

پروژه نهایی پایگاه داده

یزدان زندیه وکیلی

ندا صدری

سپهر عباس پور

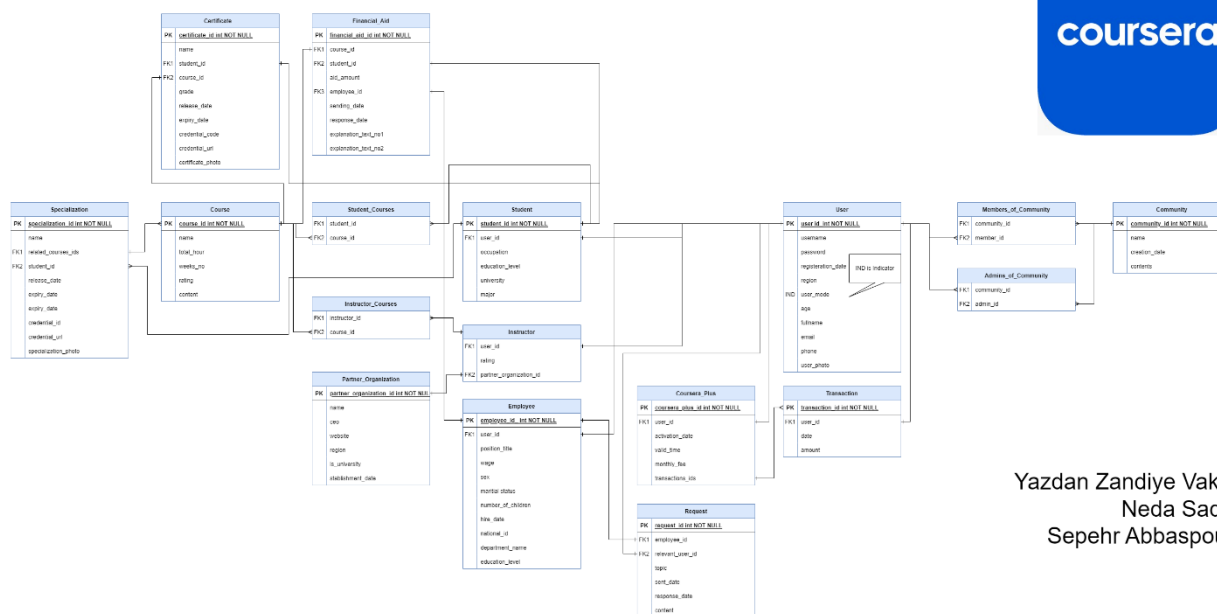
رضا تاج گذاری

مقدمه

Coursera یک پلتفرم آموزش آنلاین است که امکان دسترسی به دوره‌ها و برنامه‌های تحصیلی ارائه شده توسط دانشگاه‌ها و موسسات آموزشی برجسته در سراسر جهان را فراهم می‌کند. این وبسایت با همکاری دانشگاه‌هایی چون استنفورد، دانشگاه میشیگان، و دانشگاه دوک، دوره‌های متنوعی در زمینه‌های مختلف مانند علوم کامپیوتر، کسب و کار، هنرهای زیبا، علوم انسانی و بیشتر ارائه می‌دهد. دوره‌ها معمولاً شامل ویدئوهای آموزشی، تکالیف، امتحانات و پروژه‌های عملی هستند و بسیاری از آن‌ها گواهینامه‌هایی را به شرکت‌کنندگان اعطا می‌کنند. Coursera همچنین برنامه‌های درجه‌بندی شده مانند مدرک کارشناسی ارشد یا تخصصی را نیز ارائه می‌دهد که می‌تواند فرصت‌های شغلی بیشتری را برای فراگیران فراهم آورد. این پلتفرم برای کسانی که به دنبال یادگیری مادام‌العمر و ارتقاء مهارت‌های خود هستند، یک گزینه ایده‌آل است.

فاز اول

در این فاز از پروژه لیست موجودیت‌های پروژه استخراج شد و تا سطح سوم نرمال سازی شد که گزارش جدا دارد و در اینجا فقط نمای نهایی را قرار می‌دهیم.



Yazdan Zandiyeh Vakili
Neda Sadri
Sepehr Abbaspour

نمای جداول موجود در پایگاه داده Coursera

نکته : همانطور که در گزارش فاز دوم ذکر شده، بعضی موجودیت ها در نهایت حذف شده اند.

فاز دوم

در فاز دوم ER Diagram ای از پایگاه داده به فرمت خواسته شده رسم شد و همانطور که می توانید مشاهده کنید، تغییراتی نسب به نسخه قبلی داشت که موجب بهبود کلی طراحی پایگاه داده شد که در گزارش مربوط به فاز دوم این موارد ذکر شده است.

فاز سوم

در نهایت در فاز آخر پروژه باز تغییراتی به وجود آمد که موجب راحتی طراحی و استفاده از این پایگاه داده می شود که به جزئیات به شرح آن می پردازیم.

جداول

۱. User: این جدول را می توان اصلی ترین جدول کل پایگاه داده فرض کرد که خود به سه نوع تقسیم می شود. دانش آموز، مدرس و کارمند که هر کدام جدول جدای خود را دارند ولی هر سه خود به تنهایی یک کاربر نیز هستند، پس کلید اصلی هر کدام از آن ها در واقع کلید خارجی این جدول است. ویژگی های این جدول عبارت است از : `user_id`, `username`, `password`, `refistration_date`, `region`, `age`, `first_name`, `last_name`, `email`, `phone` می باشد که برای مواردی مانند ایمیل و شماره موبایل برای درج `Constraint` ایجاد شده که با فرمت اصلی تطابق داشته باشد.

۲. Student: در این جدول علاوه بر ویژگی های جدول `User` که با استفاده از کلید خارجی دریافت می شود، ویژگی های دیگری مانند `education`, `major`, `university` نیز وجود دارد که برای سطح تحصیل یک `Data Type` طراحی شده که شامل کارشناسی، کارشناسی ارشد و یا دکتری می باشد.

۳. Instructor: علاوه بر ویژگی های جدول `User` دو ویژگی دیگر دارد، امتیاز مربوط به مدرس و موسسه همکار با مدرس (`rating`, `partner_organization`)

۴. Employee: برای این جدول ویژگی های زیاد تری نسبت به مدرس و دانش آموز در نظر گرفته شده است و ویژگی های آن عبارت از جنسیت ، وضعیت تاهل ، تعداد بچه، کد ملی (که برای چک کردن آن Constraint داریم) و است.

۵. Partner_organization: در این جدول موجودیت سازمان های همکار پیاده سازی شده است و ویژگی هایی نظیر وبسایت، نام مدیر عامل ، اسم شرکت، تاریخ تاسیس و ... تعبیه شده است که تنها با موجودیت مدرس ارتباط دارد.

۶. Request: یک جدول برای درخواست هایی است که در سامانه گذاشته می شود و هر کاربر می تواند درخواست بدهد و یک کارمند مربوطه باید به آن پاسخ دهد. ویژگی های مربوط به آن id کارمند مربوطه و هم چنین کاربر درخواست دهنده ، متن درخواست و ... است.

۷. Community: در این موجودیت برای تمامی گروه ها و کلنل هایی که داخل وبسایت موجود است تعبیه شده است که هر کاربری جدای از اینکه چه نقشی دارد می تواند عضو آن باشد و محتوی آن برای ساده سازی فقط در یک ویژگی به نام content ذخیره شده است.

۸. Course: یکی از اصلی ترین موجودیت ها، دوره های آموزشی موجود هستند که شامل ویژگی هایی مانند امتیاز، محتوی، تعداد ساعت و ... هستند که با موجودیت های مدرس و دانش آموزان و Specialization ها در ارتباط اند.

۹. Specialization: حرفه ها در واقع موجودیتی هستند که بعد از گذراندن چندین دوره می توانید دریافت کنید و رابطه چند به یک با موجودیت Course دارد، ولی برای هندل کردن روابط آن دو از جدول کمکی استفاده شده است.

۱۰. Coursera_plus: این درواقع حکم اکانت ویژه را دارد که می توانند کاربر ها تهیه کنند. برای همین داخل آن ویژگی هایی مانند مقدار پولی که گذاشته شده و ... موجود است که داخل جدول تعبیه شده است.

۱۱. Transaction: همانطور که می بینید یک کاربر می تواند یا اکانت ویژه و یا یک دوره را بخرد، پس نیاز به ماهیت تراکنش ها داریم که نام کاربر مربوطه به همراه تاریخ و مقدار مبلغ جا به جا شده و ... را نگه می دارد.

جدول های ارتباطی

ما در این پایگاه داده ۴ جدول ارتباطی داریم که هر چهارتا ساختار یکسانی دارند و دارای کلید مجزا نمی باشند و زوج دو کلید خارجی داخل آن با یکدیگر کلید اصلی را تشکیل می دهند.

۱. Rel_community_member: برای ارتباط بین اعضا و گروه ها

۲. Rel_course_specialization: برای ارتباط بین دوره ها و حرفه ها

۳. Rel_instructo_course: برای ارتباط بین مدرسین و دوره ها

۴. Rel_student_course: برای ارتباط دانشجویان و دوره ها

داده های ورودی

در این بخش ما برای هر جدول متناسب با ویژگی های آن یک کد پایتون مربوط نوشتیم که دستور وارد کردن داده های که درست کردیم را در یک فایل ذخیره کند و در نهایت آن را روی پایگاه داده اعمال کردیم. تمامی کدها و داده های ورودی فایل هایشان موجود است و فرستاده شده است. در این بخش برای بعضی ویژگی ها مانند سن افراد به صورت رندوم مقدار تولید کردیم ولی برای ویژگی هایی مانند اسم و رشته و ... از مقادیر Scrap شده از اینترنت استفاده کردیم که در یک کتابخانه به نام Faker از قبل تعبیه شده بود. به دلیل تولید داده ها به صورت اتوماتیک، تعداد رکورد های زیادی به پایگاه داده اضافه شد، به طور مثال ۳۰۰۰ کاربر داریم که ۲۰۰۰ تای آن ها دانشجوی، ۸۰۰ عدد مدرس و ۲۰۰ عدد کارمند هستند و ۱۰۰۰ عدد دوره داریم و ... و در کل تعداد داده ها سعی شده که برای هر موجودیت به طوری باشد که نسبت در دنیای واقعی در نظر گرفته شده باشد.

فایل های فرستاده شده

۱. یک فایل اصلی طراحی پایگاه داده به نام Coursera
۲. Query های طراحی شده (۱۰ عدد)
۳. کد های تولید اطلاعات
۴. اطلاعات تولید شده
۵. گزارش و عکس فاز اول و دوم
۶. گزارش فاز سوم

نکات مهم

۱. این پروژه با استفاده از PostgreSQL زده شده است
۲. تمامی مراحل ساخت ختی ورود اطلاعاتی که گفته شد در فایل Coursera.sql موجود است و به همین دلیل تعداد خط های کد زیاد شده است.