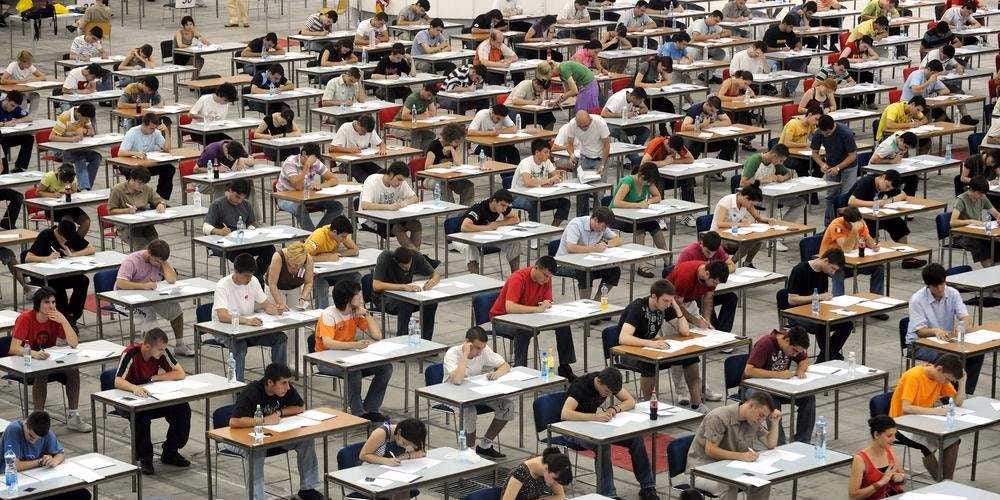


**26th Mar., 2022**

**Midterm Exam**



**CS-GY 6083 - B, Spring-2022. Principles of Database Systems.**

**MIDTERM EXAM [100 points with 20% weight]**

**TIME LIMIT: 2 hours and 30 minutes:**

**03/26/2022 11:00 AM to 1:30PM EST**

**THIS IS AN ONLINE – OPEN BOOK EXAM. PLEASE LOGIN TO ZOOM MEETING USING YOUR NET ID (DO NOT LOGIN WITH YOUR PERSONAL EMAIL ACCOUNT). Find the Zoom meeting details under ZOOM top bar menu of the course website with title “Midterm Exam S2022”**

* **WRITE YOUR ANSWERS UNDER EACH QUESTION (DO NOT REMOVE QUESTION) IN THIS WORD DOCUMENT AND SUBMIT IT ON OR BEFORE 1:30PM TO Course site> ASSIGNMENTS > MIDTERM. Save and submit the exam submission document in format <Your Net id>\_Midterm\_Spring\_2022\_6083B.PDF. YOU MAY RESUBMIT YOUR ASSIGMENT UNLIMITED TIME BEFORE THE SUBMISSION DEADLINE. The latest submission will be considered for the grading. PLEASE MUTE YOUR MICROPHONE DURING ENTIRE EXAM DURATION and KEEP VIDEO ON SINCE IT IS ONLINE EXAM.**
* **This exam has 4 sections A, B, C, and D. All sections and questions have grading points. There is NO negative points for any wrong answers.**
* **IF YOU HAVE ANY QUESTION DURING THE EXAM, PLESAE SEND YOUR QUESTION PRIVATELY TO PROFESSOR/TA ON ZOOM MEETING CHAT WINDOW. DO NOT SPEAK IN MICROPHONE.**
* **USE Oracle Data Modeler for ERD diagrams, no hand drawing will be graded. Insert snapshot of database design models in same PDF document. NO ZIP FILE WILL BE ACCEPTED. NO ANY TYPE OF COPY WILL BE GRADED.**
* **All Objects created in ERD should have your initial as prefix, e.g. AP\_DEPT**

**GOOD LUCK!**

1. **Answer following questions briefly [20 points]**
   1. **Explain difference between Primary Key, Unique Key, and Candidate key**
2. **The primary key is a minimal group of attributes (could have more than one attribute) that uniquely identifies an individual row of each table in a database.**
3. **The unique key can also identifies an individual row in a table. The main difference between primary key and unique key is that primary key cannot be NULL while unique key could be NULL.**
4. **The candidate key is a group of attributes which can also uniquely identifies rows in a table. But since a table can only have “one” primary key, the candidate key could not be considered as the primary key.**
   1. **Explain composite, multivalued, and derived attribute with example of each**
5. **Composite attribute:**

**The composite attribute is an attribute which is made of several components (mostly more than one) and each component can also be a single attribute.**

**For example, the “address” attribute of a bank account can be a composite attribute since it is typically made of separated address information such as street number, city, state and zip code etc. Thus, the “address” attribute could be a composite attribute.**

1. **Multivalued attribute:**

**The multivalued attribute is an attribute that can have more than one**

**Value for each row of a table.**

**For example, suppose there is a “student” table of a university management database and the table has attributes like “Selected Course” and “Phone number”. A student typically registers for several course per semester and they may also have multiple phone numbers. Thus the “Selected Course” and “Phone number” attribute could be a multivalued attribute.**

1. **Derived attribute:**

**The derived attribute is an attribute that can be obtained or inferred from other corresponding attributes.**

**For example, a “student” table has 2 attributes “Date of birth” and “Age”. Obviously, the “Age” attribute could be calculated using the “Date of birth” attribute and the current date. Thus the “Age” attribute is a derived attribute.**

* 1. **Explain importance of good quality data and database design in prospective of business value.**

**There are several advantages of the good quality data and the database design:**

1. **Data redundancy and inconsistency.**

**Good quality of data and design of database could handle the data redundancy and inconsistency problems. Therefore it can help the company to save money on the expense of the data management.**

1. **Data concurrency and security.**

**Typically, for modern business cases, the customer of a service or an software application is growing gigantically. It means that huge amount of requests could be sent at the same time and there may have some criminals trying to hack the application. With a well designed database, it could provide reliable concurrency control services and a secured environment for all users’ data.**

1. **Data searching**

**Nowadays, normal people and technical staff will do a lot of search in their daily life and work. A well-designed database could provide a very efficient search service to them and thus improve their performance in daily life.**

* 1. **What is database constraint? Explain three examples of domain constraints.**

1. **The database constraint is a kind of implementation of the bossiness rule when designing the database. It is used to maintain the data integrity and consistency.**

**(1) Is this attribute mandatory or not?**

**(2) What is the type of this attribute?**

**(3) What is the size of this attribute?**

1. **Draw ERD (relational model) for following problem set [50 points]**

**EDUCATION AND SYSTEM EXCELLENCE (EASE) is a university research department. You have been engaged in summer research program to develop a database system for EASE. EASE has initially identified following entities and attributes for data management.**

**STUDENT: student ID, name, email address, gender, major**

**FACULTY: faculty ID, name, age, status (PART TIME or FULL TIME)**

**COURSE: course name, credit, prerequisite course(s)**

**EASE has also clarified following business rules:**

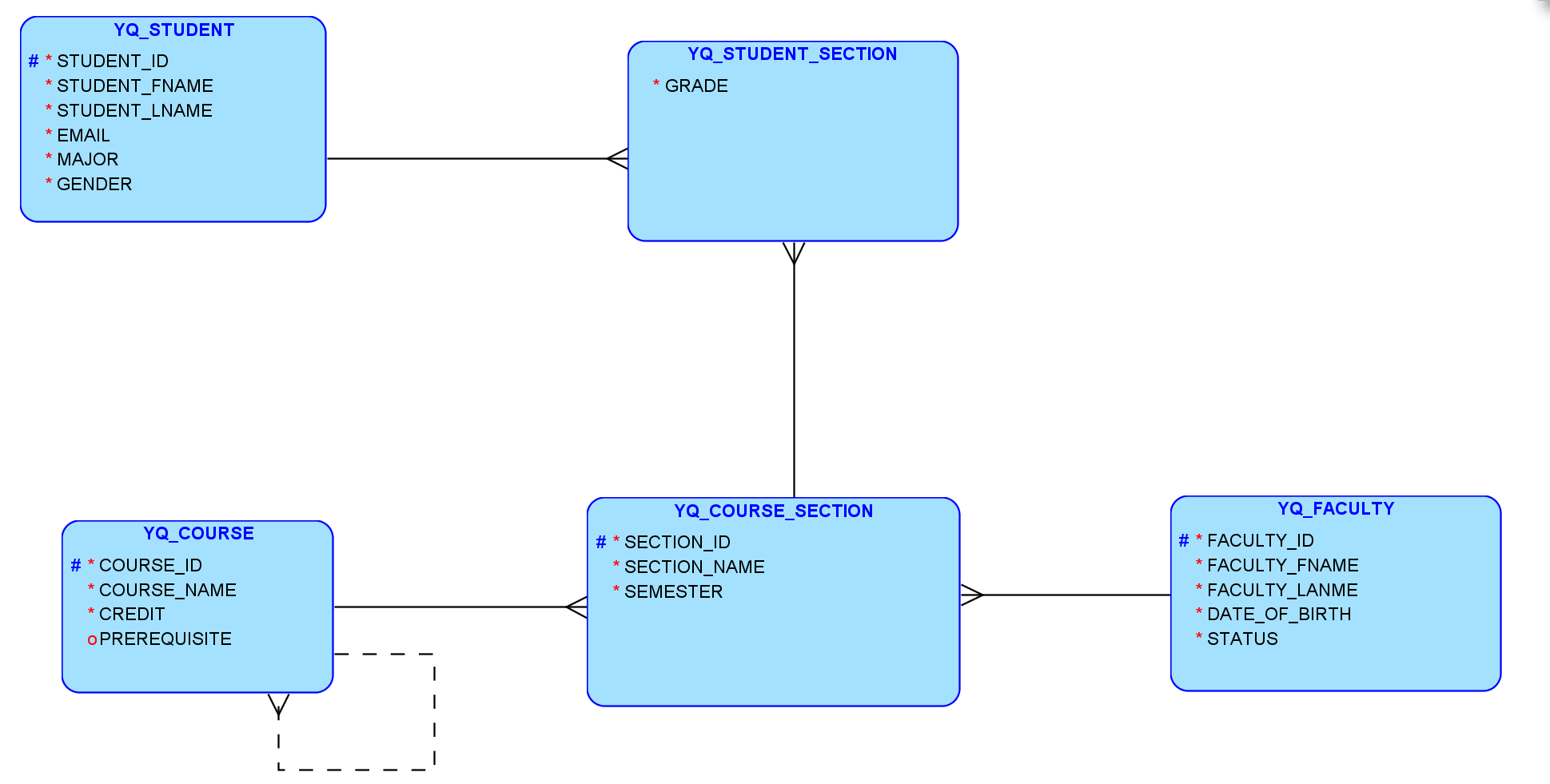
1. **Course can have one or more sections. Each course section is unique by its name, e.g. S2021A (for Spring 2021 A), S2021B (for Spring 2021 B), S2021C (for Spring 2021 C) etc.**
2. **The course can have multiple sections in each semester term**
3. **Each course section is taught by only one faculty**
4. **Same course can be taught by multiple faculties**
5. **Faculties can teach same course for multiple sections**
6. **Students can take multiple sections of different courses**
7. **The course may have one or more prerequisite courses**

**EASE intend to record students letter grade (A+, A, B+, B, C, F) for each section that student takes.**

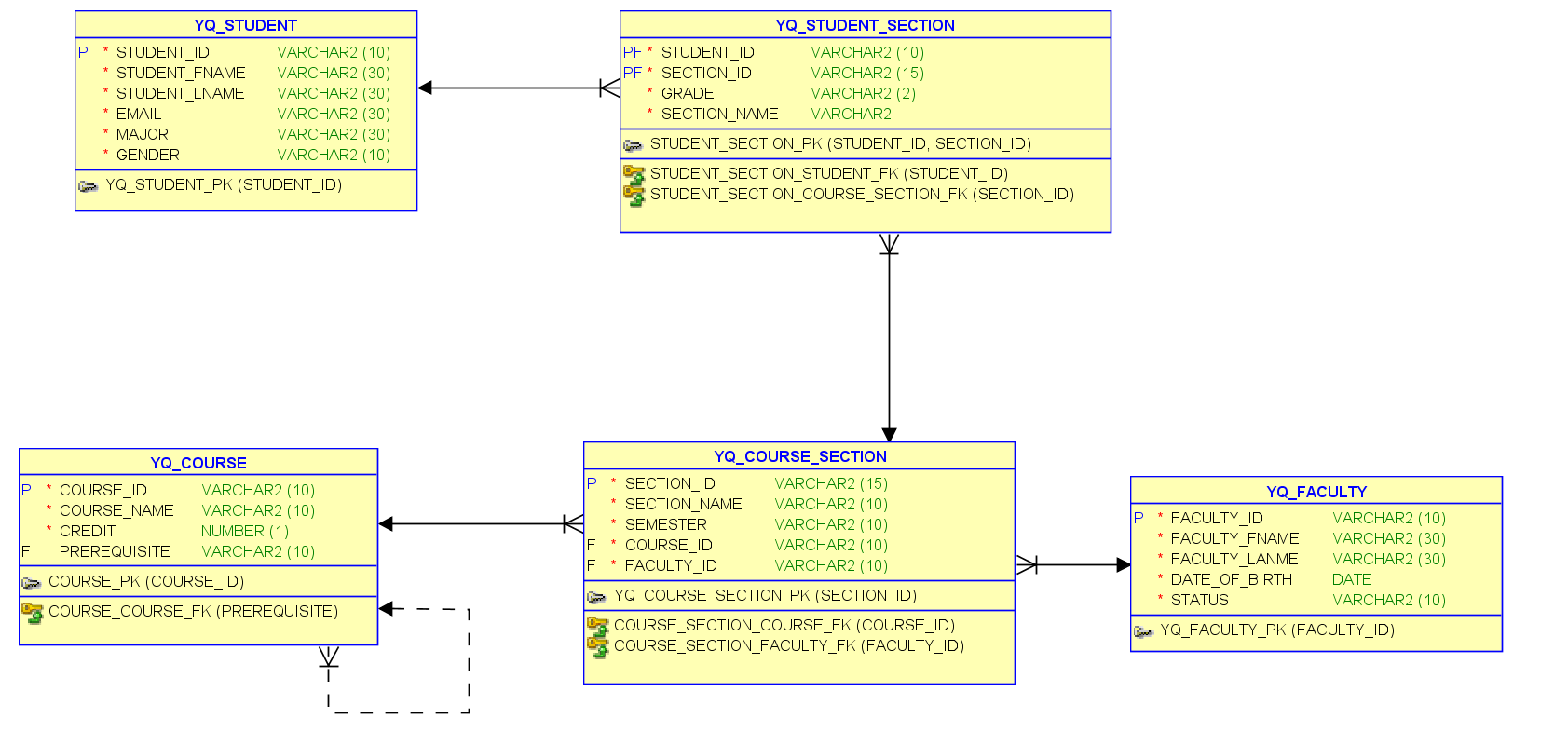
**Using Oracle Data Modeler, draw logical data model satisfying business rules as above. All entities in your submission should have your initial as prefix, such as AP\_COUNTRY. Identify/add appropriate primary key of each table, resolve composite/multivalued/derived attributes, if any. Use proper datatype for each attribute and mark them optional or mandatory, as appropriate. Define appropriate relationship among entities that satisfies the business rules stated as above. Convert Logical model into Relational Model and make necessary adjustments into relational model.**

**Insert screenshots of both logical and relational model. Cleary state any assumptions that you have made other than stated by EASE (if any) to support your solution. No need to submit DDL code.**

**The logical model is listed below:**

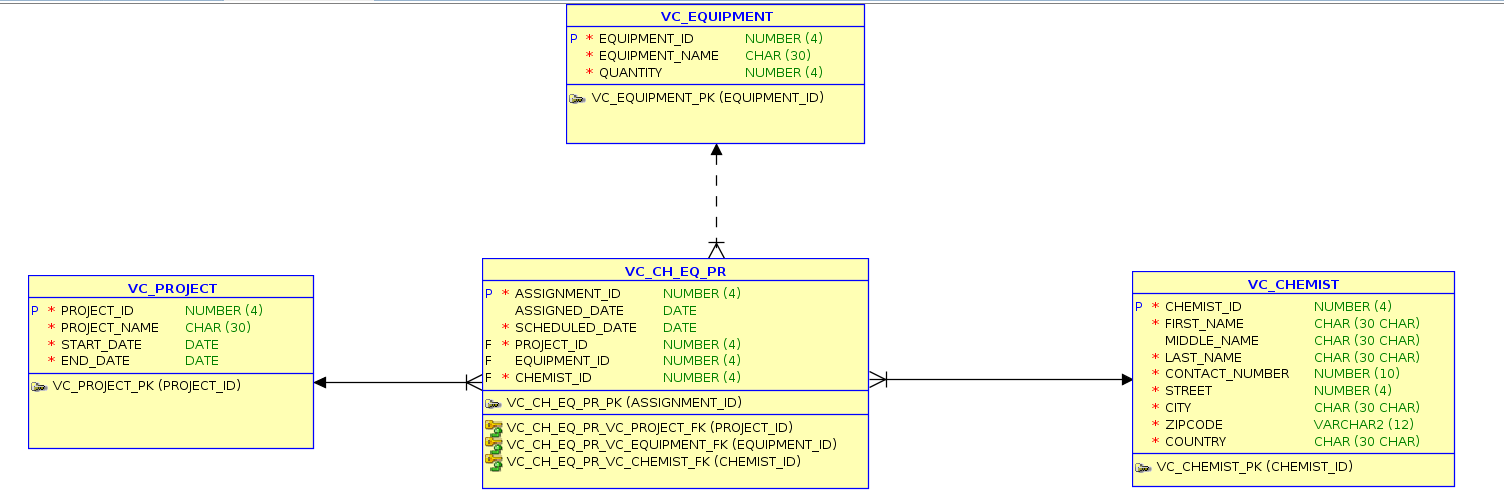


The relational mode is listed below:



Possible Assumptions:

1. The attribute “DATE OF BIRTH” is added to the “FACULTY” table to replace the “Age” attribute because the “Age” attribute could be a derived attribute thus it could be calculated from the “date of birth” and the current date.
2. The unique attribute “SECTION\_ID” is added to the intersection table of “FACULTY” and “COURSE” as an surrogate key.
3. In the “YQ\_STUDENT\_SECTION” table, the attribute “STUDENT\_ID” and the attribute “SECTION\_ID” makes up the compositeA primary key of this table.
4. **For given relational model below, write SQL for requirements asked [15 points]**



* + 1. **Enforce the business rule that EQUIPMENT\_ID attribute in VC\_EQUIPMENT table has minimum 4 digits**

**ALTER TABLE VC\_EQUIPMENT ADD CONSTRAINT CHECK\_VC\_EQUIP\_ID (EQUIPMENT\_ID >= 1000);**

* + 1. **Drop column ASSIGNMENT\_ID in intersect table and create a primary key for the intersect table, VC\_CH\_EQ\_PR**

**ALTER TABLE VC\_CH\_EQ\_PR DROP ASSIGNMENT\_ID;**

**ALTER TABLE VC\_CH\_EQ\_PR ADD CONSTRAINT VC\_CH\_EQ\_PR\_PK PRIMARY KEY(PROJECT\_ID, CHEMIST\_ID);**

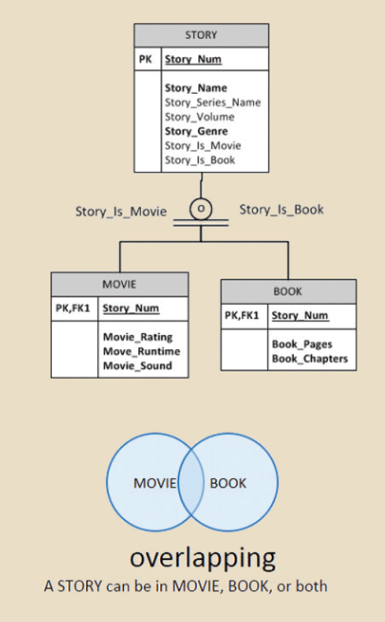
* + 1. **Add an attribute PROJECT\_DESCRIPTION in VC\_PROJECT table with maximum size of 100 characters**

**ALTER TABLE VC\_PROJECT ADD PROJECT\_DESCRIPTION varchar2(100);**

1. **Consider following EERD (Extended ERD) [15 points]**

**It is the example of Supertype and Subtypes with total specialization rule and overlapping constraint.**

**Using Oracle Data Modeler, create Logical and Relational model with proper datatype and size of each attributes. The bold face attributes in picture are mandatory and others are optional. Instead of Story\_Is\_Movie and Story\_is\_Book attributes, create a discriminator attribute Story\_Type with possible values as “M” or “B”. Supertype and Subtypes should have your initial, e.g. AP\_STORY, AP\_MOVIE, AP\_BOOK.**



Examples of **Story\_Genre** are

Action

Comedy

Drama

Fantasy

Horror

Mystery

Romance

Thriller

Examples of **Story\_Series\_Name** are

Star Wars

Harry Potter

James Bond Franchise

WARNER BROS

Disney

**Story\_Volume** is typically identified sequentially e.g. "volume 3" or "volume III" of book publications.

Example of **Movie\_Rating** are G, PG, PG13, R etc.

**Movie\_Runtime** is length of movie in minutes.

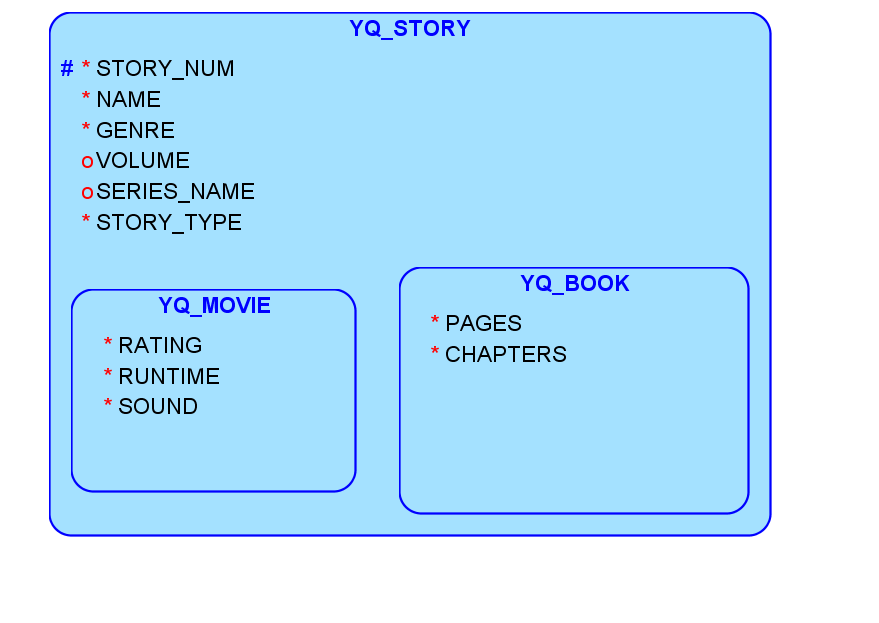
Examples of **Movie\_Sound** are Dolby Stereo, Dolby Surround, DTS, and Sonic etc.

**Book\_Pages** is total number of pages of book

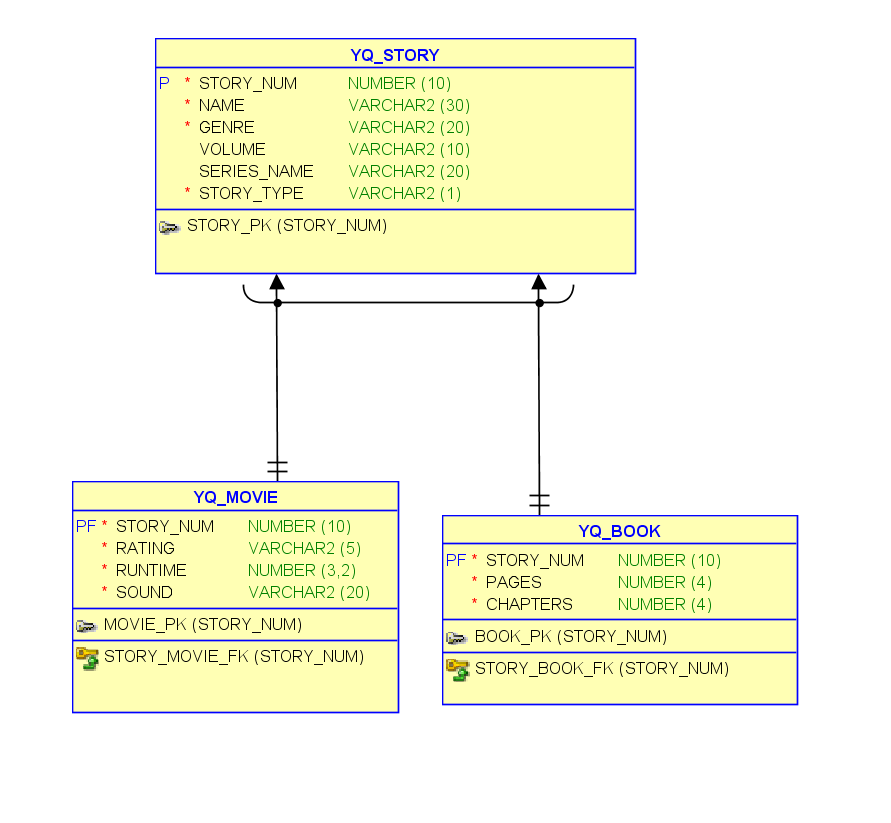
**Book\_Chapters** is total number of chapters in book.

**Insert Logical Model, Relational model, and any assumptions made to support your solution. No need to submit DDL code.**

1. **The logical model is listed below:**



1. **The relational model is listed below:**



**====================END OF THE DOCUMENT =========================**