1. What is database? What is mean by DBMS. Write down the advantages of DBMS over file management system.

Data is the plural form of datum, which means a single logical interrelated data and a description of these data, designed to meet the information needs of an organization.

DBMS is the abbreviations of Database Management System. It is defined as the software system that allows user to define, create, maintenance and control access to the database. In another since DBMS is an intermediate layer between programs and data.

The following are the different advantages of DBMS over file management system—

1) Reduction of Redundancy: Centralized control of data by DBA (Database Administrator) avoids unnecessary duplication of data and effectively reduces the total amount of data storage required. It also element the extra processing necessary to trace the required data in a large mass of data. Another advantages of avoiding duplication is, the elimination of inconsistencies that tend to be present in redundant data files.

2) Shared data: A database allows the sharing of data under its control by any no of application, programs or users.

3) Integrity: Centralized control can also ensure that adequate checks are incorporated in the DBMS to provide data integrity. Data integrity means that the data contained in the database is both accurate and consistence. Therefore, data values being entered for storage could be checked to ensure that they fall within a specified range and are of correct format. Another integrity that should be incorporated in the database is to ensure that if there is a reference to certain object, hat object must exist.

4) Security: Data is a vital importance of an organization and may be confidential. Such confidential data must not have accessed by an unauthorized person. The DBA who has the ultimate responsibility for the data in the DBMS, can ensure that proper access procedures are followed, including proper authentication schema for access to the DBMS, and additional checks before permitting access to sensitive data. Different level of security could be implemented for various type of data and applications.

5) Conflict Resolution: Since the data is under the control of DBA, he should resolve the conflicting requirement of various user and application.

6) Data Independency: Data independency is considered form two points of view: physical data independency and logical data independency. Physical data independency allows changes in physical storage devices or organization of the files to be made without requiring changes in the conceptual view or any of the external views. Logical data independence implies that application programs need not to be changed if fields are added to an existing record nor do they have to changed, if fields not used by application program are deleted.

2. Explain three level architecture of DBMS.

A database system is a collection of interrelated data and a set of programs that allow users to access and modify these data. A major purpose of a database system is to provide users with an abstract view of the data. That is, the system hides certain details of how the data are stored and maintained.

Many database-system users are not computer trained, developers hide the complexity from users through several levels of abstraction, to simplify users' interactions with the system—

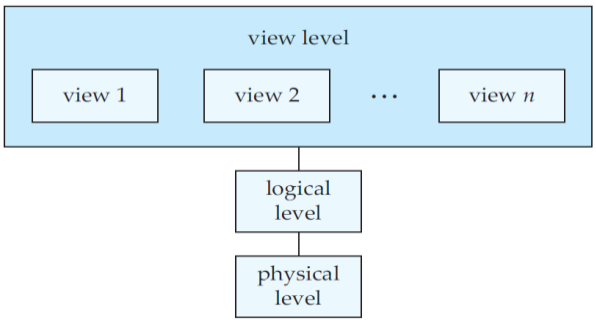
1) Physical Level (Internal view): We find this view at the lowest level of abstraction, closes to the physical storage method used. It indicates how the data will be stored and describes the data structure and access method used by the database. The internal view is expressed by the internal schema, which contains the definition of the stored record, the method of representing the data fields and the access aids used.

2) Logical Level (Conceptual/Global Layer): At this level of database abstraction all the database entities and the relationship among them are included. One conceptual view represents the entire database. This conceptual view is defined by conceptual schema, it describes all the records and relationship included in the conceptual view and, therefore, in the database. There is only one conceptual schema per database. This schema also contains the method deriving the objects in the conceptual view form the objects in the internal view.

3) View Level (External or user view): The external or user view is at the highest level of data abstraction, where only those portion of the database of concern to a user or application program are included. Any number of user view may exists for a given global conceptual view.

Each external view is described by means of a schema called external schema. The external schema consist of the definition of the logical records and the relationship in the external view. The external schema also consist the method of deriving the objects in the external view from the objects in the conceptual view. The objects include entities, attributes, and relationships.

The interrelationship among these three levels of abstraction is shown in the following figure—



3. Define Instance and Schema.

Databases change over time as information is inserted and deleted. The collection of information stored in the database at a particular moment is called an *instance* of the database. The overall design of the database is called the *database schema*.

Database systems have several schemas, partitioned according to the levels of abstraction. At the lowest level is the physical schema; at the intermediate level is the logical schema and at the highest is a subschema. In general, database system supports one physical schema, one logical schema and several subschemas.

4. What do you mean by data independence? What are different level of independence.

The ability to modify a schema definition in one level without effecting a schema definition in the next higher level is called data independence. There are two level of data independence.

1) Physical Data Independence: It is the ability to modify the physical schema without causing application programs to be rewritten. Modification at physical level are occasionally necessary to improve performance.

2) Logical Data Independence: It is the ability to modify the logical schema without causing application programs to be rewritten. . Modification at logical level are necessary whenever the logical structure of the database is altered.

Logical data independence is more difficult to archive than the physical data independence. Since application programs are heavily dependent on the logical structure of the data than they access.

5. Write short note on (i) DDL (ii) DML.

1) DDL: The DDL is the abbreviation of Data Definition Language. A database schema is specified by a set of definition expressed by a special language called a DDL . The result of compilation of DDL statements is a set of tables that is stored in a special file called *data dictionary* or the *data directory.*

A data dictionary is a file that contains metadata – that is data about data. This file is consulted before actual data are read or modified in the database system.

The storage structure and the access method used by the database system are specified by a set of definitions in a special type of DDL called data storage and definition language. The result of compilation of these definition is a set of instruction to specify the implementation details of the database schemas.

Some of the DDL commands of SQL are – CREATE, ALTER, DROP etc.

2) DML: DML is the abbreviation of Data Manipulation Language. By data manipulation we mean

- The Retrieval of information stored in the database.
- The Insertion of new information into the database.
- The Deletion of information from the database.
- The Modification of information stored in the database.

DML is a language that enables users to access or manipulate data as organized by the appropriate data model. There are basically two types:

- Procedural DMLs require a user to specify what data are needed and how to get those data.
- Declarative DMLs (also referred to as nonprocedural DMLs) require a user to specify what data are needed without specifying how to get those data.

6. Define DBA. Write down the responsibilities of DBA.

One of the main reason of using DBSs is to have central control of both the data and the programs that access those data. The person who has such central control over the system is called the database administrator (DBA).

The functions of the DBA include the following—

1) Schema Definition: The DBA creates the original database schema by writing a set of definitions, that is translated by the DDL compiler to set of tables that is stored permanently in the data dictionary.

2) Storage Structure and Access Method Definition: The DBA creates appropriate storage structure and access method by witting a set of definitions which is translated by data storage and DDL compiler.

3) Schema Physical Organization Modification: Programmers accomplish the relatively rare modifications either to the database scheme or to the description of the physical storage organization by writing a set of definitions that is used by either the DDL compiler or the data-storage and Data Definition Language compiler to generate modification to the appropriate internal system tables.

4) Granting of authorization for data access: The granting of different type of authorization allows the DBA to regulate which part of the database various user can access. The authorization information is kept in a special system structure that is consulted by the database system whenever access to data is attempted in the system.

5) Integrity Constrain Specification: the data values in the database must satisfy certain consistency constrain. For example, perhaps the number of hours an employee may work in one weak may not excide a specific limit. Such a constrain must specified explicitly by the DBA. The integrity constrain are kept in a special system structure that is consulted by the database system whenever an update takes place in the system.

7. Explain the overall structure of a DBMS.

A database system into modules that deal with each of responsibilities of the overall system. Some of the functions of the database system may be provided by the computer's operating system. In most cases, the computer operating system provides only the most basic services, and the database system must build on that base. Thus, the design of database system must include consideration of interface between database system and operation system.

The functionality of a database system can broadly divide into query processor components and storage manager components.

The query processor includes—

1) DML Compiler: It translates DML statements in a query language into low level instructions that the query evaluation engine understand. In addition, The DML compiler attempts to transform a user's request into an equivalent but most efficient form. Thus finding a good strategy for executing the query.

2) **Embedded DML pre-compiler:** It converts DML embedded in an application program to normal procedure calls in the host language. The pre-compiler much interacts with the DML compiler to generate appropriate code.

3) **DDL interpreter:** It interprets the DDL statements and record the in a set of tables containing metadata.

4) **Query evaluation engine:** It executes low level instruction generated by the DML compiler.

The storage manager components provide the interface between the low level data stored in the database and the application program and queries submitted to the system. The storage manager components include—

1) **Authorization and integrity manager:** It tests for the satisfaction of the integrity constrains and checks the authority of the users to access data.

2) **Transaction manager:** It ensure that database remains in a consistence state despite system failure and that concurrent transaction executing proceed without conflicting.

3) **File manager:** It manages the allocation of space on disk storage and data the data structures used to represent information needed on disk.

4) **Buffer Manager:** It is responsible for fetching data from disk storage into main memory, and deciding what data to cache in memory.

In addition, several data structures are required as part of the physical system implementation—

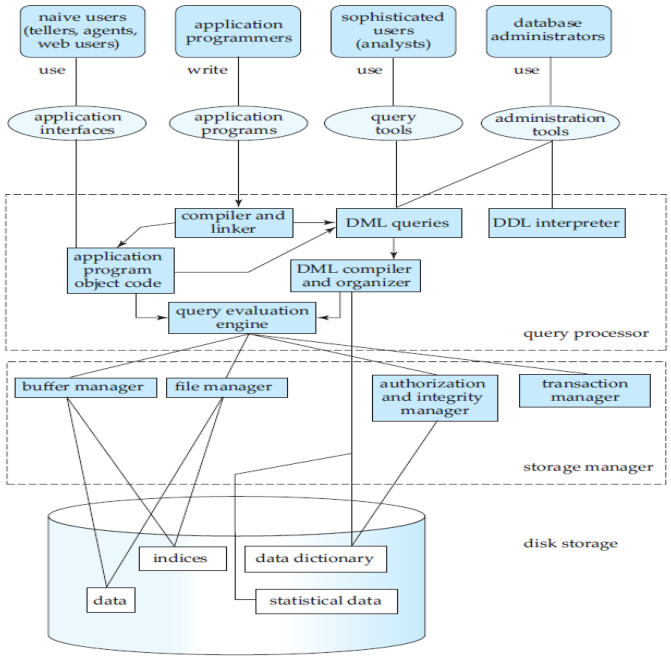
i) Data Files: It stores the database itself.

ii) Data Dictionary: It stores metadata about the structure of the database. The data dictionary is used heavily.

iii) Indices: It provides fast access to data items that hold particular values.

iv) Statistical Data: It stores statistical information about the data in the database. This information is used by the query processor to select effective ways to execute queries.

The following figure shows the various components of DBMS and the connections among the components—



8. What do you mean by data model? Explain various types of data model.

Underlying the structure of DBA is the data model – a collection of conceptual tools for describing data, data relationships, data semantics and consistency constraint.

The various data model that have been proposed fall into 3 different groups—

1) Object Based Model: Object based logical models are used in describing data at the logical and the view levels. They are characterized by the fact that they provide fairly flexible structuring capabilities and allow data constraint to be specified explicitly.

Some of the popular object based logical models are—

a) The entity relationship model.
b) The object oriented model.
c) The semantic data model.
d) The functional data model.

2) Record based logical models: Record based data models are used in describing data at the logical and view levels. In constraint to object based models, they are used both to specify the overall logical structure of the database and to provide a high-level of description of the implementation.

Record based models are so named because the database is structured in fixed format records of several types. Each record type defines a fixed no of field, attributes, and each field is usually of fixed length.

The three most widely accepted record based data models are—

i) Relational Model: The relational model uses a collection of table to represent both data and the relationship among those data. Each table has multiple columns and each column has unique name.

The following figure presents a simple relational database comprising of two tables: one shows bank customer and the other shows the a/c that belongs to those customers.

Customer Name	SNO	Street	City	A/C No
Johnson	37539	Alma	Banker	S123
Modena	51970	Street2	Purlin	T535
Shyam	37539	Street3	Charta	S123

A/C No	Balance
T535	72,364,62
R137	8,38,18,84
S123	8,281,41

The following are the different characteristic of relational data model—

(1) **Tuples:** Each row of data is known as tuple.
(2) **Attributes:** Each column in a relation is called as attribute.
(3) **Relation:** In relational data model table is considered as relation.
(4) **Cardinality:** It is defined as the no of tables in a relation.
(5) **Degree of relation:** The no of attributes in a relation is defined as its degree.
(6) **Domain:** It is defined as set of all possible values that an attribute may validly contain.
(7) **Body of relation:** It consist of an unordered set of zero or more tuples.

Advantages:

i. **Simple:** It is a simpler model as it frees the designers from the actual physical data storing details. This allows them to concentrate on the logical views of the database.

ii. **No Anomalies:** This model does not suffer from insert, delete, update anomalies. The retrieval operation is very simple and symmetric.
iii. **Structural Independent:** The relational data model does not depend on the navigational data access system. So, changes in the database structure do not affect the data access.
iv. **Easier design, implementation, maintenance and uses:** Both data as well as structural independence are provided by the relational model which makes database design maintenance administration and uses much easier than other model.
v. **Better query capabilities:** The relational database model provides vary powerful, flexible and easy-to-use query facilities.

Disadvantages:

i. The relational data models need more powerful computing hardware and storage devices to perform RDMS-assigned tasks.
ii. The easy way to design relational data model makes it a poor system in terms of efficiency.
iii. SQL does not provide an efficient way to browse alphabetically through an index. Therefore, there may be some difficulties in accessing data.

ii) **Hieratical Model:** It is one of the oldest database models. One of first hieratical database's Information Management System (IMS) way developed by IBM.

In a hieratical model data are viewed as a collection of relation (also called segments).

A hieratical relation is defined as a collection of relation that are connected together as a hieratical tree by logical associations. Each segments contain multiple instances. Hence,even if the segments are connected as a chain by logical associations, the instance structure can be a structure with multiple branches. The segments pointed to by the logical association is referred to as the child segments and the other segments is called as the parent segment.

For example, consider the two segments as the customer information (customer_name, ssn, customer_state, customer_city, ac_no) and (ac_no, balance). Now, the hieratical model for these two segments is as follows—

Advantages:

i. **Simplicity:** Since database is based on hieratical structure relationship between various layer is logically simple.
ii. **Data Sharing:** Because all the data are held in a common database, data sharing becomes practical.
iii. **Data Security:** Hieratical model was the 1st database model that offered the data security that is provided and enforced by the DBMS.
iv. **Data Integrity:** Because this model is based parent-child relationship, there is always a link between parent segments and the segments under it. The child segments are always automatically referred by its aren't, so this model provides data integrity.
v. **Efficiency:** This model is very affective. When a database contains large volume in one-to-many relationship when the users required large no of transaction.

Disadvantages:

i. **Implementation Complexity:** Although this model is conceptually simple and easy, it is quit complex to implement. The database designer should have a very good knowledge of the physical data storage characteristics.
ii. **Lack of flexibility:** Hieratical database is less flexible. The changes in the new relations or segments often yield very complex system management task. A deletion one segment may **leis** to the involuntary deletion of all the segments under it. Such an error could be very costly.
iii. **DBMS Problems:** If we make any changes in the database structure of hieratical database, then we need to make the necessary changes in all the application programs that access the database. Thus maintaining the database and application programs become very difficult.
iv. **Operational Anomaly:** Hieratical model suffers from the insert, update and delete anomalies. The relational operation and asymmetric.
v. **Implementation Limitation:** Many of the common relationship do not conform to the one-to-many relationship as required by the hieratical model.

ii) **Network Model:** Data in the network model are represented by collection of records and relationship among data are represented by links which can viewed as pointer. A network model is generalization of hieratical model. In this model, a segment can have multiple parent segment. In general, the segments are grouped as levels but logical associations can exist between segments belonging to any level. Although the logical associations are directional in nature, any two levels can have both types of derationed associations. The diagrammatic representation of different logical associations between the segments resemble graph. In general, these logical associations between instance are implemented by pointers in the data set on by ereating connecting relations that reflect the many-to-many associations between instances.

In network database terminology, Asa relationship is a set of at least two type of records.

a. An owner record (equivalent to parent)
b. A member record (equivalent to child)

The following figure shows the structure of a network data model—

FIG

Advantages:

i. **Simplicity:** Network model is simple and easy to design.
ii. **Capable of handling more relationship types:** The network model handles, one-to-one, one-to-many and many-to-many relationships which helps in modeling the real life application.
iii. **Easy to access:** The data access in this model is easier and superior. An application can access an owner record and all the member records within a set.
iv. **Data integrity:** The network model does not allow a member to exists without an owner. Thus a user must define the owner record then the member record. This ensure data integrity.
v. **Data Independence:** The network model is batter then the hieratical model in isolating the program from the complex physical storage details.
vi. **Database Standers:** The hieratical model does not have any universal standard for the database design and modeling, but the network model is based on the standards formatted by DBTG.

Disadvantages:

i. **System Complexity:** As all the records are maintained using pointers so the whole database structure becomes very complex.
ii. **Operational Anomalies:** Since network model uses pointer for relation so its implementation becomes quite complex.
iii. **Not user friendly:** The network model is not a design for a user friendly system as it is a highly skill oriented system.
iv. **Abuse of structural independence:** Since the data access method in the network database model is navigational system, making structural changes and impossible in some cases. If changers are made to the database storage, then all application program need to be modified they can access data.

3) Physical Data Model: Physical data models are used to describe data at the lowest level. In contrast to logical data models, there are few physical data models is used to if the widely known once are—

i. Unifying Model.
ii. Frame-Memory model.

9. Compare Hieratical, Network & Relational Data models.

No.	Hieratical	Network	Relational
i.	Relationship between record is of parent-child type.	Relationship between records is expressed in the form of pointer links.	Relationship between record is represented by a relation that contains a key for each record involved in the relationship.
ii.	many-to-many relationship can't be expressed in this model.	many-to-many relationship can be implemented.	many-to-many relationship can be easily implemented.
iii.	It is simple, straight forward and natural method of implementing record relationship.	Record relationship implementation is very complex due to the use of pointers.	Record relationship implementation is very easy through the use of a key or composite key field.
iv.	This type of model is useful only when there is some hieratical characteristics in the database.	Network model is useful for representing such record which have many-to-many relationship.	Relationship model is useful for representing most of the real world object and relationship among then.
v.	In order to represent links among records pointers are used. Thus relationship among records are physical.	In network model the record relations are also physical.	Relation does not maintain physical connection among record. Data is organized logically I the forms of rows columns and stored in table.
vi.	Searching for record is very difficult, since one can relate a child only after going through its parent.	Searching a record is easy, since there are multiple access pass to data element.	A unique index key field is used to search for a data element.