# Database Management Systems
## *1st Assignment*

Muhammad Karbalaee

November 12, 2022

## 1 Define data. What are the different types of data?

Data is every sort of fact about events, people, objects and generally speaking, raw information is called data. Data should be storable
There are three types of data.

- Structured

- Unstructured

- Semi-structured

**Structured**
Structured data is highly organized and has a model. Structured data is categorized as quantitative data. SQL is used to manage structured data.
**Unstructured**
Unstructured data is a sort of data that is not preset by a data model or schema. NoSQL databases are used to manage this category of data. Unstructured data is categorized as qualitative data.
**Semi-structured**
Semi-structured data is like a bridge between structured data and unstructured data. It's more complex in comparison to structured data and is easier to store than unstructured data. Examples of semi-structured data is JSON, CSV and XML.

## 2 Define database. What are the advantages of using a database system?

A database system is a repository that stores and manages data files and records.
Generally there are two methods of keeping data records, file systems or databases.
File systems store data records directly into files and reading, writing and updating data must be performed manually by computer programs that programmers create.
For example; To set a constraint on the data type that the user of the program can input for the age field, the programmer of that software should manually check the input data type and apply the restrictions.
On the other hand, database can satisfy this need easily since databases are basically already-created softwares that have implemented simple-to-use APIs that make the life of programmers easier.
Furthermore, databases can reduce redundancy, increase security, maintain integrity, concurrent access, and avoid inconsistencies.
In terms of redundancy, databases manage data records carefully to avoid data records being stored twice unnecessarily.
They insure security by applying access levels which simply means defining which users can access which data and what they can do with it.
Integrity is guaranteed by databases through predefined constraints. For example; the database designer can set constraints that an student cannot take more than 20 credits in each semester, this makes the software more reliable and integrated.
Sometimes two or more users want to access on data record at the same time. This is known as concurrent access. Databases help developers with that too since they handle all these by themselves.
These features make databases the preferred method to store data rather than file systems.

## 3 What is a transaction?

Transaction is defined as a set of consecutive operations, operations being SELECT, DELETE, UPDATE and so on. Transactions are defined to be breaking-safe, Meaning if one of the operations fail, the transaction is not done and is reverted. For example; While you ask the bank to transfer some amount of money to another account, if any of the

operations in the middle of this transaction fail, the transaction is reverted and your money doesn't get lost.

In other words, database transactions take place fully or not at all. While a transaction is happening, changes to database are unstable meaning they are not finalized, but once the transaction is complete, the changes are committed and finalized. This feature is so helpful for reverting changes in case the transaction fails.

# 4 What are locks in a DBMS? Briefly explain them.

Locks in DBMS are a method of ensuring data isolation in a transaction that requires mutually exclusive data access. This means that by locking certain data in database, Access for that data while another transaction is performed is banned. There are two types of lock.

- Shared lock

- Exclusive lock

Shared lock let's other transactions to read data while another transaction is taking place, but writing is not allowed. In the other hand, exclusive lock allows neither reading nor writing operations on the locked data.

# 5 What are ACID Properties? Explain Atomicity and Consistency.

ACID stands for atomicity, consistency, isolation and durability. ACID properties are a set of features or principles that every transaction should have to ensure reliability of the transaction which maintains the database integrity.

Atomicity means that either all the operation in a transaction should take place successfully or not at all. This simply means, if one of the operations within a transaction fails, the whole transaction is failed. There is no such thing as partially correct transaction.

Consistency means that after the transaction is committed and performed completely, all sources of data should have the latest data. For example, if a money transfer transaction is debited, both the sender and receiver of the money should have the results in their database.

# 6 What is meaning of "data in database is integrated".

This means that overall, data stored in the database is complete, consistent and correct. An integrated database is one that users can rely on the data which they get from. For example; if they query a data, they are sure that what the database returns to them is reliable and correct and they can use the data without performing additional manual checks themselves.

# 7 What is logical data independence and why is it important?

In three-schema database design, logical data independence in defined as the independence between the conceptual level and external level of the database, meaning, changes in the conceptual level do not effect the external level of the database. This is pretty important since it helps lower-cost maintenance since enhancing the logic inside database doesn't require changing the program using the database. Moreover, It helps security.

# 8 Name the main steps in database design. Explain two of them.

There are six steps in database design:

- Requirement analysis

- Conceptual design

- logical design

- Schema refinement

- Physical design

- Application and security design

The first step in database design is requirement analysis which means thinking of what data is going to be stored in the database and what application is going to be built on top of the database and what it is going to do. Also, performance requirements are considered in this step.

Conceptual design step is performed using the data gathered in the requirement analysis step to create a high-level description of the data to be stored in

the database, as well as the constraints which this
data should follow.