ابتدا باید جدول مقصد را تعریف کرده باشید تا به ارور نخورید)

در نهایت دستوراتی که با اجرای آن جدول ها را ساخته اید، در فایلی به نام Create_tables.sql ذخیره نمایید.

گام سوم: وارد کردن داده های گام اول با دستور Insert into :

در این مرحله، باید دیتاست های با فرمت CSV. که ساخته بودید را ایمپورت کنید (با نام هایی که با جداول اصلی تان اشتباه نشود) و با دستور Insert into داده هایشان را در ستون های جداول متناظری که ساخته بودید اضافه کنید. در نهایت پس از اتمام این کار، آن جدول هایی که از CSV ایمپورت شده بودند را پاک کنید چون دیگر نیازی به آنها نیست.

در انتهای این گام، دستوراتی که با آنها داده ها را insert کرده اید را در فایلی به نام Insertions.sql ذخیره کنید.

توجه: شاید در این گام برایتان این سوال پیش بیاید که چرا از همان اول فقط دیتاست ها را ایمپورت نکردیم و مثل سوالات تمرین عملی، از همان ها پیش نرفتیم؟ چه نیازی بود که جدولی خالی هم بسازیم و از یکی در دیگری داده بریزیم؟ پاسخ این است که در واقعیت، ابتدا جداول داده ساخته می شوند و بعد داده ها به مرور و پس از عرضه ی محصول در آنها جای می گیرد. پس این که اول جداول خالی را به درستی بسازیم و بعد داده ها اضافه شوند، به واقعیت نزدیک تر است. و نکته مهم (و جایی که ممکن است موقع insert دردرسر بکشید) اینجاست که دیتاتایپ جداول را به درستی و مطابق داده هایی که در آینده اضافه می شوند پیش بینی کرده باشیم.

گام چهارم: کوئری

در این گام، شما باید تعداد ۲۰ کوئری را با ایده خودتان روی جداول اعمال کنید و هر کدام را در فایلی به نام `Query[number].sql` ذخیره کرده و در پوشه `Queries` قرار دهید. دقت کنید که انتخاب کوئری ها با خودتان است اما سعی کنید آنها معنی دار باشند و گزارشی به دردبخور از داده ها را بخواهند. ضمناً صرفاً روی ستون های یک جدول نباشند و در حداقل نصف آنها (۱۰ کوئری)، دو جدول یا بیش تر با هم جوین خورده باشند.

گام پنجم: حذف و به روز کردن داده ها

در فایل هایی به نام `Delete1.sql` تا `Delete3.sql`، دستوراتی بنویسید که به ترتیب ستون، ردیف و سلولی از یکی از جدول هایتان را پاک کند. همچنین در فایل هایی به نام `Update1.sql` تا `Update3.sql`، دستوراتی بنویسید که به ترتیب ستون، ردیف و سلولی از یکی از جدول هایتان را با مقداری جدید بروزرسانی کند.

در نهایت همه موارد خواسته شده در گام ها را زیپ کرده و در کوئرا بارگزاری کنید.