

## پروژه اول درس پیاده‌سازی پایگاه داده‌ها

### داده‌های دیتابیس:

در فایل ضمیمه داده‌های مربوط به امتیازدهی و تگ‌گذاری برای فیلم‌های سینمایی در یک سرویس پیشنهاد فیلم آمده است. این داده‌ها ۲۵ میلیون رای برای ۶۲ هزار فیلم را در بردارند که توسط ۱۶۲ هزار کاربر امتیازدهی شده‌اند. تمامی کاربران حداقل ۲۰ فیلم را امتیازدهی کرده‌اند. داده‌ها در فایل‌های CSV به قرار زیر وجود دارند:

فایل ratings.csv: حاوی امتیازاتی است که کاربران به فیلم‌ها داده‌اند. امتیازات در قالب سیستم 5-star است.

فایل tags.csv: حاوی تگ‌هایی است که هر کاربر به هر فیلم داده است. هر تگ یا یک تک کلمه است یا عبارتی کوتاه.

فایل movie.csv: حاوی اطلاعاتی در مورد فیلم‌هاست. این اطلاعات شامل عنوان فیلم و ژانرهای مربوطه است. یک فیلم ممکن است بیش از یک ژانر داشته باشد که به وسیله علامت "|" از یکدیگر جدا شده‌اند. ژانرها شامل موارد زیر است:

- Action \*
- Adventure \*
- Animation \*
- Children's \*
- Comedy \*
- Crime \*
- Documentary \*
- Drama \*
- Fantasy \*
- Film-Noir \*
- Horror \*
- Musical \*

Mystery \*

Romance \*

Sci-Fi \*

Thriller \*

War \*

Western \*

فایل‌های scores.csv و tags-relevance.csv: این فایل‌ها میزان مربوط بودن تگ‌های فیلم‌ها را بررسی می‌کند.

### توضیحات پروژه:

برای انجام پروژه جداول مورد نیاز را با توجه به سطوح نرمال سازی در Microsoft SQL Server 2017 بسازید و داده‌های این فایل‌ها را داخل آن جداول قرار دهید.

یک برنامه به هر زبان برنامه نویسی که مسلط هستید بنویسید که با دیتابیس اتصال برقرار کند و اطلاعات مربوط به پرس و جوهای زیر را در فرم‌های مرتبط نمایش دهد:

- ۱- پرامتیازترین فیلم‌ها که تعداد rate آن‌ها بزرگتر از ۵۰ است.
- ۲- در هر سال، عنوان فیلمی که بیشتر از همه rate داشته است.
- ۳- در هر ژانر، فیلمی که rate بزرگتر از 3.0 دارد را مشخص کنید.
- ۴- اختلاف بین میانگین نرخ rate فیلم‌ها قبل از سال ۲۰۰۵ و بعد از آن را به دست آورید.
- ۵- کاربرانی که به تمام فیلم‌ها rate یکسانی داده‌اند را بیابید.
- ۶- در مورد هر فیلم، تگی که بالاترین امتیاز مرتبط بودن با فیلم را دارد چند بار توسط کاربران استفاده شده است.

برای هر پرس و جو در صورت نیاز از indexing یا partitioning استفاده کنید و زمان اجرای آن‌ها را با استفاده از indexing و partitioning و بدون استفاده از آن‌ها به دست آورید.