BBM471 DATABASE MANAGEMENT SYSTEMS

Homework 3

1. Consider the given schema of a sample database.

```
Worker(WORKER_ID, FIRST_NAME, LAST_NAME, SALARY, JOINING_DATE, DEPARTMENT)
```

a) Write an SQL query to fetch the count of employees working in the department 'Admin'.

```
SELECT COUNT(*) FROM worker WHERE DEPARTMENT = 'Admin';
```

b) Write an SQL query to fetch the list of employees with the same salary.

```
Select distinct W.WORKER_ID, W.FIRST_NAME, W.Salary
from Worker W, Worker W1
where W.Salary = W1.Salary
and W.WORKER_ID != W1.WORKER_ID;
```

c) Write an SQL query to show the second highest salary from a table.

```
Select max(Salary) from Worker
where Salary not in (Select max(Salary) from Worker);
```

d) Write an SQL query to fetch the departments that have less than five people in it.

```
SELECT DEPARTMENT, COUNT(WORKER_ID) as 'Number of Workers' FROM Worker GROUP
BY DEPARTMENT HAVING COUNT(WORKER_ID) < 5
```

e) Write an SQL query to fetch the names of workers who earn the highest salary.

```
SELECT FIRST_NAME, SALARY from Worker WHERE SALARY=(SELECT max(SALARY) from
Worker);
```

2. Consider the following relations containing airline flight information:

Flights(flno: integer, from: string, to: string, distance: integer, departs: time, arrives: time)
Aircraft(aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)

Note that the Employees relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft (otherwise, he or she would not qualify as a pilot), and only pilots are certified to fly. Write the following queries in relational algebra

a) Find the eids of pilots certified for some Boeing aircraft.,

$$\pi_{eid}(\sigma_{aname='Boeing'}(Aircraft \bowtie Certified))$$

b) Find the eids of employees who make the highest salary.

$$\rho(E1, Employees)$$
$$\rho(E2, Employees)$$
$$\rho(E3, \pi_{E2.eid}(E1 \bowtie_{E1.salary>E2.salary} E2)$$
$$(\pi_{eid}E1) - E3$$

Subject : Designing an ER Diagram

Advisor : Res. Assist. (Nebi Yılmaz, Selim Yılmaz)

Entrant's ID: ..... Name & Surname: .....

Quiz 1

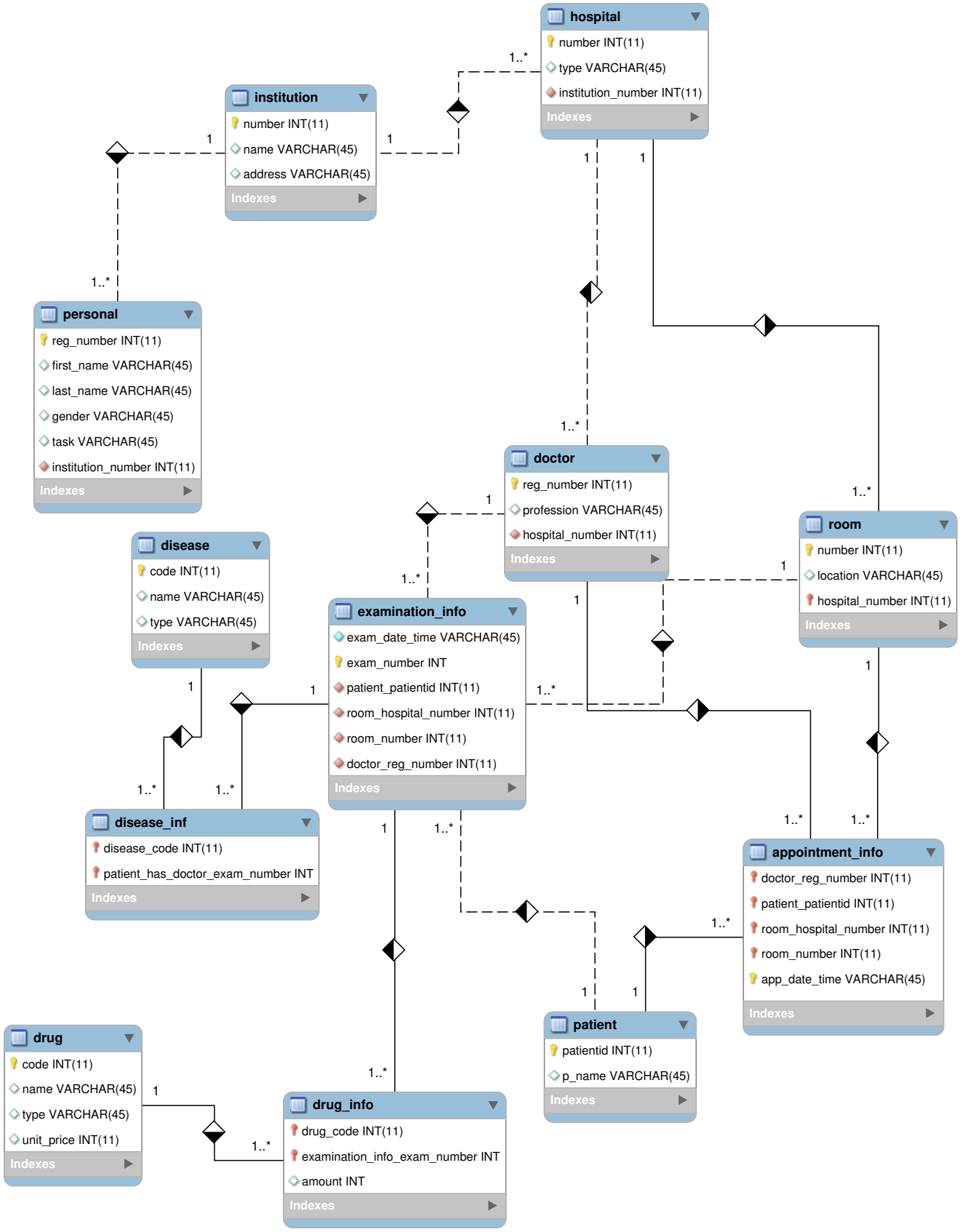
A database named "**Public health employee**", should have the entity sets and relations explained below:

- Number, name, and address of **each public institution**.
- Number and type of **each public hospital** (each public hospital is regarded as a public institution)
- Number (unique in hospital) and location of each **patient exam room** of public hospital.
- Registration number, first name, last name, gender, task, and the institution information of **each public employee**. (It is assumed that no personal can simultaneously take on multiple job)
- Registration number and profession of **each doctor**. Each doctor is a public employee. Like public employee, each doctor also works at public institution; but it should be a hospital.
- The code, name and type of **each disease**.
- The code, name, type, and unit price of **each drug**.
- **Appointment information:** An appointment should keep these information: patient who asks appointment, *doctor* who responses, *date and time*, *appointment room*. It is assumed that an examination is performed at the hospital at which the doctor works; a patient can ask appointment from any doctor; a doctor can give appointment only a patient at a given moment.
- **Examination information:** *Patient* and *doctor*, *examination date* and *time*, *examination room*. The relation between examination and appointment is not one-to-one. In general, each appointment corresponds to examination unless patient breaks the appointment. In emergency case, no appointment is required for examination. It is assumed that a doctor can examine a patient at a given moment.
- **Disease information:** *Examination number*, *patient*, *disease name*. Making a diagnosis is done after examining the patient. A doctor can make one or more disease diagnosis.
- **Drug information:** *Examination number*, *patient*, *drug name*, *amount of drug provided*. A doctor can prescribe zero, one or more drugs.

Considering the functionalities and constraints above, you are to **construct an ER diagram**.

Grading:

- Entity sets: 50 p.
- Relations: 50 p.


hospital
number INT(11)
type VARCHAR(45)
institution_number INT(11)
Indexes
institution
number INT(11)
name VARCHAR(45)
address VARCHAR(45)
Indexes
personal
reg_number INT(11)
first_name VARCHAR(45)
last_name VARCHAR(45)
gender VARCHAR(45)
task VARCHAR(45)
institution_number INT(11)
Indexes
doctor
reg_number INT(11)
profession VARCHAR(45)
hospital_number INT(11)
Indexes
room
number INT(11)
location VARCHAR(45)
hospital_number INT(11)
Indexes
disease
code INT(11)
name VARCHAR(45)
type VARCHAR(45)
Indexes
examination_info
exam_date_time VARCHAR(45)
exam_number INT
patient_patientid INT(11)
room_hospital_number INT(11)
room_number INT(11)
doctor_reg_number INT(11)
Indexes
disease_inf
disease_code INT(11)
patient_has_doctor_exam_number INT
Indexes
appointment_info
doctor_reg_number INT(11)
patient_patientid INT(11)
room_hospital_number INT(11)
room_number INT(11)
app_date_time VARCHAR(45)
Indexes
patient
patientid INT(11)
p_name VARCHAR(45)
Indexes
drug
code INT(11)
name VARCHAR(45)
type VARCHAR(45)
unit_price INT(11)
Indexes
drug_info
drug_code INT(11)
examination_info_exam_number INT
amount INT
Indexes
1
1..*

Quiz 2 **BBM473 Database Management Sys. Lab.**

Name Surname:

1. Based on *Student database* below, you are expected to implement each query that carries out the operations given below.

   Faculty(FacultyID, Name, DeanRegNo)
   Department(DepartmentID, Name, HeadRegNo, FacultyNo)
   Student(StudentNo, Name, Surname, Sex, DateofBirth, EduPeriod, DepartmentNo)
   Instructor(InstructorRegNo, Name, Surname, Sex, Position, DepartmentNo)
   Course(CourseID, CourseName, THour, PHour, Credit, DepartmentNo)
   Section(CourseID, SectionID, Inst_inCharge)
   Stu_Course(StudentNo, CourseID, SectionID, MO, Grade)

a) Show the students' number, name, and surname whose department number is 1 and that have not taken any course offered by department 3.

b) Find out id and name of the departments having at least one student and offering no course.

c) Find out the departments' id, name, and research assistant count having at least two research assistants (abbreviated as *RA* in database).

d) Find out the students' number, name, and surname whose total credit is greater than 10.

[1] In order to learn how to load a database, stored as sql script file, to MySQL Workbench, perform the following steps.

Quiz 2 **BBM473 Database Management Sys. Lab.**

Name Surname:

1. Download database as sql file from

   vtlab.cs.hacettepe.edu.tr/files/BBM473-2015-2016BaharQuiz3relation.sql

2. Connect server, *localhost*.
3. From file menu, press **Open SQL Script**, *or use "Ctrl+Shift+O" buttons*.
4. From toolbar of script file, press **Execute** button ().
5. Database is loaded, but you may need to refresh your database list to see your loaded new database. Use **refresh** () button in schemas window.

Submission:

There are four questions each of which should be saved separately. To do that:

- You should take one screenshot for each question that shows your sql script and output.
- Then you should name each screenshot file with its question index. (e.g. "1a.png" for question a)
- Before submission, make all of your files compressed zip file and name it your department id beginning with "b".
- Send your zip file as an e-mail to **vtlab@cs.hacettepe.edu.tr.**

Quiz3 – Solutions

1. Based on *Student database*, you are expected to implement each query that carries out the operations given below.

   Faculty(FacultyID, Name, DeanRegNo)
   Department(DepartmentID, Name, HeadRegNo, FacultyNo)
   Student(StudentNo, Name, Surname, Sex, DateofBirth, EduPeriod, DepartmentNo)
   Instructor(InstructorRegNo, Name, Surname, Sex, Position, DepartmentNo)
   Course(CourseID, CourseName, THour, PHour, Credit, DepartmentNo)
   Section(CourseID, SectionID, Inst_inCharge)
   Stu_Course(StudentNo, CourseID, SectionID, MO, Grade)

a) Find out the number, name, and surname of the students who have taken all courses offered by "Computer" department.

```
select StudentNo, StudentName,StudentSurname from student where StudentNo in
(select student_StudentNo from student_course where
student_course.course_CourseID in  (select CourseID from course, department where
course.department_DepartmentID=department.DepartmentID and
department.DepartmentName='Computer') group by student_StudentNo having
count(*)=(select count(*) from course, department where
course.department_DepartmentID=department.DepartmentID and
department.DepartmentName='Computer'))
```

b) For each department stored in database, find out the summation of difference between theorical and practical hours of the course offered by that department.

```
select department.DepartmentName, sum(CourseTheoreticalH-CoursePracticalH) as
Difference from course, department where course.department_DepartmentID
=department.DepartmentID group by DepartmentName
```

c) Find out the number, name, and surname of the students whose average grade is over 50 (include 50) and department is either *Computer* or *Electric.* Note: Consider only mandatory courses taken for calculation of average grade. (Mandatory is abbreviated as *M* in database).

```
select StudentNo,StudentName,StudentSurname from student,department where
(department.DepartmentName="Computer" or
department.DepartmentName="Electric") and
student.department_DepartmentID=department.DepartmentID and student.StudentNo
in (select student_StudentNo from student_course where Man_or_Opt="M" group by
student_StudentNo having avg(grade)>=50);
```

Quiz3 – Solutions

d) Find out id and name of the courses whose all sections are given by only one research assistant. In other words, find out such a course's id and name that its sections are not taught by more than one research assistant. Note: A course having a single section should also be displayed if it is provided by a research assitant. (Research assistant is abbreviated as *RA* in database)

```
select CourseID, CourseName from course where CourseID in (select course_CourseID
from (select course_CourseID,count(*) from section,ınstructor group by
course_CourseID, ınstructor_InstructorRegNo) as temptbl group by course_CourseID
having count(course_CourseID)=1 ) and CourseID in (select distinct(course_CourseID)
from section, ınstructor where ınstructor.InstructorPosition="RA" and
section.ınstructor_InstructorRegNo=ınstructor.InstructorRegNo)
```

Quiz2 – Solutions

1. Based on *Student database*, you are expected to implement each query that carries out the operations given below.

   Faculty(FacultyID, Name, DeanRegNo)
   Department(DepartmentID, Name, HeadRegNo, FacultyNo)
   Student(StudentNo, Name, Surname, Sex, DateofBirth, EduPeriod, DepartmentNo)
   Instructor(InstructorRegNo, Name, Surname, Sex, Position, DepartmentNo)
   Course(CourseID, CourseName, THour, PHour, Credit, DepartmentNo)
   Section(CourseID, SectionID, Inst_inCharge)
   Stu_Course(StudentNo, CourseID, SectionID, MO, Grade)

a) Show the students' number, name, and surname whose department number is 1 and that have not taken any course offered by department 3.

```
select StudentNo, StudentName,StudentSurname from student where
student.department_DepartmentID=1 and StudentNo not in (select
student_course.student_StudentNo from course, student_course where
course.CourseID=student_course.course_CourseID and
course.department_DepartmentID=3)
```

b) Find out id and name of the departments having at least one student and offering no course.

```
select DepartmentID, DepartmentName from department where (select count(*)
from student where student.department_DepartmentID
=department.DepartmentID)>0 and (select count(*) from course where
course.department_DepartmentID=department.DepartmentID)=0
```

c) Find out the departments' id, name, and research assistant count having at least two research assistants (abbreviated as *RA* in database).

```
select DepartmentID, DepartmentName, count(*) as RA_Count from department,
ınstructor where department.DepartmentID=ınstructor.department_DepartmentID
and ınstructor.InstructorPosition="RA" group by department.DepartmentID having
count(*)>=2
```

d) Find out the students' number, name, and surname whose total credit is greater than 10.

```
select StudentNo,StudentName,StudentSurname
from student
 where StudentNo in (select student_StudentNo
                     From student_course, course
                     where student_course.course_CourseID=course.CourseID
                     group by student_course.student_StudentNo
                     having sum(all course.CourseCredit)>10)
```

Quiz2 – Solutions

Subject : Relational algebra

Advisor : Res. Assist. (Nebi Yılmaz, Merve Ozdes)
Entrant's ID: ..... Name & Surname: .....

Quiz 4

1. Considering the `University Schema` below:

   Classroom (building, room_number, capacity)
   Department (dept_name, building, budget)
   Course (course_id, title, dept_name, credits)
   Instructor (ID, name, dept_name, salary)
   Section (course_id, sec_id, semester, year, building, room_number, time_slot_id)
   Teaches (ID, course_id, sec_id, semester, year)
   Student (ID, name, dept_name, tot_cred)
   Takes (ID, course_id, sec_id, semester, year, grade)
   Advisor (s_ID, i_ID)
   Timeslot (time_slot_id, day, start_time, end_time)
   Prereq (course_id, prereq_id)

   Write the following queries in relational algebra, using the university schema.

   a. Find the names of all students who have taken at least one Comp. Sci. course.

   b. Find the ID s and names of all students who have not taken any course offering before Spring 2009.

   c. For each department, find the maximum salary of instructors in that department. You may assume that every department has at least one instructor.

   d. Find the lowest, across all departments, of the per-department maximum salary computed by the preceding query.

2. Find the results of the following relational algebra queries applied on the database provided thereafter:

P(A,B,C)

A	B	C
1	11	X
2	11	Y
3	11	X
4	12	Z

R(C,D)

X	1
X	2
Y	1

a. $X1 \leftarrow (\pi_{C,A} P) - R$

b. $X2 \leftarrow \pi_{B,C} P \div R$

c. $X3 \leftarrow \sigma_{A=D} (P \times R)$

d. $X4 \leftarrow (P \bowtie R)$

Grading:

- 60 p. for 1st answer (15 p. for each)
- 40 p. for 2nd answer (10 p. for each)

Quiz2 – Solutions

1.

a)

$$X1 \leftarrow \pi_{dept\_name}(\sigma_{dept\_name='Comp.Sci.'}(Department))$$

$$X2 \leftarrow \pi_{course\_id}(X1 \bowtie Course)$$

$$X3 \leftarrow \pi_{id}(X2 \bowtie Takes)$$

$$X4 \leftarrow \pi_{names}(X3 \bowtie Student)$$

b)

$$X1 \leftarrow \pi_{id}(\sigma_{(year<2009) \ \vee (year=2009 \wedge semester='fall')}(Takes))$$

$$X2 \leftarrow \pi_{id}(Student) - X1$$

$$X3 \leftarrow \pi_{id,name}(X2 \bowtie Student)$$

c)

$$X1 \leftarrow \pi_{id,salary,dept\_name}(Instructor)$$

$$X2(id\_buffer,salary\_buffer,dept\_name\_buffer) \leftarrow X1$$

$$X3 \leftarrow \pi_{id,salary,dept\_name}\ (X1 \bowtie_{salary<salary\_buffer \wedge dept\_name=dept\_name\_buffer} X2)$$

$$X4 \leftarrow \pi_{salary,dept\_name}\ (X1 - X3)$$

d)

$$X4 \leftarrow \pi_{id,salary}(X3^{*})$$ **calculated in (c)*

$$X5(id\_buffer,salary\_buffer) \leftarrow X4$$

$$X6 \leftarrow X4 \times X5$$

$$X7 \leftarrow \pi_{id}(\sigma_{(salary>salary\_buffer)}(X6))$$

$$X8 \leftarrow X4 - X7$$

Quiz2 – Solutions

2.

a)

Y	2
X	3
Z	4

b)

Such an operation cannot be performed since relation *R* has distinctive attribute labeled "*D*" so is not sub-part of relation *P*.

c)

1	11	X	X	1
1	11	X	Y	1
2	11	Y	X	2

d)

1	11	X	1
1	11	X	2
2	11	Y	1
3	11	X	1
3	11	X	2

Note: The solutions given here are not unique and there may exist solutions apart from these.

Quiz 5 **BBM473 Database Management Sys. Lab.**

Name Surname:

1. *The relation below records info of a car producer company and the sale operations*.

Customer Name	Model	Shipping Address	Producer	Phone	Price(x100$)
Alan Smith	Golf	35 Palm St, Miami	Volkswagen AG	(090) 555 6688	250
Roger Banks	Fiesta	47 Camp. Road, Boston	Ford MC	(090) 600 9090	300
Evan Wilson	Golf, Focus	28 Rock Av, Denver	Volkswagen AG, Ford MC	*Both*	450
Alan Smith	Fiesta	47 Camp. Road, Boston	Ford MC	(090) 600 9090	300

Considering that relation, perform:

a) 1NF decomposition:

b) 2NF (but not 3NF) decomposition:

c) 3NF decomposition:

2. Suppose that we decompose the schema *r (A, B, C, D, E)* into *r1 (D, A, E) and r2 (D, B, C)*

a) Show that this decomposition is a lossless decomposition if the following set *F* of functional dependencies holds:

F: $D \rightarrow AE$
$EB \rightarrow C$
$A \rightarrow B$
$C \rightarrow D$

b) Give a lossless-join decomposition into BCNF of schema R.

Quiz 6 **BBM473 Database Management Sys. Lab.**

Name Surname:

1. **Material(MaterialID, MaterialName, MaterialType)** (E / WE / R / DA)
   District(DistrictID, DistrictName) (E / WE / R / DA)
   Unit(DistrictID, UnitID, UnitName) (E / WE / R / DA)
   Stock(DistrictID, MaterialID, CurrentAmount, AmountIn, AmountOut) (E / WE / R / DA)
   Purchase(DistrictID, MaterialID, PurcDate, PurcAmount, UnitPrice) (E / WE / R / DA)
   Consumption(DistrictID, UnitID, MaterialID, ConsDate, ConsAmount) (E / WE / R / DA)
   Transfer(DistrictID1, DistrictID2, MaterialID, TransDate, TransAmount) (E / WE / R / DA)

   a) For each element given in the stock database above, state whether it is **E**ntity, **W**eak **E**ntity, **R**elation or having **D**escriptive **A**ttribute by circling appropriate options(s).

   b) Draw the Entity-Relational (ER) diagram of stock database above.

2.

   a) Write an appropriate question for the relational algebra below.

$$X1 \leftarrow \pi_{\text{DistrictID, MaterialID}}(\sigma_{(\text{TotalAmount}=0)}(\text{Stock}))$$

$$X2 \leftarrow X1$$

$$X3 \leftarrow ((\sigma_{(X1.\text{MaterialID} = X2.\text{MaterialID}) \wedge (X1.\ \text{DistrictID} \neq X2.\ \text{DistrictID})}\ (X1 \times X2)$$

$$X4 \leftarrow \pi_{X1.\text{MaterialID}}\ (X3)$$

$$X5 \leftarrow \pi_{\text{MaterialID, MaterialName}}\ (\text{Material} \boxtimes X4)$$

Name Surname:

b) Write a relational algebra that outputs the name of each district having all "food" type materials in its stock. (Take current amount of stock into consideration, which should be greater than 0)

3. Using an SQL query, find the name of each district such that there exists at least one material (current amount of material is greater than zero) in its stock.

4. Suppose you are given a relation R with four attributes ABCD. For each of the following sets of FDs, assuming those are the only dependencies that hold for R, do the following:

   i. $A \rightarrow B, BC \rightarrow D, A \rightarrow C$
   ii. $AB \rightarrow C, AB \rightarrow D, C \rightarrow A, D \rightarrow B$

   a. Identify the candidate key(s) for R.

   b. If R is not in BCNF, decompose it into a set of BCNF relations that preserve the dependencies.

1. Given the following description of the database, answer the following questions

"The Customer relation store all customers of the rental service, their names and the countries where they are located. Movie stores all movies available for rent, their titles and the year of production. When a customer rents a movie, a record is inserted in the relation Rent; a customer is allowed to rent a movie only once. After watching the movie, a customer is allowed to rate the movie, with a score from 1(worst) to 5 (best): this information is stored in the Rate table. Notice that a customer may rate a movie only if she/he has also rented the movie."

a) Draw the ER Diagram.
b) Write the relational schema.
c) The *Canadian rating* of a movie is the average score received by that movie from customers in Canada. Write a SQL query that computes the average *Canadian rating* of each movie; order your answers by the *Canadian rating*, starting from the highest rating.
d) Write a SQL query that returns the cid's and names of all customers who rented all the movies where "Alice" gave a score of 5. For example, if Alice gave a score of 5 to "Mad Max" and "Hunger Games" and gave no other score of 5, then your query should return the customers who rented both "Mad Max" and "Hunger Games".

2. Consider relation R(A, B, C, D, E) with functional dependencies (FDs):

S = {A→B; A→D; B→C; BC→D; BE→A}

a) What are the candidate keys of R?
b) A canonical cover also called as irreducible set is a simplified and reduced version of the given set of functional dependencies whose closure is exactly same as that of the closure of given set of functional dependencies. Compute the canonical cover of the set of FDs in S.
c) For each FD in T, state whether or not it violates BCNF and 3NF. Clearly explain your answer for each FD.

A→B	
B→C	
B→D	
BE→A	

d) Compute a decomposition of R into 3NF and explain your solution. Clearly mark the candidate key or keys of each relation in the result of the decomposition.

3. Consider the following database schema and relational algebra queries:

R(A,B), S(A,B,C), T(B,D,E)

$Q_1 = \sigma_{A=1,B>2}((R \bowtie S) \bowtie T)$ $Q_3 = \pi_E(\sigma_{D=1}(T \bowtie S))$

$Q_2 = (\sigma_{A=1,B>2}(R \bowtie S)) \bowtie T$ $Q_4 = \pi_B(S) \bowtie \pi_{B,E}(\sigma_{D=1}(T))$

a) Prove or disprove that Q1 and Q2 are equivalent.
b) Prove or disprove that Q3 and Q4 are equivalent.
c) If in part (a) or (b), the given pair of the queries are equivalent, explain which one might generate a faster evaluation plan than the other.

4. Briefly answer the following questions:

a) What are two physical properties that the relational database exploit, which are not visible at the logical (data model) level?
b) Can one look at a database instance and determine a set of functional dependencies? Why or why not?
c) Between dependency preservation and losslessness, which is more important and why?
d) Why is there a difference between Third Normal Form and Boyce-Codd Normal Form?
e) Describe how a relational database management system (DBMS) evaluates a query. (i) What happens from the moment the DBMS receives a SQL statement in the form of a string to the moment it returns results to the user? More complete answers will receive more points. (ii) How does the presence of an index affect what happens?