

به نام خدا

فاز صفر پروژه درس طراحی پایگاه داده

(طراحی یک سیستم ذخیره و بازیابی اطلاعات با استفاده از فایل)



نیمسال دوم ۹۹-۹۸

نام درس	استاد درس
طراحی پایگاه داده	دکتر زارع پور
تاریخ تعریف فاز صفر	مهلت تحویل فاز صفر
۱۳۹۸/۱۱/۳۰	۱۳۹۸/۱۲/۱۵

● مقدمه

همان‌گونه که در جلسات اول کلاس گفته شد، برای طراحی یک سیستم ذخیره و بازیابی اطلاعات دو دیدگاه کلی وجود دارد. دیدگاه اول مبتنی بر فایل و دیدگاه دوم مبتنی بر پایگاه داده‌هاست. هدف از این فاز پروژه این است که شما با طراحی و پیاده‌سازی یک سیستم مبتنی بر فایل، با محدودیت های این دیدگاه آشنا شوید.

● اطلاعات مسئله

فرض کنید که از شما خواسته شده است که یک سامانه برای ذخیره و بازیابی اطلاعات مربوط به یک سینما بدون استفاده از یک سیستم مدیریت پایگاه داده (DBMS) طراحی کنید.

◀ اطلاعاتی که در مورد هر فیلم قرار است ذخیره شود عبارت‌اند از:

۱. شناسه فیلم **FilmID** (یک عدد ۴ رقمی. این ویژگی یکتایی مقدار دارد).
۲. نام فیلم **FilmName** (حداکثر ۱۰۰ کاراکتر)
۳. نام کارگردان **DirectorName** (حداکثر ۱۰۰ کاراکتر)
۴. سال تولید **ProductionYear** (یک عدد ۴ رقمی)
۵. ژانر فیلم **Genre** (حداکثر ۲۰ کاراکتر)

◀ اطلاعاتی که در مورد هر هنرپیشه قرار است ذخیره شود عبارت‌اند از:

۱. شناسه هنرپیشه **ArtistID** (یک عدد ۴ رقمی. این ویژگی یکتایی مقدار دارد).
۲. نام هنرپیشه **ArtistName** (حداکثر ۱۰۰ کاراکتر)
۳. سن هنرپیشه **Age** (یک عدد حداکثر ۳ رقمی)
۴. نام فیلم(هایی) که در آن‌ها نقش‌آفرینی کرده است **ArtistFilms**.

● خواسته مسئله

خواسته مسئله از سه بخش اصلی و دو بخش امتیازی تشکیل شده است:

بخش اول: طراحی و پیاده‌سازی سیستم ذخیره، جستجو، حذف و بروزرسانی

یک سیستم ذخیره و بازیابی اطلاعات مبتنی بر فایل طراحی کنید که قابلیت ذخیره و بازیابی اطلاعات را از طریق یک رابط کاربری (UI) ساده داشته باشد. (اگر از رابط کاربری گرافیکی استفاده کنید، نمره اضافی تعلق خواهد گرفت). عملیاتی که این سیستم باید قادر به انجام آن باشد، عبارتند از:

(۱) ذخیره کردن اطلاعات مربوط به هر فیلم **Film** :

Add FilmID: --- , FilmName: --- , DirectorName: --- , ProductionYear: --- , Genre: ---

صفات در دستور ورودی با کاما از هم جدا می‌شوند و --- مقدار هر صفت است.

نکته: در صورت تکراری بودن با خطا (**Error**) مواجه شویم.

مثال:

Add FilmID: 1123 , FilmName: The Salesman , DirectorName: Asghar Farhadi , ProductionYear: 2016 , Genre: Drama

(۲) ذخیره کردن اطلاعات مربوط به هر هنرپیشه **Artist** :

Add ArtistID: --- , ArtistName: --- , Age: --- , ArtistFilms: ---

نکته: در صورت تکراری بودن با خطا (**Error**) مواجه شویم.

مثال:

Add ArtistID: 2243 , ArtistName: Shahab Hosseini , Age: 46 , ArtistFilms: The Salesman,About Elly,A Separation

(۳) پیدا کردن یک فیلم با استفاده از نام کامل(یا بخشی از آن) و شناسه (**FilmID**):

Find Film --- By FilmName/FilmID

مثال:

Find Film The Salesman By FilmName or
Find Film Salesman By FilmName or
Find Film 1123 By FilmID

2243/Shahab Hosseini/46/The Salesman
2243/Taraneh Alidoosti/36/The Salesman

۴) قابلیت حذف یک هنرپیشه و فیلم توسط شناسه آن:
که --- ArtistID است.

Remove ArtistID ---

Remove FilmID ---

که --- FilmID است.
۵) قابلیت تغییر مشخصات یک هنرپیشه یا یک فیلم:

Update Artist --- Set X to A

Update Film --- Set X to A

این دستور مقدار صفت X مربوط به فیلم یا هنرپیشه با نام کامل --- را به A تغییر می دهد. به طور مثال:

Update Artist Shahab Hosseini Set Age to 47

Update Film The Salesman Set Genre to Thriller

بخش دوم: اطمینان از وجود نام فیلم

در هنگام درج (یا ویرایش) مشخصات یک هنرپیشه با چه مکانیزمی می توان مطمئن شد که نام فیلمی که برای آن هنرپیشه ثبت شده است، قبلاً در لیست فیلم ها حضور داشته است. این مکانیزم را پیاده سازی کنید.

بخش سوم: نمایه سازی

برای افزایش سرعت بازیابی اطلاعات یک نمایه متراکم (Index) روی فیلدهای (شناسه هنرپیشه ArtistIDIndex ، نام هنرپیشه ArtistNameIndex) ایجاد کنید. توجه: در صورت پیاده سازی حداقل یکی از نمایه ها (ترجیحاً شناسه هنرپیشه) به صورت درخت B-Tree یا B-Tree + نمره اضافی تعلق خواهد گرفت.

بخش امتیازی اول: نمایش اطلاعات هنرپیشه و فیلم های مربوطه به او.

یک دستور در سامانه خود طراحی کنید که بتواند اطلاعات زیر در مورد همه فیلم ها را با فرمتی مناسب نمایش دهد. اطلاعاتی که قرار است نمایش داده شوند به شرح زیر است:

نام فیلم، نام هنرپیشه های آن، نام کارگردان، سال تولید، ژانر فیلم. به طور مثال خروجی به شکل زیر باشد:

The Salesman/Shahab Hosseini, Taraneh Alidoosti/Asghar Farhadi/2016/Drama

...

بخش امتیازی دوم: ارتباط بین فایل های فیلم و بازیگر

در این بخش با استفاده از یک فایل جدا از دو فایل دیگر باید نشان داده شود که کدام بازیگر در کدام فیلم بازی می کند. درج داده تنها زمانی ممکن است که بازیگر و فیلم درج شده در دو جدول بازیگر و فیلم موجود باشند. در صورت حذف یا بروزرسانی در فایل مرجع، تغییرات در این فایل نیز اعمال می گردد.

راهنمایی: درج داده در این فایل به صورت مقابل می باشد:

● نکات مهم در مورد فرمت ذخیره سازی اطلاعات

نکته: توجه داشته باشید که فرمت ذخیره سازی اطلاعات در فایل مربوط به فیلم ها و هنرپیشه ها و ارتباط بین آن ها (قسمت امتیازی) به صورت زیر می باشد:

اطلاعات مربوط به هر رکورد به صورت زیر ذخیره می شوند:

۱- مقدار صفت: نام صفت

- ۲- هر دو صفت با "/" از یکدیگر جدا می شوند.
۳- هر دو رکورد با **Enter** از یکدیگر جدا شوند.
۴- برای رکورد ها شماره گذاری انجام دهید. (هنگام اندیس گذاری از اعداد استفاده کنید.) به طور مثال:

۱- رکورد اول
۲- رکورد دوم
....

مثال:

Film.txt

1-1123/The Salesman/Asghar Farhadi/2016/Drama

2-7756/A Separation/Asghar Farhadi/2011/Mystery

3-8911/Shirin/Abbas Kiarostami/2008/Drama

4-3567/The Truman Show/Peter Weir/1998/Drama

Artist.txt

1-2243/Shahab Hosseini/46/The Salesman,A Separation

2-9966/Taraneh Alidoosti/36/The Salesman,Shirin

ArtistFilm.txt

1123/2243

1123/9966

Or

8911/9966

نکات مهم در مورد تحویل تمرین:

- انجام این فاز از پروژه به صورت انفرادی می باشد.
- پاسخ هر سوال باید دقیق و متناسب با سوال باشد. از ذکر مطالب مبهم، نامرتبط و زائد خودداری شود.
- پاسخ ها باید با کلمات خودتان بیان شوند. مطالب منابع دیگر را عینا کپی نکنید. در صورت استفاده از منبع خاصی برای پاسخ به سوال، اسم آن منبع ذکر گردد.
- در صورت کشف تقلب، بر اساس مقررات آموزشی با آن برخورد خواهد شد.
- کلیه محتویات تمرین (مستند، کد منبع، خروجی نرم افزارها و غیره) بایستی در قالب یک فایل فشرده با نام **P0_9XXXXXXX_Fullname** در سایت کوئرا ارسال گردد. برای پاسخ می توانید جواب های خود را بر روی کاغذ نوشته و اسکن کنید.
- تکلیف بایستی فقط یکبار فرستاده شود. در صورت ارسال چندین نسخه در زمان های مختلف، فقط نسخه آخر بررسی می شود. همچنین تاخیر در ارسال پاسخ، طبق ضوابطی معین مشمول کسر نمره خواهد بود.
- در صورت نقض هر کدام از نکات گفته شده، نمره کسر می گردد.
- برای پیاده سازی این سیستم می توانید از زبان برنامه نویسی پایتون یا سی شارپ استفاده کنید. تنها مجاز به استفاده از دستورات استاندارد این دو زبان هستید و مجاز به استفاده از کتابخانه ها و یا ماژول های آماده نیستید.
- برای تست کردن سیستم یک مجموعه ورودی / خروجی استاندارد تهیه و بر روی سایت کوئرا قرار خواهد گرفت که شما می توانید با استفاده از این فایل ها، صحت عملکرد برنامه خود را ارزیابی کنید.
- کدها و فایل های مربوط به این تمرین به صورت یک فایل زیپ شده به همراه یک راهنمای کاربری کوتاه می بایست تا تاریخ مشخص شده از طریق سایت کوئرا بارگذاری شود. در پوشه ارسالی صرفا می بایست فایل های مربوط به این پروژه به ترتیب زیر قرار داشته باشد:

۱- فایل مربوط به کد اصلی برنامه: **main.py** یا **main.cs**

۲- فایل مربوط به اطلاعات هنرپیشه ها: **Artist.txt**

۳- فایل مربوط به اطلاعات فیلم ها: **Film.txt**

۴- فایل مربوط به نمایه (**Index**) روی فایل هنرپیشه ها.

ArtistIDIndex, ArtistNameIndex.txt

توجه: در صورت پیاده سازی نمایه با ساختار **B-Tree** (قسمت امتیازی)، آن را در یک فایل با عنوان **B-treeIndex.txt** ذخیره کنید.

اطلاعات مربوط به **ArtistNameIndex** نیز به صورت زیر ذخیره می شوند:

(که این اعداد شماره فیلم های این هنرپیشه در فایل فیلم می باشد.) **ArtistName1: 1123-7756**

ArtistName2 : 1123-8911

...

نحوه ذخیره سازی اطلاعات مربوط به **ArtistIDIndex.txt** به صورت مشابه می باشد.

- پس از ارسال کدها در زمانی مقرر، برای هر دانشجو زمانی مشخص خواهد شد که کد خود را برای دستیاران آموزشی درس اجرا و تشریح نماید.
- در صورت پیدا شدن کدهای مشابه و یا کپی کردن این فاز از پروژه (تمرین) و یا هر کدام از تمرین های دیگر، بار اول کل نمره تمرین ها و پروژه (۶نمره) صفر خواهد شد و بار دوم نمره کل نمره درس صفر منظور خواهد شد.
- در صورت بروز ابهام در مورد سوالات، می توانید سوالات خود را از دستیاران حل تمرین بپرسید.

ارزش انسان ز علم و معرفت پیدا شود بی هنر گر دعوی بیجا کند رسوا شود

در مسیر زندگی هرگز نمی افتد به چاه با چراغ دین و دانش گر بشر بیناشود

موفق باشید.