Review questions

1. Describe the three-schema architecture.

+ Theo kiến trúc ANSI-PARC, một CSDL có 3 mức biểu diễn

+ Kiến trúc 3 mức giúp tách biệt các ứng dụng người dùng với CSDL vật lý.

* + Mức vật lý: Sử dụng mô hình dữ liệu vật lý để mô tả cấu trúc lưu trữ vật lý của CSDL
  + Mức logic/quan niệm: Dấu đi chi tiết về cấu trúc lưu trữ vật lý. Dùng mô hình dữ liệu logic để mô tả cái gì được lưu trữ trong CSDL và mối quan hệ giữa các dữ liệu đó.
  + Mức ngoài/ view: Mô tả một phần của CSDL cho một nhóm người dùng quan tâm và dấu đi phần còn lại của CSDL khỏi nhóm người dùng đó.

1. Define the following terms and give an example for each term: an attribute, the domain of an attribute, a relation schema, a relation, n-tuple, degree of a relation, a relationship, a relation instance (state), a relational database schema, a relational database state, integrity constraints.

* An attribute is a column of the table
* Domain of an attribute is the range of value that a cell in this column can be set
* A relation schema is all of attributes that exist in a relation
* A relation is a table that includes rows and cols, the data in a relation have to gurantee constrainsts related that relation.
* n-tuple: is rows
* Degree of a relation is the numbers of attributes of its relation schema
* A relationship: biểu diễn mối kết hợp giữa các quan hệ trong CSDL (mối quan hệ dựa trên khóa ngoại).
* A relation instance: Tập tất cả các bộ của một quan hệ tại một thời điểm.
* a relational database schema: một tập các lược đồ quan hệ có liên quan với nhau về một lĩnh vực cụ thể.
* a relational database state is a set of relation states such that every relational database state satisfies the integrity constraints specified on relational database schema.
* integrity constraints: Tập các quy tắc mà mọi dữ liệu trong CSDL phải tuân theo nhằm đảm bảo tính toàn vẹn của cơ sở dữ liệu.

1. Why are tuples in a relation not ordered?

* In a relational database, tuples in a relation are not ordered because the relational model is based on the set theory concept. A set is an unordered collection of distinct elements, and similarly, a relation in a database is considered to be a set of tuples

1. Why are duplicate tuples not allowed in a relation?

* Because each relation has its own primary key that have constraints about unique and not null, if duplicate tuples exist, it will not ensure integrity constraints.

1. What is the difference between a key and a superkey?

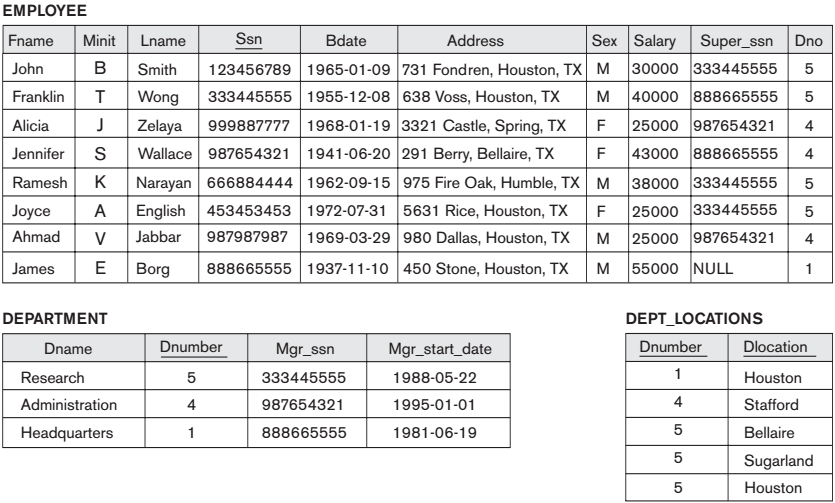
* SQL keys are used to identify rows in a table uniquely. SQL keys can either be a single column or a group of columns. A Super key is a single key or a group of multiple keys that can uniquely identify tuples in a table

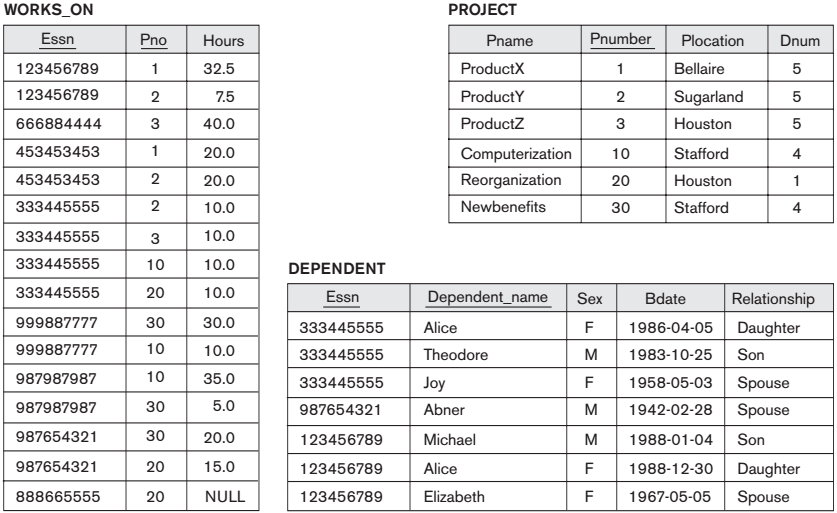
1. Discuss the entity integrity and referential integrity constraints. Why is each considered important?

* An entity is any person, place, or thing to be recorded in a database. Each table represents an entity, and each row of a table represents an instance of that entity. To identify each row in a table, the table must have a primary key. The primary key is a unique value that identifies each row. This requirement is called the entity integrity constraint.
* A referential integrity constraint is also known as foreign key constraint. A foreign key is a key whose values are **derived from the Primary key of another table**.

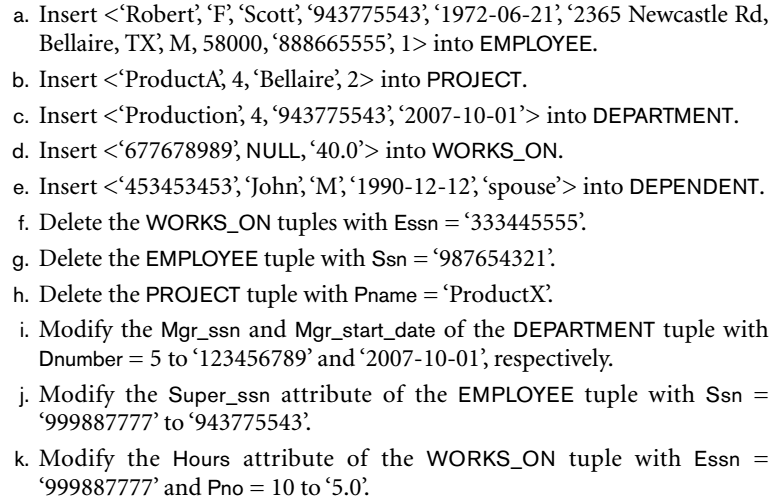
Exercises.

1. Given one possible database state for the COMPANY relational database schema as follows:





- Suppose that each of the following Update operations is applied directly to the database state shown above. Discuss all integrity constraints violated by each operation, if any, and the different ways of enforcing these constraints.



1. Insert <'Robert', 'F', 'Scott', '943775543', '1972-06-21', '2365 Newcastle Rd, Bellaire, TX’, M, 58000, '888665555’, 1> into EMPLOYEE.

**Đúng**

1. Insert <'ProductA', 4, 'Bellaire', 2> into PROJECT.

+ Sai, muốn thực hiện được dòng lệnh này ta phải chèn thêm dòng dữ liệu cócột Dnumber là 2 cho bảng DEPARTMENT

1. Insert <'Production', 4, '943775543', '2007-10-01'> into DEPARTMENT.

+ Sai do Dnumber = 4 đã xuất hiện 1 lần rồi, muốn thực hiện được dòng lệnh này ta phải thay đổi giá trị cột Dnumberkhác 4 để không bị trùn

1. Insert <'677678989', NULL, '40.0'> into WORKS\_ON.

+ Sai do Pno la một thuộc tính trong khóa chính trong bảng WORKS\_ON nên có ràng buộc not null và unique, khong thể nhận giá trị null

1. Insert <'453453453', 'John', 'M', '1990-12-12', 'spouse'> into DEPENDENT.

+ Đúng

1. Delete the WORKS\_ON tuples with Essn = '333445555'.

+ Đúng

1. Delete the EMPLOYEE tuple with Ssn = '987654321'

+ Sai, muốn thực hiện được câu lệnh trên ta phải thay đổi giá trị Mgr\_ssn, vàMgr\_start\_date ở bảng DEPARTMENT, xóa bộ có giá trị ở cột Essn "987654321”ở bảng DEPENDENT, xóa các bộ có giá trị ở cột Essn “987654321” ở bảngWORKS\_ON. Sau đó thực hiện lệnh

1. Delete the PROJECT tuple with Pname = 'ProductX'.

+ Sai, cột Pnumber ở bảng PROJECT là khóa ngoại tham chiếu với cột Pno bảng WORKS\_ON. Theo câu lệnh trên, xóa bộ có giá trị cột Pname là “ProductX” sẽ làm mất giá trị “1” ở cột Pnumber vì thế các bộ có giá trị cột Pno ở bảng WORKS\_ON sẽ không có giá trị để tham chiếu vào vì thế vi phạm ràng buộc

1. Modify the Mgr\_ssn and Mgr\_start\_date of the DEPARTMENT tuple with Dnumber = 5 to '123456789' and '2007-10-01', respectively.

+ Đúng

1. Modify the Super\_ssn attribute of the EMPLOYEE tuple with Ssn = '999887777' to '943775543'

+ Sai do giá trị ssn = 943775543 không tồn tại

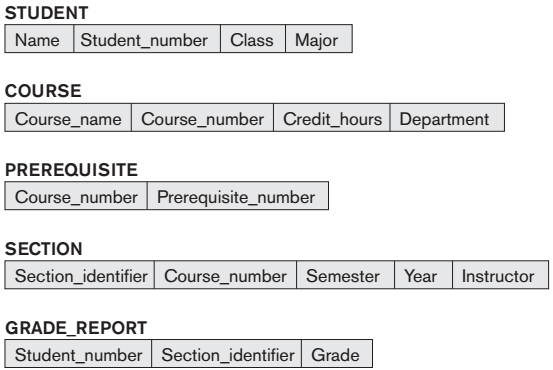
1. Modify the Hours attribute of the WORKS\_ON tuple with Essn '999887777' and Pno = 10 to 5.0'.

+ Đúng

( pi chiếu, sigma chọn)

* Which of the following operations are not correct? Why? Show the results of the correct operations:

1. ΠSsn, Fname, Lname, Salary(Employee)
2. σSex = F and Salary > 30000(Employee)
3. σSex = F and Salary > 30000 (ΠSsn, Fname, Lname, Salary(Employee))
4. ΠSsn, Fname, Lname, Salary (σSex = F and Salary > 30000(Employee))
5. Department X Dept\_Location
6. Department |X| Dept\_Location
7. Compare the results of question e and f
8. Employee – Dependent
9. ΠSsn(Employee) - ΠEssn(Dependent)
10. σSex = F and Salary > 30000(Employee) ∪ σSex = M and Salary < 30000(Employee)
11. ΠSsn(σDno=5(Employee)) ∩ ρSsn(ΠEssn(σPno =1(Works\_on)))
12. Explain the sense of the result of question j and k.
13. Given a University database as follows:



Consider the above figure. In addition to constraints relating the values of columns in one table to columns in another table, there are also constraints that impose restrictions on values in a column or a combination of columns within a table. One such constraint dictates that a column or a group of columns must be unique across all rows in the table. For example, in the STUDENT table, the Student\_number column must be unique (to prevent two different students from having the same Student\_number).

- Identify the column or the group of columns in the other tables that must be unique across all rows in the table.

- Identify the primary key, foreign key in each relational schema.

- Specify the following queries on the University database schema:

a. Retrieve the student\_number and the name of all students in the university.

Π Student\_number, Name(STUDENT)

b. Retrieve the student\_number and the name of all students in “Computer Science” major

c. Retrieve the prerequisite course number and course name of the course which has course number “1111’’

d. Retrieve the section identifier, course number and course name of courses which are opened in semester 1, year 2019

e. Retrieve the section identifier, course number and course name of courses which are not opened in semester 1, year 2019

f. Retrieve the section identifier, course number and course name of courses which are opened in year 2018 or 2019.

g. Retrieve the section identifier, course number and course name of courses which are opened in year 2018 and 2019.