Instructions:

1. There are 12 questions in this question paper. All questions are compulsory.
2. For each question, please justify how you arrived at your answer. Incomplete answers will lead to loss of marks.
3. Assumptions made in each question should be clearly stated. The assumption should be rational and reasonably accurate. Wild assumptions should not be considered for arriving at an answer.
4. DO NOT COPY.

1. In terms of execution time performance, which of the two interfaces- Statement or Prepared Statement performs better and Why?

   Based on the answer given above, when (in what situations) should we use Statement and Prepared Statement interfaces respectively!   [3 Marks]

2. Consider the following relations for a company database: [5 Marks]

   employee (*employee_name*, street, city)
   works (*employee_name*, company_name, salary)

   where the primary keys are underlined. Using JDBC API, write a program to do the following tasks:
      a. find companies whose employees earn a higher salary, on average, than the average salary at "First Bank Corporation".
      b. Give all employees of "First Bank Corporation" a 10 percent raise unless the salary becomes greater than $100,000; in such cases, give only a 3 percent raise.
      c. Delete all tuples in the works relation for employees of "Small Bank Corporation"

   You can refer to the Java code given below, which uses the JDBC API. Assume that the userid, password, machine name, database, etc. are all okay.
   You only need to focus on the incomplete part of the code (//Write your code here...), and write appropriate code for each of the tasks mentioned above.

```
import java.sql.*;
public class Mystery {
        public static void main(String[] args) {
        try {
                Connection con=null; Class.forName("oracle.jdbc.driver.OracleDriver");
                con=DriverManager.getConnection("jdbc:oracle:thin:star/X@//edgar.cse.lehigh.edu:1521/XE");

                // Write your code here.....
```

} catch(Exception e){e.printStackTrace();} }

}

3. Consider the student table shown below with attributes - *stu_id int, stu_name varchar (20), marks int* **[2 Marks]**

A sample instance of the student table is shown below:

stu_id	stu_name	marks
101	Ravi	65
102	Sumit	32
103	Rekha	76
104	Yong	34
105	Hem	78

Write a query which gives the following output. In other words, a SQL query which displays the remarks and grades of all students along with their id, name, and marks, is required. You may use the following criteria to get the remarks and grades. If marks are greater than 32, then the student has passed, else failed. Along with this, if the marks are greater than 75, then grade is AA, and if the marks are greater than 64, then grade is AB, and if the marks are greater than 31, then grade is BB.

stu_id	stu_name	marks	remarks	grade
101	Ravi	65	Pass	AB
102	Sumit	32	Fail	BB
103	Rekha	76	Pass	AA
104	Yong	34	Pass	BB
105	Hem	78	Pass	AA

4. Consider the following relational database schema consisting of the four relation schemas: **[5 Marks]**

passenger (pid, pname, pgender, pcity)

agency (aid, aname, acity)

flight (fid, fdate, time, src, dest)

booking (pid, aid, fid, fdate)

Answer the following questions using relational algebra queries;

a. Get the details of flights that are scheduled on both dates 01/12/2022 and 02/12/2022 at 16:00 hours.

b. Get the details of flights that are scheduled on either of the dates 09/11/2022 or 22/11/2022 or both at 16:00 hours.

c. Find the agency names for agencies who do not have any bookings for passenger with id 123.

d. Find the passenger names for those who do not have any bookings in any flights.

e. Find the details of all male passengers who are associated with Jet agency.

5. Why would you choose a database system instead of simply storing data in operating system files? When would it make sense not to use a database system?   [6 Marks]
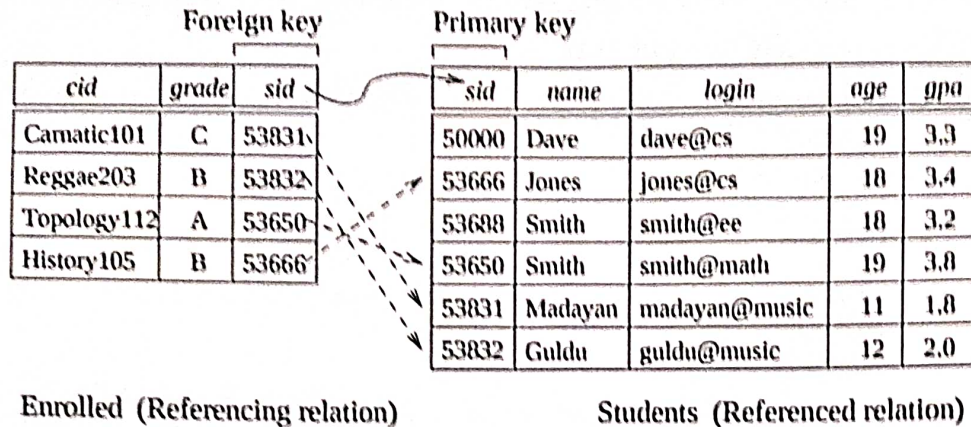
6. *Felonious Gru* wants to store information (names, addresses, descriptions of embarrassing moments, etc.) about the many minions on his payroll. Not surprisingly, the volume of data compels him to buy a database system. To save money, he wants to buy one with the fewest possible features, and he plans to run it as a stand-alone application on his PC clone. Of course, *Gru* does not plan to share his list with anyone. Indicate which of the following DBMS features *Gru* should pay for; in each case also indicate why *Gru* should (or should not) pay for that feature in the system he buys.   [3 Marks]

DBMS Features: 1. A security facility. 2. Concurrency control. 3. Crash recovery. 4. A view mechanism. 5. A query language.

7. Describe the structure of a DBMS in terms of various layers studied in the class. If your operating system is upgraded to support some new functions on OS files (e.g., the ability to force some sequence of bytes to disk), which layer(s) of the DBMS would you have to rewrite in order to take advantage of these new functions?   [4 Marks]

8. With respect to the following relations (Students and Enrolled), answer the following questions: [6 Marks]

a. What should we do if an Enrolled row is inserted, with a sid column value that does not appear in any row of the Students table?

b. What should we do if a Students row is deleted?

c. What should we do if the primary key value of a Students row is updated?

Foreign key   Primary key

cid	grade	sid
Carnatic101	C	53831
Reggae203	B	53832
Topology112	A	53650
History105	B	53666

sid	name	login	age	gpa
50000	Dave	dave@cs	19	3.3
53666	Jones	jones@cs	18	3.4
53688	Smith	smith@ee	18	3.2
53650	Smith	smith@math	19	3.8
53831	Madayan	madayan@music	11	1.8
53832	Guldu	guldu@music	12	2.0

Enrolled (Referencing relation)   Students (Referenced relation)

9. Explain the difference between logical and physical data independence.   [3 Marks]

10. The *semijoin* of relations R and S, written R ⋉ S, is the set of tuples t in R such that there is at least one tuple in S that agrees with t in all attributes that R and S have in common. Give two different expressions of relational algebra that are equivalent to R ⋉ S.   [3 Marks]

11. Consider general form of two relational algebra queries shown below: [4 Marks]

   a. $\pi_L(\sigma_C (R1 \times R2 \times R3 \ldots\ldots X R_n)$

   b. $\pi_L(\sigma_C (R1 \bowtie R2 \bowtie R3 \ldots\ldots \bowtie R_n)$

Here, L is an arbitrary list of attributes, and C is an arbitrary condition. The operators x and ⋈ represent the cartesian product and natural join operations respectively. Show how to express the two relational algebra queries shown above in SQL.

12. Let R be a relation with schema: (A1,A2, ... ,An,B1,B2, ..., Bm)

and let S be a relation with schema (B1,B2,....Bm); that is, the attributes of S are a subset of the attributes of R. The quotient of R and S, denoted R ÷ S, is the set of tuples t over attributes A1, A2, ... ,An (i.e, the attributes of R that are not attributes of S) such that for every tuple s in S, the tuple *ts*, consisting of the components of t for A1 ,A2,... ,An and the components of s for B1, B2,... , Bm, is a member of R.
Give an expression of relational algebra, using the operators we have studied in class, that is equivalent to R ÷ S.   [3 Marks]