CIS430/530 Database Systems FINAL

Last Name: ________________________________ First Name_______________________ ID: ________________

1. Answer T for True or F for False for the following statements from 1 to 10 as follow:

EMPLOYEE (fname, minit, lname, ssn, birthdate, address, sex, salary, superssn, dno) KEY: ssn
DEPARTMENT (dname, dnumber, mgrssn, mgrstartdate) KEY: dnumber. PROJECT (pname, pnumber, plocation, dnum) KEY: pnumber.
WORKS_ON (essn, pno, hours) KEY: (essn, pno)
DEPENDENT (essn, dependent-name, sex, bdate, relationship) KEY: (essn, dependent-name)

```
Create View VDept_Info (Dept_Name, No_of_Emps, Total_Sal) As
Select          D.dname, COUNT(E.dno), Sum(E.salary)
From            Department D LEFT OUTER JOIN Employee E,
Where          D.dnumber = E.dno
Group By       D.dname;
```

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Headquarters	1	888665555	19-Jun-71
Administration	4	987654321	01-Jan-85
Research	5	333445555	22-May-78
Automation	7	123456789	06-Oct-05

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	987654321	5
Franklin	T	Wong	333445555	08-Dec-45	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, SPring, TX	F	25000	987654321	4

[F] 1. COUNT (E.dno) does not count the duplicate dno values.

[T] 2. COUNT (E.dno) does not count Null values.

[F] 3. Group By will ignore NULL values if there are Nulls in the Group By column.

[F] 4. In three levels of the architecture of RDBMS, Table schema belongs to Internal Physical level.

[F] 5. Once a view is created with a Select query, it is final. You cannot change the content or the schema of the existing view unless you drop it and recreate again with the same view name.

[F] 6. View is temporary, so once it is created, it lasts only during the session. Once you log out, all the views you created since you logged in are dropped. Next time you log in again, they are gone from your database.

[T] 7. View is always updated. That is, if two newly hired employee tuples are inserted into the table Employee after the view VDept_Info was created, the column values for No_of_Emps and Total_Sal of the view VDept_Info are automatically updated with the changes of the Employee table when you submit the query: Select * from VDept_Info;

[F] 8. For the department dnumber 7 that no employee is working for, Sum (E.salary) will return 0.

[F] 9. When you want to change Dept_Name 'Automation' of the view VDept_Info to 'Development' after the view VDept_Info was created, you can directly update it to the view VDept_Info with "Update VDept_Info Set Dept_Name = …".

[T] 10.To process the query Q1 below, the nested Subquery in Q1 can be evaluated only once at the beginning, then the saved result of the subquery can be used to evaluate the rest of the outer query all at once.

```
 Q1:
SELECT      E.SSN, E.LNAME
 FROM       EMPLOYEE E
WHERE       1    <=          (SELECT
            COUNT(D.Essn)
                FROM   Dependent D
              WHERE  E.ssn = D.Essn );
```

Answer the following questions by writing the appropriate SQL query using the COMPANY database. The database schema consists of the following tables.

EMPLOYEE (fname, minit, lname, ssn, birthdate, address, sex, salary, superssn, dno) KEY: ssn
DEPARTMENT (dname, dnumber, mgrssn, mgrstartdate) KEY: dnumber. PROJECT (pname,
pnumber, plocation, dnum) KEY: pnumber.
WORKS_ON (essn, pno, hours) KEY: (essn, pno)
DEPENDENT (essn, dependent-name, sex, bdate, relationship) KEY: (essn, dependent-name)

2. Fill in the two blank lines 1) and 2) below with correct SQL clauses in the following SQL Q2:

For each department that has more than 2 employees, retrieve the department name and its employees (ssn and last name) who are making more than $40,000.

Q2:

```
Select      Distinct  D.dname, E.ssn, E.lname
From        Employee E, Department D
Where       E.salary > 40000 and
            E.Dno = D.Dnumber and
            E.Dno in (Select E1.Dno
                      From Employee E1, Department D1
                      Where E1. Dno = D1. Dnumber
                      1)  GROUP BY E1.Dno
                      2)  HAVING(E1.ssn)>2);
```

3. Answer the following questions on table T1 and table T2 below.

T1

T1.a1	T1.b1
1	1
2	1
3	2

T2

T2.a2	T2.b2
1	2
3	1
3	3
4	4

3-1) What is the result of the following query Q1? Show all the rows and columns with values in the result that Q1 will return.

Q1:
Select * From T1 Left Outer Join T2 On T1.a1 = T2.a2;

T1.a1	T1.b1	T2.a2	T2.b2
1	1	1	2
2	1	NULL	NULL
3	2	3	1
3	2	3	2

3-2) What is the result of the following query Q2? Show all the rows and columns with values in the result that Q2 will return.

Q2:
Select * From T1 Right Outer Join T2 On T1.a1 = T2.a2;

T1.a1	T1.b1	T2.a2	T2.b2
1	1	1	2
3	2	3	1
3	2	3	3
NULL	NULL	4	4

4. What will be the content of table ACCOUNT when Transaction 1 and Transaction 2 are done as below. Explain your answer briefly.

We assume that:

1) AccountID is declared as PK of ACCOUNT table and
2) The database server throws errors when trying to Insert, update or delete invalid rows or nonexisting rows in a database.
3) Each Transaction can start with its own transaction count and isolation status. We assume there is no commit (write) on ACCOUNT table since 1:00 pm when Transaction 2 reads (select) at the time 1:14 pm below, which means that Transaction 2 is allowed to read (Select) the main ACCOUNT table that everyone sees.

ACCOUNT ---- at the time 1:00 pm when Transaction 1 begins

AccountID	Balance
1111	15000
2222	25000

```
Transaction 1 by User 1:

BEGIN TRY
   BEGIN TRANSACTION    ---- at the time 1:00 pm
      DELETE FROM ACCOUNT Where AccountID = ‘1111’;
      INSERT INTO ACCOUNT VALUES (‘3333’, 35000);
      ROLLBACK;
                                      INSERT INTO ACCOUNT VALUES (‘4444’, 45000);
      SELECT * FROM ACCOUNT                                           ;   <----------- (1)
----                                  at the time 1:14 pm    COMMIT TRAN  ---- at the
time 1:20 pm
END TRY
BEGIN CATCH
   IF @@TRANCOUNT > 0
      ROLLBACK TRAN     --RollBack in case of Error
END CATCH
```

```
Transaction 2 by User 2:

BEGIN TRY
                                      BEGIN TRANSACTION   ---- at the time 1:14 pm
      SELECT * FROM ACCOUNT                       ;     <--------------- (2) at the time 1:14 pm
                                      INSERT INTO ACCOUNT VALUES (‘4444’, 45000);
      INSERT INTO ACCOUNT VALUES (‘5555’, 55000);
      COMMIT TRAN   <--------------- at the time 1:25 pm
END TRY
BEGIN CATCH
   IF @@TRANCOUNT > 0
      ROLLBACK TRAN     --RollBack in case of Error
END CATCH
```

4-1) What is the output of SELECT * FROM ACCOUNT;   <----------- (1) ---- at the time 1:14 pm in the Transaction 1 by User1 at the time 1:14 pm (AFTER Rollback and his last Insert and right BEFORE Commit)

AccountID	Balance
1111	15000
2222	25000
4444	45000

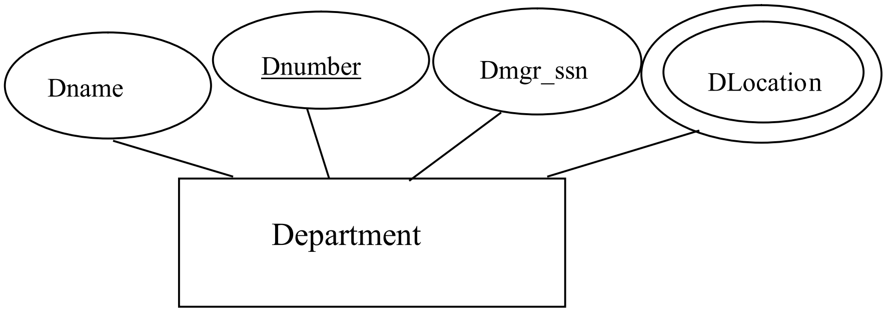
; <---4-2) What is the output of SELECT * FROM ACCOUNT -------- (2) ---- at the
time in the Transaction 2 by User2 at the time 1:14 1:14 pm
(right BEFORE he makes his changes and BEFORE COMMIT by Transaction 1)

AccountID	Balance
1111	15000
2222	25000

4-3) What is the output of SELECT * FROM ACCOUNT; to any users After Transaction 2 Committed at the time 1:25 pm (AFTER Transaction 2 Committed)

AccountID	Balance
1111	15000
2222	25000
4444	45000
5555	55000

5. The table created from the E-R diagram with the multivalued attribute below is valid? If not, transform the following table into a correct Database Schema. Show a correct Database Tables with the schema and the content

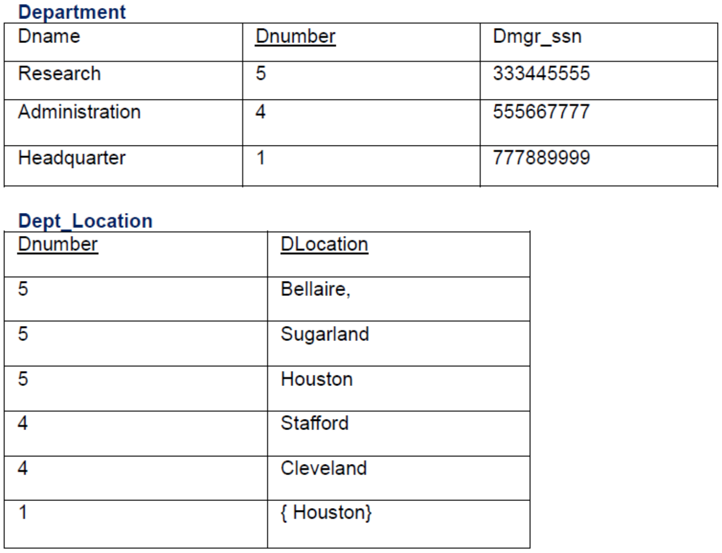


Department

Dname	<u>Dnumber</u>	Dmgr_ssn	DLocation
Research	5	333445555	{ Bellaire, Sugarland, Houston }
Administration	4	555667777	{ Stafford, Cleveland }
Headquarter	1	777889999	{ Houston}

Draw the correctly transformed schema of the table(s) for the Department above to create a correct database schema and Show all the tuple values together.

Mark(Underline) Primary Key of Each Table.
Specify Foreign Key as (FK) next to the column name if any in your table.



Department

Dname	<u>Dnumber</u>	Dmgr_ssn
Research	5	333445555
Administration	4	555667777
Headquarter	1	777889999

Dept_Location

<u>Dnumber</u>	<u>DLocation</u>
5	Bellaire,
5	Sugarland
5	Houston
4	Stafford
4	Cleveland
1	{ Houston}

6. JDBC (Java Database Connectivity) provides its API with Java.SQL classes to allow a java application to play as a Client to a RDBMS Server. Complete the following JDBC codes to write the output in a console as below.

```
…
import java.sql.Connection;  ...
public class MyJDBCSQLConnection {
public static void main(String[] args) throws ClassNotFoundException, SQLException {

Class.forName("com.mysql.jdbc.Driver");
Connection cnn = DriverManager.getConnection(
                                "jdbc:mysql://localhost:3306/company", "csuperson", "euclid");
if (cnn == null) {
System.out.println("Problems: No connection...");
System.exit(0);
}
System.out.println("Success: connection is available...");

Statement stmt = cnn.createStatement();

String mySQL = (1) "_Select * from employee where dno = ? _";

PreparedStatement stm = (2)_cnn.prepareStatement__ (mySQL, ResultSet.FETCH_FORWARD);

stm.setString(1, "4");

(3)_ResultSet_____________ rs  =  stm.executeQuery();

System.out.println("\nPreparedStatement - Employees working for Dno 4\n");

(4)_while (rs.next())_ {

          System.out
                    .println(rs.getString("ssn") + " "
                              + rs.getString("dno") + " "
                              + rs.getString("Fname") + " "
                              + rs.getString("Lname"));

}  rs.close();
cnn.close();
} //Main
```

Console Output:

PreparedStatement - Employees working for Dno 4

987654321 4 Jennifer Wallace
987987987 4 Ahmad Jabbar
999887777 4 Alicia Zelaya

Fill the missing codes at (1), (2), (3) and (4) in the JDBC codes above.

(1) Prepared SQL Query to Submit?

(2) Method call to submit your Dynamic SQL query with a parameter in (1)?

(3) Class Name that handles to process the Query Result returned?

(4) Missing a line of the code for a loop in (4) to process the query result per tuple in a table structure?

7. Suppose that you have a table Customer where:

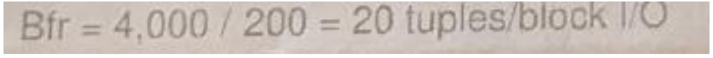
- Block size B = 4000 bytes
- (1 block is 4096 bytes but we will use 4000 bytes here for simplicity)
- Each record (tuple) size R of Customer Table = 200 bytes
- The number of records r = 40000 records (tuples) in the Customer table
- The search Key of the Customer Table is LName

blocking factor bfr means how many tuples of a table can fit in one block size(4000 bytes) in a disk of the database file system.

Q:
Select * From Customer Where Customer.Lname = “Smith”;

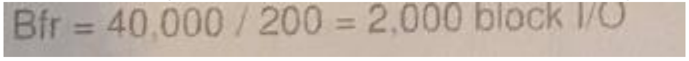
7-1) To find a record for a Customer Lname = Smith,

a) What is your blocking factor bfr (# of tuples per block) for the table Customer?



Bfr = 4,000 / 200 = 20 tuples/block I/O

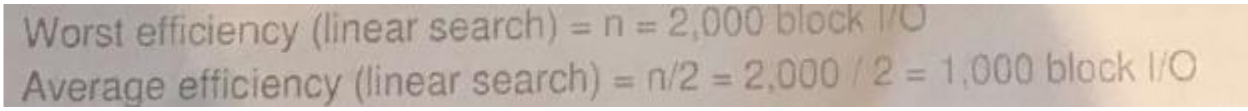
b) How many blocks in disk are needed to store your entire Customer able?



Bfr = 40,000 / 200 = 2,000 block I/O

c) How many block access (disk I/O accesses) would be needed to search (find) a record for the Customer Smith when there is no Index built on the search key column Lname and the Lname field is not sorted in the disk?

Note that this is an average sequential search cost without Index built on the column Lname of the table that is not sorted in the disk.

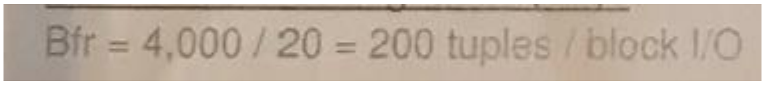


Worst efficiency (linear search) = n = 2,000 block I/O
Average efficiency (linear search) = n/2 = 2,000 / 2 = 1,000 block I/O

- Block size B = 4000 bytes (to simplify calculations)
- Each record (tuple) size R of Customer Table = 200 bytes
- The number of record r = 40000 records (tuples) in the Customer table
- The search Key of the Customer Table is LName
- Customer table with the Lname field is sorted in the disk

7-2) To find a record for a Customer Lname = Smith,
When an index (in an index table format) is built on the key column: Lname field with each index entry has the column size VLname=10 bytes and a block pointer column (as a record pointer) PR =10 bytes, so each index entry (tuple) size is (10 + 10) = 20 Byte:

a) What is blocking factor BfrI (# of index entries per block) for your index file (table)??

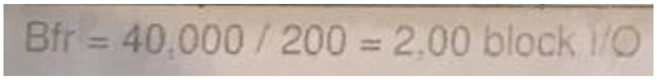


Bfr = 4,000 / 20 = 200 tuples / block I/O

b) How many blocks are needed to store your index file when your Customer table with the Lname field is sorted in the disk.

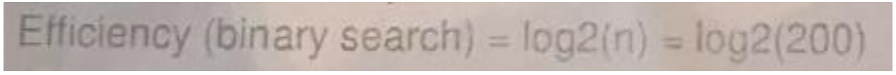
Hint: Calculate how many index entries need to be stored when the Lname field is sorted in Customer table in disk. Decide whether you need one index entry per each record of the Customer table in disk or one index entry per block of the Customer table in disk ?

Your index file size = index entry size (20 byte) * total # of index entries needed



Bfr = 40,000 / 200 = 2,00 block I/O

c) How many disk block access would needed to search (find) a record for the Customer Smith using such Index built on Lname as in b) when Lname is sorted in the Customer table in disk? (Hint: Index file is always sorted on the search key column values so you can use a search algorithm to find a last name)



Efficiency (binary search) = log2(n) = log2(200)