

6th Nov., 2021

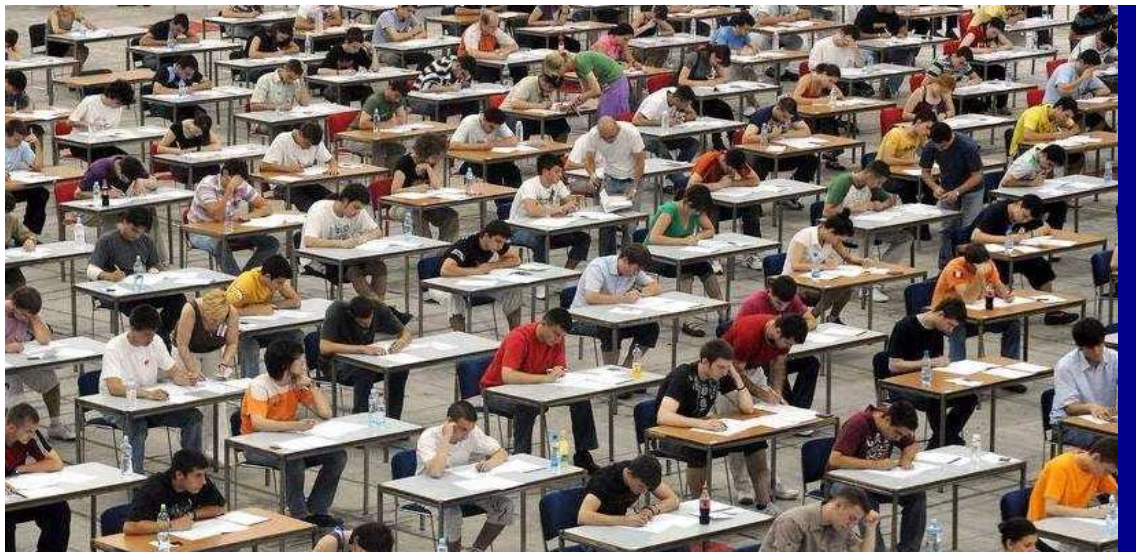
Midterm Exam



NYU

TANDON SCHOOL
OF ENGINEERING

CS-GY 6083 - B, Fall-2021.
Principles of Database Systems.



MIDTERM EXAM [100 points with 20% weight]

TIME LIMIT: 2 hours and 30 minutes:

11/06/2021 10:00 AM to 12:30 PM EST

Please read instructions carefully before writing exam

- Write your name, student id, and net id below
- Last Name: Li First Name: Minrui
- Net ID: ml7136 Student ID: N17048569

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 F2021”

- WRITE YOUR ANSWERS UNDER EACH QUESTION IN THIS WORD DOCUMENT AND SUBMIT IT ON OR BEFORE 12:30PM TO Course site > ASSIGNMENTS > MIDTERM. Save and submit the exam submission document in format <Your Net id>_Midterm_Fall_2021_6083B. **YOU MAY RESUBMIT YOUR ASSIGNMENT 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 5 sections A, B.C.D, and E. 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 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 Word/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!

A) Answer following questions briefly [18 points]

- i) **Explain the difference between COMPOSITE attributes and MULTI-VALUED attributes with example? How to resolve COMPOSITE and MULTI-VALUED attributes for the Database Design?**

A COMPOSITE attribute is an attribute which consists of several SIMPLE attributes. For example, NAME is a composite attribute, because it can be broken into 3 SIMPLE attributes: FirstName, MiddleName and LastName.

A MULTI-VALUED attributes is an attribute which could take more than 1 value. For example, SKILL is a MULTI-VALUED attribute, because one person may have more than 1 Skill.

COMPOSITE attribute could be resolved by splitting it into several SIMPLE attributes.

MULTI-VALUED attribute should be represented as a new table.

- ii) **Explain Primary Key, Unique Key, Candidate Key, Foreign Key, and Surrogate key with examples**

Primary Key: primary key is the unique attribute in a table, it can't be NULL and cannot be repeated. For example, NYU StudentID.

Unique Key: unique key is the unique attribute, but it can be NULL. For Example, someone's phone number, it's unique but someone's may not have phone number.

Candidate Key: candidate key is also the unique attribute, it can be primary key as well. But since we already have a primary key, we don't need these keys to be primary key. For example, NYU students' email, it's unique and it can't be NULL, but since we use the StudentID as primary, students' email will become candidate key.

Foreign Key: foreign key is the attribute is one table that refers to the primary key of another table. For example, SkillID is a foreign key, it refers to the primary key in table SKILL.

Surrogate Key: surrogate key is the attribute that can be declared as primary key instead of natural key. For example, when can use ShipmentID as the primary of table SHIPMENT, instead of using the combination of CustomerID and VendorID as the composite primary key.

iii) **Explain the terms Data, Metadata, Information, and Discriminator attribute with example.**

Data: data is the items stored in database, it could be numeric, char, varchar, date, etc.

Metadata: metadata is data of data, such as data type and data size.

Information: information is something useful derived from data, such as the average salary, average grade.

Discriminator Attribute: discriminator attribute is an attribute in supertype which determines what subtype it should be. For example, the EMP_TYPE is the discriminator attribute of table EMP. When EMP_TYPE is H, it's a hourly employee; when EMP_TYPE is S it's a salaried manager; when EMP_TYPE is C, it's CONSULTANT.

B) For given data sets, answer questions [12 points]

Table name: CUSTOMER

Primary key: CUS_CODE

Foreign key: AGENT_CODE

| CUS_CODE | CUS_LNAME | CUS_FNAME | CUS_INITIAL | CUS_RENEW_DATE | AGENT_CODE |
|----------|-----------|-----------|-------------|----------------|------------|
| 10010 | Ramas | Alfred | A | 05-Apr-2018 | 502 |
| 10011 | Dunne | Leona | K | 16-Jun-2018 | 501 |
| 10012 | Smith | Kathy | W | 29-Jan-2019 | 502 |
| 10013 | Olowski | Paul | F | 14-Oct-2018 | |
| 10014 | Orlando | Myron | | 28-Dec-2018 | 501 |
| 10015 | O'Brian | Amy | B | 22-Sep-2018 | 503 |
| 10016 | Brown | James | G | 25-Mar-2019 | 502 |
| 10017 | Williams | George | | 17-Jul-2018 | 503 |
| 10018 | Farriss | Anne | G | 03-Dec-2018 | 501 |
| 10019 | Smith | Olette | K | 14-Mar-2019 | 503 |

Table name: AGENT (only five selected fields are shown)

Primary key: AGENT_CODE

Foreign key: none

| AGENT_CODE | AGENT_AREACODE | AGENT_PHONE | AGENT_LNAME | AGENT_YTD_SLS |
|------------|----------------|-------------|-------------|---------------|
| 501 | 713 | 228-1249 | Alby | 132735.75 |
| 502 | 615 | 882-1244 | Hahn | 138967.35 |
| 503 | 615 | 123-5589 | Okon | 127093.45 |

- I. **Which table is Parent table and which table is Child table? What is degree and cardinality of the relations between these tables? Is the relationship mandatory or optional?**

AGENT is parent table, CUSTOMER is child table.

Degree is 2.

Cardinality is one-to-many.
The relationship is optional.

II. Please explain possibilities of various effects when record of AGENT_CODE 503 is deleted from AGENT table. Explain all possible situation with Delete Rules.

If the rule is restrict, then we can't delete the record of AGENT_CODE 503, because there exists related rows in CUSTOMER.

If the rule is cascade, then when we delete the AGENT_CODE 503 in AGENT table, all the related rows in CUSTOMER table will be deleted as well.

If the rule is set-to-NULL, then after we delete the AGENT_CODE 503, all the related rows in CUSTOMER will be set to NULL.

C) For given dataset below, draw ERD (Logical and Relational Mode) [15 points]

Table name: EMPLOYEE

| EMP_CODE | EMP_LNAME | JOB_CODE |
|----------|-----------|----------|
| 14 | Rudell | 2 |
| 15 | McDade | 1 |
| 16 | Ruellardo | 1 |
| 17 | Smith | 3 |
| 20 | Smith | 2 |

Table name: BENEFIT

| EMP_CODE | PLAN_CODE |
|----------|-----------|
| 15 | 2 |
| 15 | 3 |
| 16 | 1 |
| 17 | 1 |
| 17 | 3 |
| 17 | 4 |
| 20 | 3 |

Table name: JOB

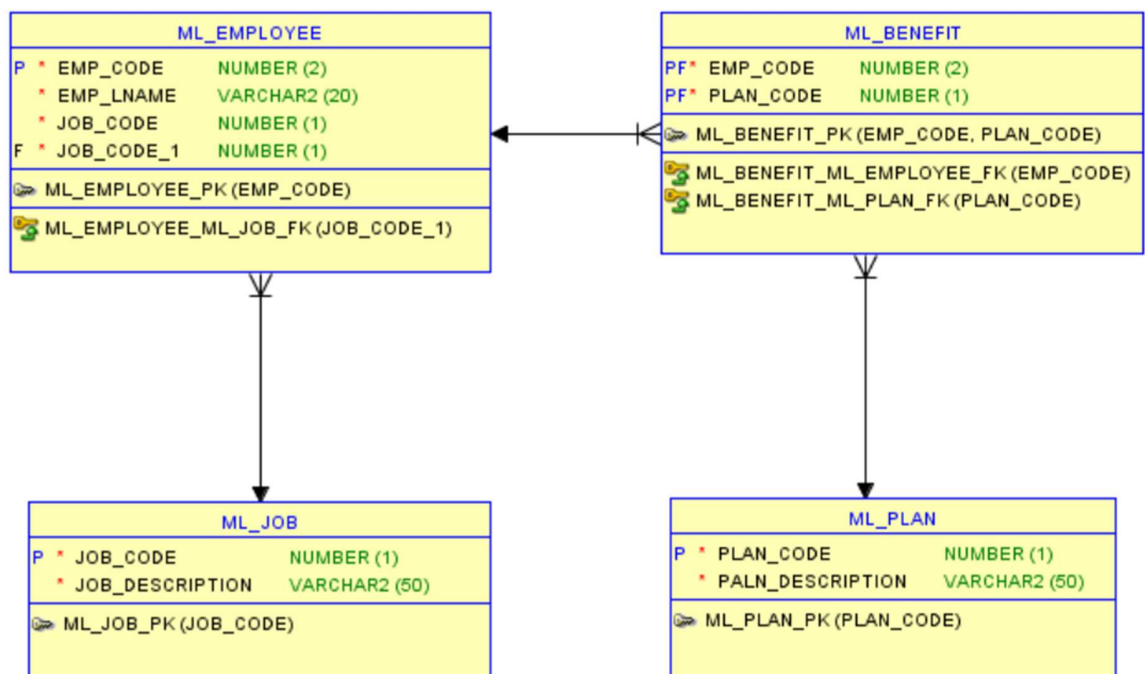
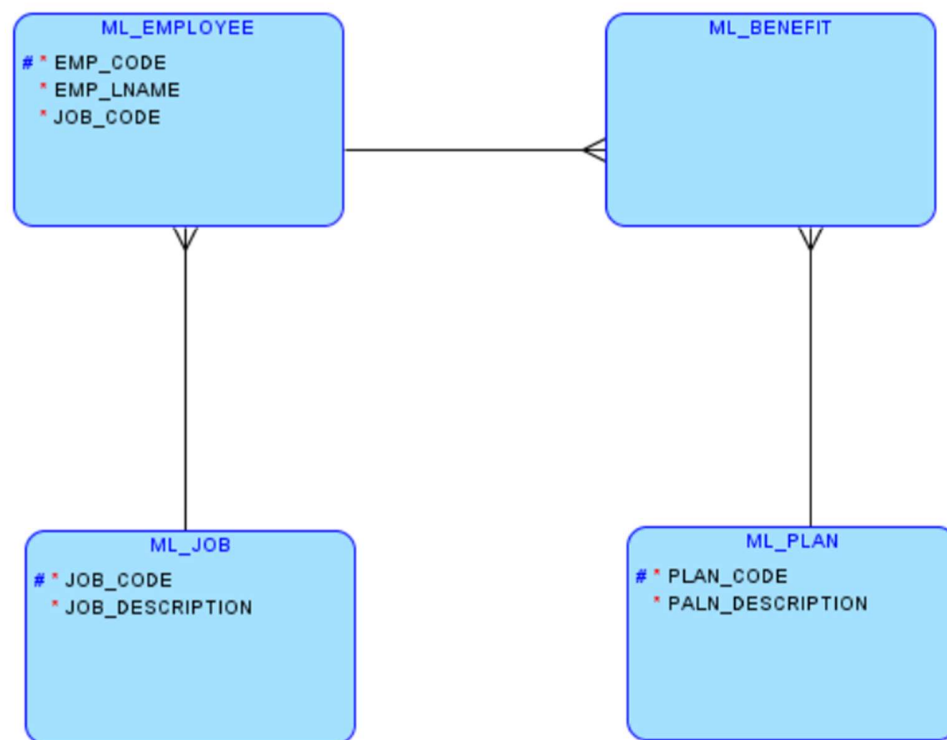
| JOB_CODE | JOB_DESCRIPTION |
|----------|-----------------|
| 1 | Clerical |
| 2 | Technical |
| 3 | Managerial |

Table name: PLAN

| PLAN_CODE | PLAN_DESCRIPTION |
|-----------|----------------------|
| 1 | Term life |
| 2 | Stock purchase |
| 3 | Long-term disability |
| 4 | Dental |

Using Oracle Data Modeler draw ERD (both Logical and Relational model) of your database solution.

If you have made any assumptions about business rules, please clearly state your assumptions to support your ERD models.



D) Data Normalization ERD [35 points]

WOW (WE OFFER WELLNESS) is a healthcare group with hospitals in affiliation. Each hospital has many doctors providing medical services. Each doctor may treat many patients for many diseases, each patient may be treated by many doctors for many diseases, and each disease can be treated by many doctors for many patients. At present, WOW stores data of each treatment in a spreadsheet. As data is growing, it is becoming very difficult to maintain data and to protect data integrity. WOW has decided to develop a centralized relational database system. Being an intern at WOW, you have been asked to design a normalized data model. WOW has provided the following sample data.

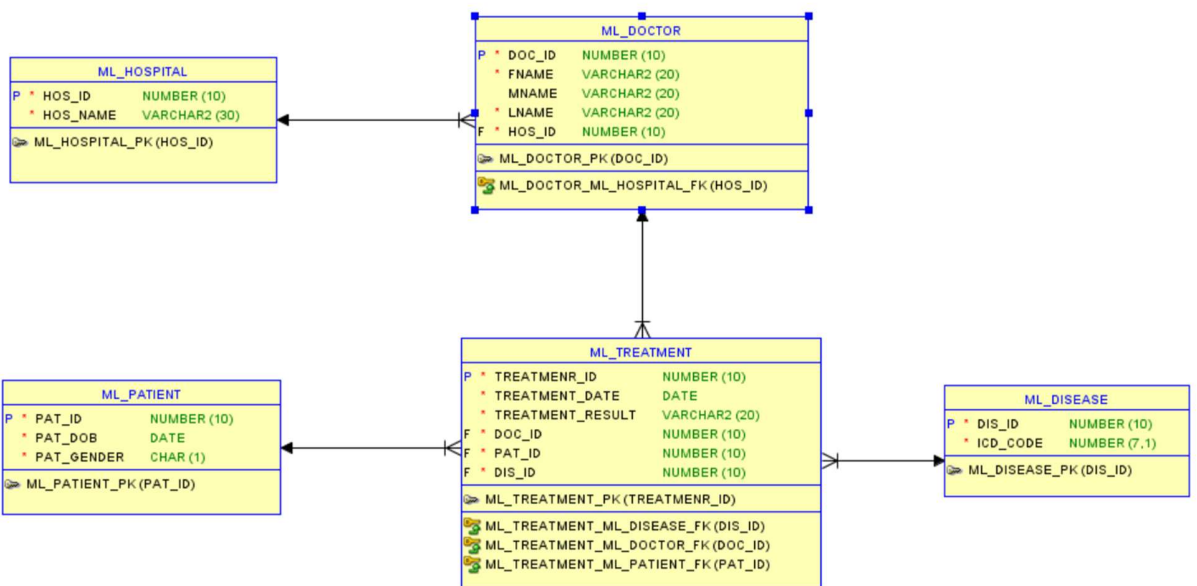
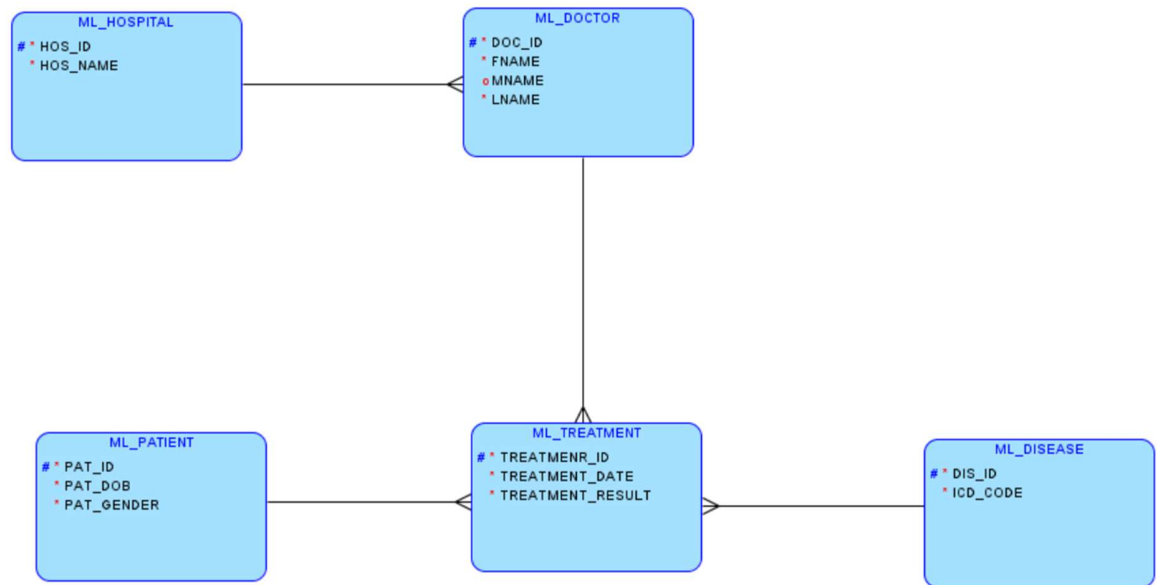
| HOSPITAL_ID | HOSPITAL_NAME | DOCTOR_ID | FNAME | LNAME | PATIENT_ID | PDOB | PGENDER | DISEASE_ID | ICD_CODE | TREATMENT_DATE |
|-------------|---------------|-----------|-------|---------|------------|-----------|---------|------------|----------|----------------|
| 800 | Langone | 100 | James | Johnson | 5000 | 3-May-80 | M | 1000 | 2 | 10-Feb-21 |
| 800 | Langone | 100 | James | Johnson | 5000 | 3-May-80 | M | 1003 | 831.1 | 19-Jul-20 |
| 800 | Langone | 100 | James | Johnson | 5002 | 14-Dec-63 | M | 1003 | 831.1 | 3-Dec-20 |
| 800 | Langone | 100 | James | Johnson | 5003 | 10-Jul-05 | F | 1002 | 480.9 | 8-Jan-21 |
| 800 | Langone | 100 | James | Johnson | 5007 | 11-Nov-99 | F | 1000 | 2 | 27-Dec-20 |
| 800 | Langone | 100 | James | Johnson | 5007 | 11-Nov-99 | F | 1002 | 480.9 | 13-Apr-20 |
| 801 | Regent | 101 | Lily | White | 5005 | 22-Apr-97 | M | 1002 | 480.9 | 1-Nov-20 |
| 801 | Regent | 101 | Lily | White | 5007 | 11-Nov-99 | F | 1002 | 480.9 | 3-Feb-18 |
| 801 | Regent | 101 | Lily | White | 5008 | 9-Mar-87 | M | 1001 | 295.1 | 26-Sep-20 |
| 801 | Regent | 101 | Lily | White | 5008 | 9-Mar-87 | M | 1003 | 831.1 | 7-Feb-21 |
| 801 | Regent | 103 | Nancy | Miller | 5000 | 3-May-80 | M | 1003 | 831.1 | 18-Oct-20 |
| 801 | Regent | 103 | Nancy | Miller | 5001 | 20-Jan-99 | F | 1002 | 480.9 | 17-Nov-20 |
| 801 | Regent | 103 | Nancy | Miller | 5006 | 30-Aug-10 | M | 1001 | 295.1 | 15-Jan-21 |
| 801 | Regent | 103 | Nancy | Miller | 5007 | 11-Nov-99 | F | 1002 | 480.9 | 10-Sep-19 |
| 802 | Riverside | 102 | David | Smith | 5001 | 20-Jan-99 | F | 1001 | 295.1 | 5-Sep-20 |
| 802 | Riverside | 102 | David | Smith | 5002 | 14-Dec-63 | M | 1001 | 295.1 | 16-Dec-20 |
| 802 | Riverside | 102 | David | Smith | 5004 | 10-Oct-00 | F | 1002 | 480.9 | 29-Jan-21 |
| 802 | Riverside | 102 | David | Smith | 5004 | 10-Oct-00 | F | 1003 | 831.1 | 22-Jun-20 |
| 802 | Riverside | 102 | David | Smith | 5008 | 9-Mar-87 | M | 1002 | 480.9 | 29-Oct-19 |

FNAME and **LNAME** represent a doctor's first name and last name, **PDOB** represents a patient's date of birth, and **PGENDER** represents a patient's gender. **ICD_CODE** represents international classification of diseases.

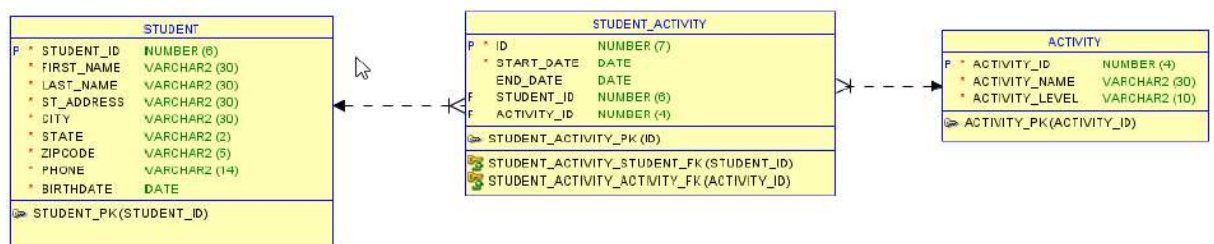
Using data normalization process, design database model up to 3rd Normal Form. Identify the primary key of each entity you define. Use appropriate data types and size of each attribute. Define proper relationships among entities.

Using Oracle Data Modeler draw ERD (both Logical and Relational model) of your database solution.

If you have made any assumptions about the business rules other than listed by WOW, please clearly state your assumptions to support your ERD models.



E) Consider following ERD and write SQL to answer the questions [20 points]



- I. **Drop the primary key of the STUDENT_ACTIVITY table and then drop the column ID.**

```
ALTER TABLE STUDENT_ACTIVITY DROP PRIMARY KEY;  
ALTER TABLE STUDENT_ACTIVITY DROP COLUMN ID;
```

- II. **Add a composite primary key for the STUDENT_ACTIVITY table for STUDENT_ID and ACTIVITY_ID.**

```
ALTER TABLE STUDENT_ACTIVITY ADD CONSTRAINT  
STUDENT_ACTIVITY_PK PRIMARY KEY (STUDENT_ID, ACTIVITY_ID);
```

- III. **Apply the business rule that ACTIVITY_LEVEL in ACTIVITY table has one of three possible values: "EASY", "MODERATE", or "CHALLENGING"**

```
ALTER TABLE ACTIVITY ADD CONSTRAINT C_ACTIVITY_LEVEL CHECK  
(ACTIVITY_LEVEL IN ('EASE', 'MODERATE', 'CHALLENGING'));
```

- IV. **Apply the business rule that STUDENT_ID in STUDENT table has a minimum of 6 digits.**

```
ALTER TABLE STUDENT ADD CONSTRAINT C_STUDENT_ID CHECK  
(STUDENT_ID BETWEEN 100000 AND 999999);
```

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