Solved Exercise from Sumita Arora

DATABASE CONCEPT

TYPE A : VERY SHORT ANSWER QUESTION

1.	**What is the function of database management system ?**
Ans.	Following are the function of database management system: - To organize information - To store information - To retrieve information As efficiently and effectively as possible
2.	**What is data redundancy? What are the problems associated with it?**
Ans.	Data redundancy means duplication of data. It causes duplicate data at different locations which destroys the integrity of the database and wastage of storage space.
3.	**How do database management system overcome the problems associated with data redundancy ?**
Ans.	Database normalization prevents redundancy and makes the best possible usage of storage .Proper use of foreign keys can minimize data redundancy and chance of destructive anomalies. However, concerns of efficiency and convenience can sometimes result in redundant data design despite the risk of corrupting the data. A solution to the problem is to place the redundant data in a separate table, one in which the data no longer will be redundant.
4.	**How do database management system ensure data security and privacy ?**
Ans.	A database management system ensure data security and privacy by ensuring that only means of access to the database is through the proper channel and also by carrying out authorization checks whenever access to sensitive data is attempted.
5.	**What are the three levels of data abstraction?**
Ans.	**a.** Internal Level **b.** Conceptual Level **c.** External Level (View Level)
6.	**What do you understand by data independence?**
Ans.	Data independence allows modification of a scheme definition in one level without affecting a scheme definition in the next higher level.
7.	**What are the two types of data independence? How they different?**
Ans.	There are two levels of data independence: Physical and Logical. Physical Data Independence modifies the scheme followed at the physical level without affecting the scheme followed at the conceptual level. Whereas Logical Data Independence modifies the conceptual scheme without causing any changes in the schemes followed at view levels.
8.	**Define the following terms :**
Ans.	**(i) Integrated database:** Integrated database means union of several otherwise distinct data files, with any redundancy among those files partially or wholly eliminated. **(ii) Shared database:** Sharing of data means that individual pieces of data in the database may be shared among several different users, in the sense that each of those users may have access to the same piece of data and each of them may use it for different purpose. The database which contains this type of data is called shared database. **(iii) View:** A view is a pseudo-table or virtual table. It displays the data. The data is derived from one or more base tables. **(iv) Database system:** A database is a collection of interrelated data and a database system is basically a computer base recordkeeping system. **(v) Data security:** Data security refers to protection of data against accidental or intentional disclosure to unauthorized persons, or unauthorized modification or destruction. **(vi) Data integrity:** Data integrity refers to maintaining and assuring the accuracy and consistency of data over its entire life cycle
9.	**Name the different data models available for database system. Which of them is the most preferred one ?**
Ans.	There are generally three data models available as following : (i)Relational Model

	(ii) Network Model (iii) Hierarchical Model Among these three level relational level is the most preferred one because it does not use pointers or links. Instead, it relates records by the values they contain.
10.	**What are the similarities and differences between network and hierarchical data models ?**
Ans.	The network model and hierarchical model are similar in the sense that data and relationships among data are represented by records and links respectively. The difference between these two is that in network model records are organized as collection of arbitrary graphs and in hierarchical model records are organized as collection of trees.
11.	**Define the following terms :**
Ans.	**(a) Relation**: A relation is a two-dimensional table. It contains number of rows(tuples) and columns(attributes). **(b) Tuple**: This is the horizontal part of the relation. One row represents one record of the relation. The rows of a relation are also called tuples. **(c) Attribute**: The columns of a table are also called attributes. The column is the vertical part of the relation. **(d) Domain**: A domain is a pool of values from which the actual value present in a given column are taken. **(e) Primary key**: It is a column (or columns) in a table that uniquely identifies each row. A primary key value is unique and cannot be null. There is only one primary key for a table. **(f) Candidate key**: It is a column (or columns) that uniquely identify rows in a table. Any of the identified candidate keys can be used as the table's primary key. **(g) Cartesian product**: The Cartesian product is a binary operation and is denoted by a cross(x). It yields a relation with all possible combinations of the tuples of the two relations operated upon. **(h) degree**: It is the number of columns (or attributes) in a table.
12.	**What do you mean by relational database?**
Ans.	A relational database is a collection of data items organized as a set of formally described tables from which data can be accessed easily.
13.	**What do you understand by unary relation, binary relation?**
Ans.	**Unary relation:** If a relation is between entities in a single entity type then it is called a *unary relation*. For example `friendship' between entities within the entity type PERSON. **Binary relation:** If a relation is between entities in one entity type and entities in another entity type then it is called a *binary relation* because two entity types are involved in the relationship type. For example is 'Purchased' between entity types CUSTOMER and PRODUCT.
14.	**What are views? How are they useful?**
Ans.	A view is a pseudo-table or virtual table. It displays the data. The data is derived from one or more base tables. **Syntax:** CREATE VIEW <view name> AS SELECT <attribute list> FROM <table(s)> WHERE <condition(s)>; The usefulness of views lies in the fact that they provide an excellent way to give people access to some but not all of the information in a table
15.	**Define the following :**
Ans.	**(i) Primary key**: It is a column (or columns) in a table that uniquely identifies each row. A primary key value is unique and cannot be null. There is only one primary key for a table. **(ii) Candidate key**: It is a column (or columns) that uniquely identify rows in a table. Any of the identified candidate keys can be used as the table's primary key. **(iii) Alternate key** : A candidate key that is not the primary key is called an Alternate Key. **(iv) Foreign key**: It is a column (or a set of columns) that refers to the primary key in another table i.e. it is used as a link to a matching column in another table.
16.	**What is an Alternate Key?**
Ans.	A candidate key that is not the primary key is called an Alternate Key. For example in Student table if there are two candidate keys – StudId and Stud_Name and StudId is the primary Key then Stud_Name is the alternate key.
17.	**What is the importance of a primary key in a table ? Explain with a suitable example**.
Ans.	A primary key is a set of one or more attributes that can uniquely identify tuples within the relation.

For example, in the following table **ITEM**, the column **Ino** can uniquely identify each row in the table, hence **Ino** is the primary key of the following table.

Ino	Item	Quantity
1	Pen	220
2	CD	530
3	DVD	450
4	Duster	110

18. What do you understand by the terms Primary Key and Degree of a relation in relational database?

Ans. A **primary key** is a column (or columns) in a table that uniquely identifies each row. A primary key value is unique and cannot be null. There is only one primary key for a table.

Number of columns(attributes) in a relation are called its **degree**.

19. What do you understand by the terms Candidate Key and Cardinality of relation in relational database?

Ans. A **candidate key** is a column (or columns) that uniquely identify rows in a table. Any of the identified candidate keys can be used as the table's primary key.

Cardinality of relation represents number of rows (or tuples) in a relation.

20. What is the function of select and project operation ?

Ans. The select operation selects tuples from a relation that satisfy a given condition. The selection is denoted by lowercase Greek letter ± (sigma).

The project operation yields a "vertical" subset of a given relation in contrast to the "horizontal" subset returned by select operation. The projection lets you select specified attributes in a specified order. Projection is denoted by Greek letter pi(π).

21. What is the Cartesian product? How is it different from join operation?

Ans. When you join two or more tables without any condition, it is called Cartesian product or Cross Join.

The difference between Cartesian Product and Join is – In Join operation we have to mention the column name from both the tables on which join has to be perform for eg. Select EmpName,Salary from Employee, SalaryRange where Employee.Salary =SalaryRange.MinSalary.

Whereas in Cartesian product we do not mention any condition for eg. Select a.Id,b.Id from Employee a,Manager b;

22. What does union operation do? What are the conditions to carry out union operation?

Ans. The union operation is a binary operation that requires two relations as its operands. It produces a third relation that contains tuples from both the operand relations. The union operation is denoted by ∪. Thus, to denote the union or two relations **X** and **Y**, we will write as **X ∪ Y.**

Following two condition must be satisfied by two operands A and B :

(i)The relations A and B must be of the same degree. That is, they must have the same number of attributes.

(ii) The domains of the *i*th attributes of A and the *i*th attributes of B must be the same.

23. What are the set difference and set intersection operations? How are they different?

Ans. The set difference operation, denoted by – (minus) allows us to find tuples that are in one relation but not in another. The expression **A – B** results in a relation containing those tuples in **A** but not in **B**.

The set intersection operation finds tuples that are common to the two operands relations. The set intersection operation is denoted by ∩. That means **A ∩ B** will yield a relation having tuples common to **A** and **B**.

The difference between these two is that in set difference operation the result will contain those tuples which are in one operand but not in other.

Whereas in set intersection operation result will contain those tuples which are common in both the operands.

Give the following relations

EMPL (FirstName, LastName, Workdept, Empno, Educlevel, Jobcode, Salary, Bonus, hiredate)

Dept (Deptno, Deptname, Mgrno, Adurdept)

Statement for creating above table :

```
CREATE TABLE EMPL
    (       FirstName       char(20),
            LastName        char(20),
            WorkDept        char(15),
            Empno           integer,
            Educlevel       char(10),
            Jobcode         integer,
            Salary          decimal,
            Bonus           decimal,
            Hiredate        date);

CREATE TABLE Dept
    (       Deptno      integer,
            DeptName        char(20),
            Mgrno           integer,
            Adurdept        char(20));
CREATE TABLE Supplier
    (       Sno         integer,
            Pno         integer);
CREATE TABLE Part
    (       Pno         integer,
            Colour      char(10));
```

Write relational expression for question 24 – 27 using the above relations.

24. **List employees' first and last names.**

Ans. SELECT FirstName, LastName from EMPL;

25. **Give all information on employees in manufacturing system.**

Ans. SELECT * from EMPL WHERE Workdept = "Menufacture";

26. **Just give employee names in manufacturing systems.**

Ans. SELECT FirstName from EMPL WHERE Workdept IN ("Menufacture");

27. **Give employee names outside manufacturing systems.**

Ans. Ans : SELECT FirstName from EMPL WHERE Workdept NOT IN ('Menufacture');

28. **List employee names and their departments for employee numbers between 100 and 300**

Ans. SELECT EmpName, DeptName from EMPL WHERE Empno BETWEEN 100 and 300;

30. **List employee numbers, names and educational levels for three educational 16, 18, and 20.**

Ans. SELECT Empno, FirstName, Educlevel from EMPL WHERE Educlevel IN (16, 18, 20);

31. **List employee numbers and name in manufacturing system for job code 54.**

Ans. SELECT Empno, FirstName from EMPL WHERE Workdept = 'Menufacture' AND Jobcode = 54;

32. **Same request as above with educational level below 15.**

Ans. SELECT Empno, FirstName from EMPL WHERE Workdept = 'Menufacture' AND Jobcode =< 15;

33. **List those employees' last name and monthly salary plus bonus, whose job code is at least three tmes their educational level and whose annual basic salary plus bonus exceed 100000.**

Ans. SELECT LastName, Salary+Bonus from EMPL WHERE Jobcode = (Educlevel * 3) AND ((Salary * 12) + Bonus) > 100000;

34. **List employee last name, work department and monthly salary for those in Services, Planning and Information center departments.**

Ans. SELECT LastName, Workdept , Salary from EMPL WHERE Deptname IN ('Services', 'Planning', 'Information center');

35. **List employee last name, work department, department number, and department name.**

Ans. SELECT LastName, Workdept , Deptno, Deptname from EMPL, Dept WHERE Empl.Deptno = Dept.Deptno ;

36. **Which employees earn more than their manager ?**

Ans. SELECT EmpName from EMPL, MANG WHERE empl.mgrid=mang.mgrid and EMPL.Salary > MANG.Salary ;

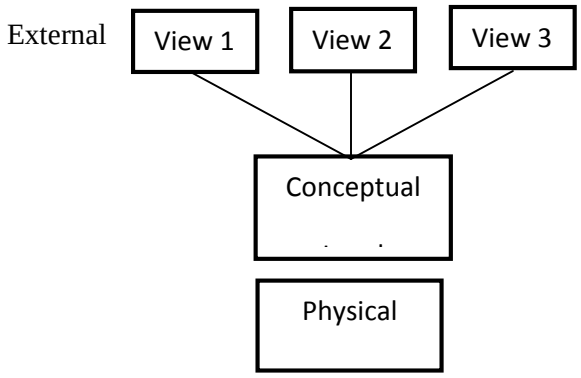
37. **List suppliers who supply only red parts.**

Ans.	SELECT SupName from Supplier WHERE colour = 'Red' ;
38.	**What is meant by functional dependency? Give example of functional dependence .**
Ans.	A functional dependency occurs when one attribute in a relation uniquely determines another attribute. This can be written A -> B which would be the same as stating "B is functionally dependent upon A." **Examples:** In a table listing employee characteristics including Social Security Number (SSN) and name, it can be said that name is functionally dependent upon SSN (or SSN -> name) because an employee's name can be uniquely determined from their SSN. However, the reverse statement (name -> SSN) is not true because more than one employee can have the same name but different SSNs.

TYPE B : SHORT ANSWER QUESTIONS

1. Draw a diagram explaining various levels of data abstraction.

Ans.



The Three Levels of Data Abstraction.

2. What is meant by "data independence"? Make a list of data independence capabilities.

Ans. Data independence is the ability to modify a scheme definition in one level without affecting a scheme definition in a higher level.

data independence capabilities :

- ✓ Physical Data Independence has ability to modify the scheme followed at the physical level without affecting the scheme followed at the conceptual level.
- ✓ Logical Data Independence has ability to modify the conceptual scheme without causing any changes in the schemes followed at view levels.

3. What are the main difference between a file-processing system and a database management system?:

Ans.

File-processing system	Database management system
It has problem of data redundancy.	It can reduce data redundancy to a large extent.
Data is insecure.	Provide more security to data.
Cannot store data permanently.	Can store data permanently.
Cannot change and access file-oriented data	Can change and access data in DBMS.
Cannot recall or recover the file-oriented data	DBMS solve such type of problem

4. This chapter has described some of the major advantage of a database system. What are the disadvantage ?

Ans. Following are the some of disadvantage of database system :

(a) Security may be compromised without good controls.
(b) Integrity may be compromised without good controls.
(c) Extra hardware may be required.
(d) Performance overhead may be significant.
(e) System is likely to be complex.

5. Explain deference between physical and logical data independence.

Ans.

Physical data independence	Logical data independence
Modifies the scheme followed at the physical level without affecting the scheme followed at the conceptual level.	Modifies the conceptual scheme without causing any changes in the schemes followed at view levels.
Less difficult to achieve.	More difficult to achieve because the
Application programs are not heavily dependent on the physical structure of the database.	Application programs are heavily dependent on the logical structure of the database.

6. Illustrate the difference between the three levels of data abstraction.

Ans.

Internal Level	Conceptual Level	External Level
Describes how the data is actually stored on the storage medium.	Describes what data are actually stored in the database and relationship existing among data.	Concerned with the data is viewed by individual users.
At this level, complex low-level data structure are described in details.	At this level, the database is described logically in terms of simple data-structures.	Only a part of the database relevant to the users is provided to them through this level.

7. What is relation? What is the difference between a tuple and an attribute?

Ans. A relation is a two-dimensional table. It contains number of rows (tuples) and columns (attributes).
This is the horizontal part of the relation. One row represents one record of the relation. The rows of a relation are also called tuples.

8. What is relation? Define the properties of a relation.

Ans. A relation is a two-dimensional table. It contains number of rows (tuples) and columns (attributes).
A relation has following properties :
(a). In any given column of the table, all items are of the same kind whereas items in different columns may not be of the same kind.
(b). For a row, each coloums must have an atomic value and also for a row, a columns cannot more than one value.
(c). All of a relations are distinct. That is, a relation does not contain two rows which are identical in every column. That is, each row of the relation can be uniquely identified by its contents.
(d). The ordering of two within a relation is immaterial. That is, we cannot retrieved any thing by saying that from row number 5, column name is to be accessed. Thee is no order maintained for rows inside a relation.
(e). The columns of a relation are assigned distinct names and the ordering of these columns is immaterial.

9. Summarise the major difference between a relation and a traditional file.

Ans. File System vs Data base Management System: -
1. Files act locally where as DBMS saves directly in a database
2. Saves in temporary locations where as DBMS in well arranged and permanent data base locations
3. In File System., transactions are not possible where as various transactions like insert, delete, view, updating etc are possible in DBMS
4. Data will be accessed through single or various files where as in DBMS, tables (schema) is used to access data.
5. A "File manager" is used to store all relationships in directories in File Systems where as a data base manager (administrator) stores the relationship in form of structural tables
6. Last.... but not the least.... Data in data bases are more secure compared to data in files!!