Homework 1

Question 1: 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.

- One would choose to use a database systems to take advantage of the characteristics of database systems, such as the use of abstraction, the ability to support multiple users and control redundancy and duplication of effort, the self describing nature of databases, and the support of multiple different views for different user groups. It would make sense not to use a database system when the application is simple and you wouldn't expect the data to change, the system does not need to support multiple users, in systems that need to respond in real time and the query speed of a database approach is not fast enough, and in embedded systems that do not have to storage for large database systems. Database systems also have built in backup and recovery mechanisms.

Question 2: The implementation of a database approach has its own challenges, such as expense. Discuss the various costs associated with the implementation of a database approach.

- Some of the various overhead costs include high initial investment in hardware, software, and training, the generality that a DBMS provides for defining a processing data, and overhead for providing security, concurrency control, recovery, and integrity. The last stated increases time costs, and can make querying and updating databases slow. The cost of specialized personnel may also be considered an extra cost associated with the implementation of a database.

Question 3: What are the responsibilities of a DBA? If we assume that the DBA is never interested in running his or her own queries, does the DBA still need to understand query optimization? Why?

- The database administer is responsible for the administration and security of a database, authorizing access to the database, and coordinating and monitoring use of the database system. The database administrator maintains the database and makes sure that it is running efficiently for the users. This means the Database administrator should understand things like query optimization so that he may fully understand how to maintain the database system so that it is running at the required level of service.

Question 4: Explain the role(s) (if any) of each of the following plays in representing information about the real world in a database?

- The data definition language is used to describe the structure and data elements of files that contains records. These records represent real-world objects. The data elements or attributes describe the characteristics of real-world objects.
- The data manipulation language is used to change or update attributes in a record. The database must always reflect the world it is based on. When real world changes occur, the database must be updated to reflect those changes.
- The buffer manager does not have a role in representing information about the real world.
- The data model is a collection of concepts used to describe the structure of a database. The data model determines *how* we represent the real world objects and their relationships.

Question 5: Compare and contrast the two-tier client-server architecture. Why is the latter architecture more appropriate for the web?

- In a two-tier client-server architecture, a server module handles the storage of data, access, and search while multiple users send queries and updates from client modules. In a three-tier system, an extra module acts as an intermediary between the clients and the server module housing the database. The intermediary machine can add an extra layer of security, by authenticating users before sending queries to the server module and is therefore more appropriate for the web. It keeps users from being able to query the server module directly and also adds better load balancing therefore being more appropriate for

Question 6: Define the term "database integrity." How does database security differ from database integrity?

- Database security refers to protecting data from malicious actors seeking to gain personal information or cause disruptions in service. Database integrity refers to how the database state actively represents the current world and data being accurate and consistent. It means that users are able to trust the information currently in the database. This is done through atomic transactions and concurrency control.

Question 7: Describe the features of a Distributed Database Management System or DDBMS?

- A DDBMS can have the actual database and DBMS software distributed over many sites connected by a computer network. The data is often replicated on multiple sites so that failure of a site will not make some data unavailable. Homogeneous DDBMSs use the same DBMS software at all the sites, while heterogeneous DDBMSs can use different DBMS software at each site DDBMSs use a client-server architecture. A middleware software can be used to access several preexisting databases stored under heterogeneous DBMSs.