Đồ án 2

Cho cấu trúc CSDL như sau:

**TRUONG(MATR, TENTR, DCHITR)**

Mô tả: Mỗi trường học có một Mã số (MATR) duy nhất để nhận biết, có tên trường (TENTR) và địa chỉ (DIACHITR).

**HS(MAHS, HO, TEN, CCCD, NTNS, DCHI\_HS)**

Mô tả: Mỗi học sinh có một mã số duy nhất (MAHS) để nhận biết, họ tên (HO, TEN), số CCCD nếu đến tuổi cấp CCCD, được sinh vào một ngày (NTNS) và cư trú tại một địa chỉ (DCHI). Số CCCD luôn duy nhất nếu có.

**HOC(MATR, MAHS, NAMHOC, DIEMTB, XEPLOAI, KQUA)**

Mô tả: Mỗi học sinh (MAHS) tham gia học tại 1 Trường (MATR) vào một năm học (NAMHOC), sẽ có một trị số Điểm trung bình chung (DIEMTB, tính trên thang 10). Tùy thuộc vào Điểm trung bình chung mà học sinh sẽ được xếp loại (XEPLOAI) học tập là “Xuất sắc”, “Giỏi”, “Khá”, “Trung bình” hoặc “Yếu”; và có một kết quả (KQUA) là “Hoàn thành” hoặc “Chưa hoàn thành”.

**Các yêu cầu:**

1. Hãy viết script **CreateSchema1.sql** tạo cơ sở dữ liệu có tên là TRUONGHOC1, chứa 3 table như mô tả trên; có các khóa chính, khóa ngoại phù hợp. Với mỗi bảng hãy tạo clustered index theo khóa chính.
2. Hãy viết chương trình bằng ngôn ngữ C++, python hoặc Java; tự động phát sinh ra dữ liệu cho các bảng trên: Bảng TRUONG có ít nhất 100 dòng; Bảng HS có ít nhất 1 triệu dòng; Mỗi học sinh trong HS có tương ứng từ 1 đến 3 dòng trong bảng HOC;
3. Hãy viết script **CreateSchema2.sql** tạo cơ sở dữ liệu có tên là TRUONGHOC2 giống như TRUONGHOC1, cả về cấu trúc lẫn dữ liệu; Ngoài ra, script còn tạo thêm trong TRUONGHOC2 các index riêng rẽ trên các thuộc tính TENTR, XEPLOAI, NAMHOC.
4. Hãy viết 1 chức năng nhận đầu vào là {tên database, tên trường; năm học; xếp loại học tập} và in ra màn hình danh sách gồm {họ tên học sinh, NTNS, Điểm TB, xếp loại, kết quả}. Đo lường và in ra thời gian truy vấn dữ liệu. (Hãy tự nhận xét thời gian thực hiện truy vấn khi làm việc trên 2 database khác nhau). Mỗi lần chạy chương trình, cũng xuất dữ liệu ra file XML. (tự động đặt tên file xml theo quy ước *têndatabase-têntrường-nămhọc-xếploại.xml*)
5. Hãy viết chương trình đọc một file xml đang có do câu 4 xuất ra, nhận vào ngưỡng điểm [thấp, cao] và in ra màn hình danh sách học sinh có điểm trung bình nằm trong ngưỡng điểm. Lưu ý sử dụng Xpath.

Quy ước nộp bài: Mỗi SV sẽ nộp 1file

-Một file nén có tên dạng *MSSV\_Họ-Tên-đệm*, chứa bên trong:

|  |  |  |
| --- | --- | --- |
| Tên Folder/File | Loại | Mô tả |
| **CreateSchema1.sql** | File | Script filetạo CSDLTRUONGHOC1 |
| **generate-db.cpp** (nếu viết bằng C++)  **generate-db.py** (nếu viết bằng python)  **generate-db.java** (nếu viết bằng java) | File | Mã nguồn tương ứng câu 2 |
| **CreateSchema2.sql** | File | Script filetạo CSDLTRUONGHOC2 |
| **query-db.cpp** (nếu viết bằng C++)  **query-db.py** (nếu viết bằng python)  **query-db..java** (nếu viết bằng java) | File | Mã nguồn tương ứng câu 4 |
| **XML** | Folder | Chứa các file *têndatabase-têntrường-nămhọc-xếploại.xml* |
| **query-xml.cpp** (nếu viết bằng C++)  **query-xml.py** (nếu viết bằng python)  **query-xml.java** (nếu viết bằng java) |  | Mã nguồn tương ứng câu 5 |
| Youtube.txt | File | File chứa link đến clip thuyết minh và chạy chương trình minh họa. Chiều dài mong đợi 5 phút. |

**Lưu ý:**

* Deadline: 30/6/2023
* Các file mã nguồn cần được ghi chú (comment) rõ ràng, bằng tiếng Việt hoặc tiếng Anh
* File video cần cân chỉnh âm thanh mức trung bình, không quá nhỏ cũng không quá lớn, không cần lồng nhạc và có xuất hiện gương mặt rõ ràng của người thuyết minh.
* Nếu Thầy thu xếp được thời gian thì sẽ mời từng em đến để hỏi thêm trực tiếp.

Chúc các em làm bài tốt.