**DATABASE AND RELATED OPERATIONS**

**Flow of presentation :**

●What is database?

●Why it’s needed

● Different types of databases

●CRUD operation

● Indexing

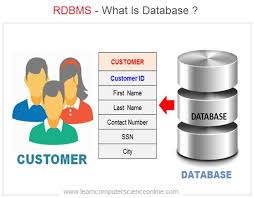
● Aggregation

● ACID properties and

CAP theorem

**What is database?**

A database is a collection of data, usually stored in electronic form. A database is typically designed so that it is easy to store and access information.



**Why database is needed?**

* A database stores and manages a large amount of data on a daily basis.
* This would not be possible using any other tool such as a spreadsheet as they would simply not work.
* A database is pretty accurate as it has all sorts of build in constraints, checks etc.
* In a database, it is easy to update data using various Data Manipulation languages (DML) available. One of these languages is SQL.
* This is ensured in databases by using various constraints for data.
* It is very easy to access and research data in a database.

**Different types of database**

|  |  |
| --- | --- |
| **RDBMS** | **MongoDM** |
| RDBMS is a relational database management system and works on relational database. | MongoDB is a non-relational, document oriented database management system and works on document based database. |
| Performance increases with increase of RAM. | Its performance increases with addition of processor. |
| Follows ACID principle, Atomicity, Isolation, and Durability. Follows CAP theorem, Consistency, | Follows CAP theorem, Consistency, Availability, and Partition tolerance. |
| Database uses Row and column | Database uses Document and Field. |
| RDBMS uses SQL to query database. | MongoDB uses BSON to query database |

**CRUD Operations On RDBMS**

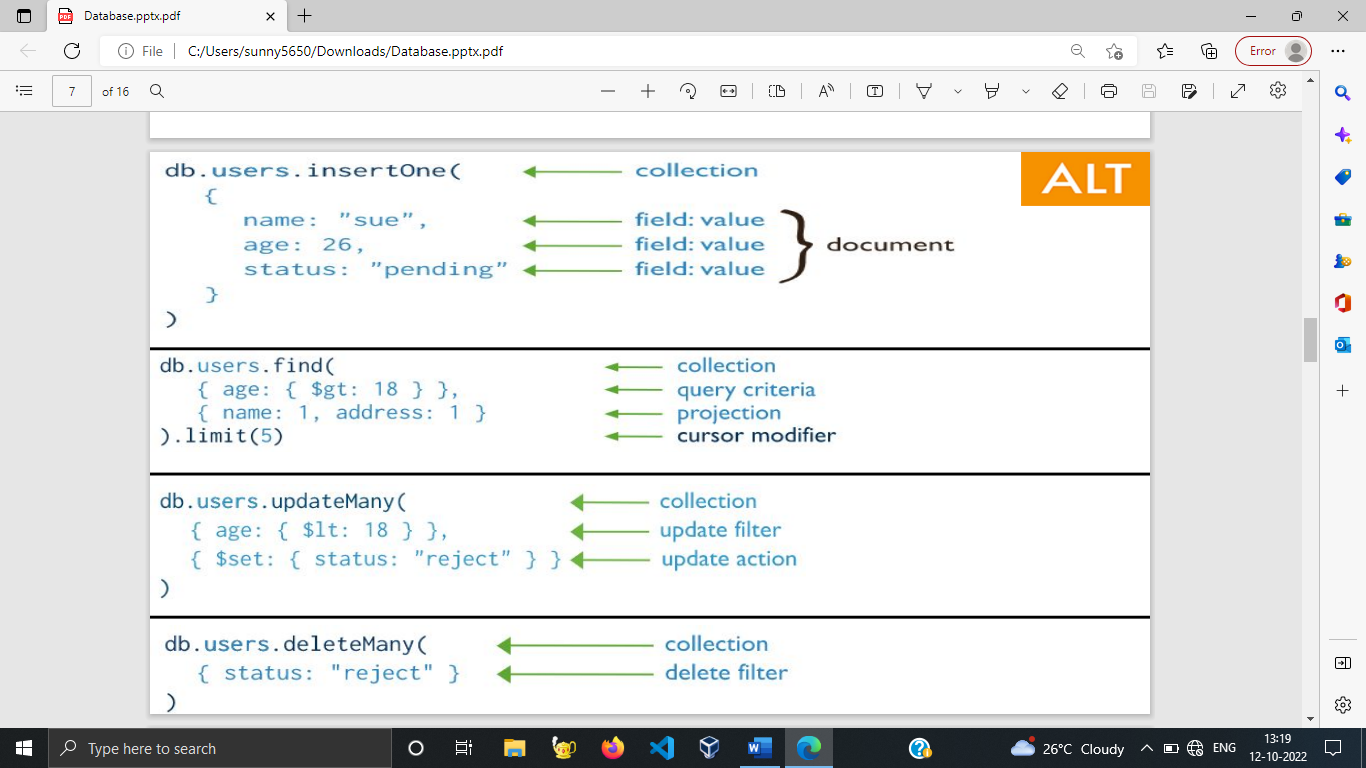
* Create -> Insert -> insert into TABLE\_NAME columns(1,2..) values(1,2..);
* Read -> select -> select \* from TABLE\_NAME;
* Update -> update -> update TABLE\_NAME set COL\_NAME='x' where condition;
* Delete -> delete -> delete from TABLE\_NAME where condition;

**CRUD Operations On MongoDBMS**

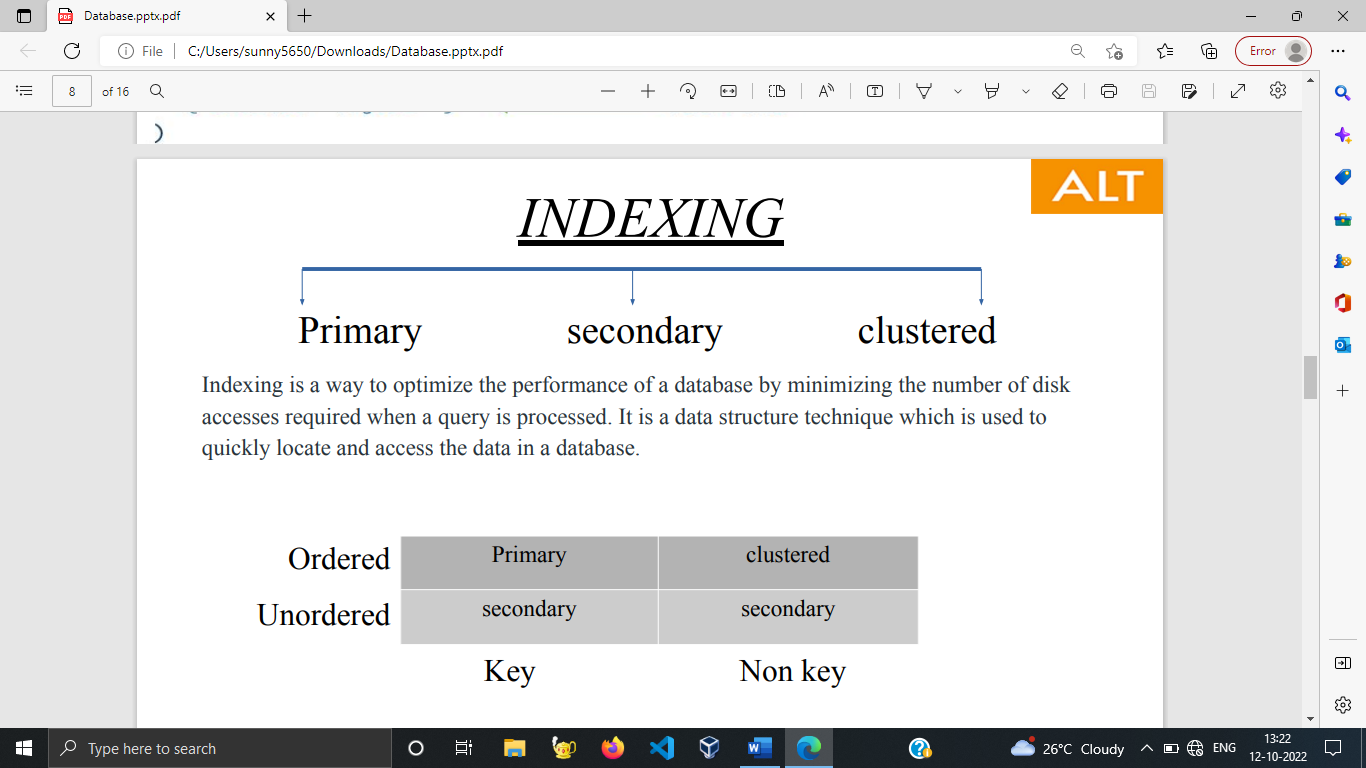
* Create -> db.collection.insertOne()

db.collection.insertMany()

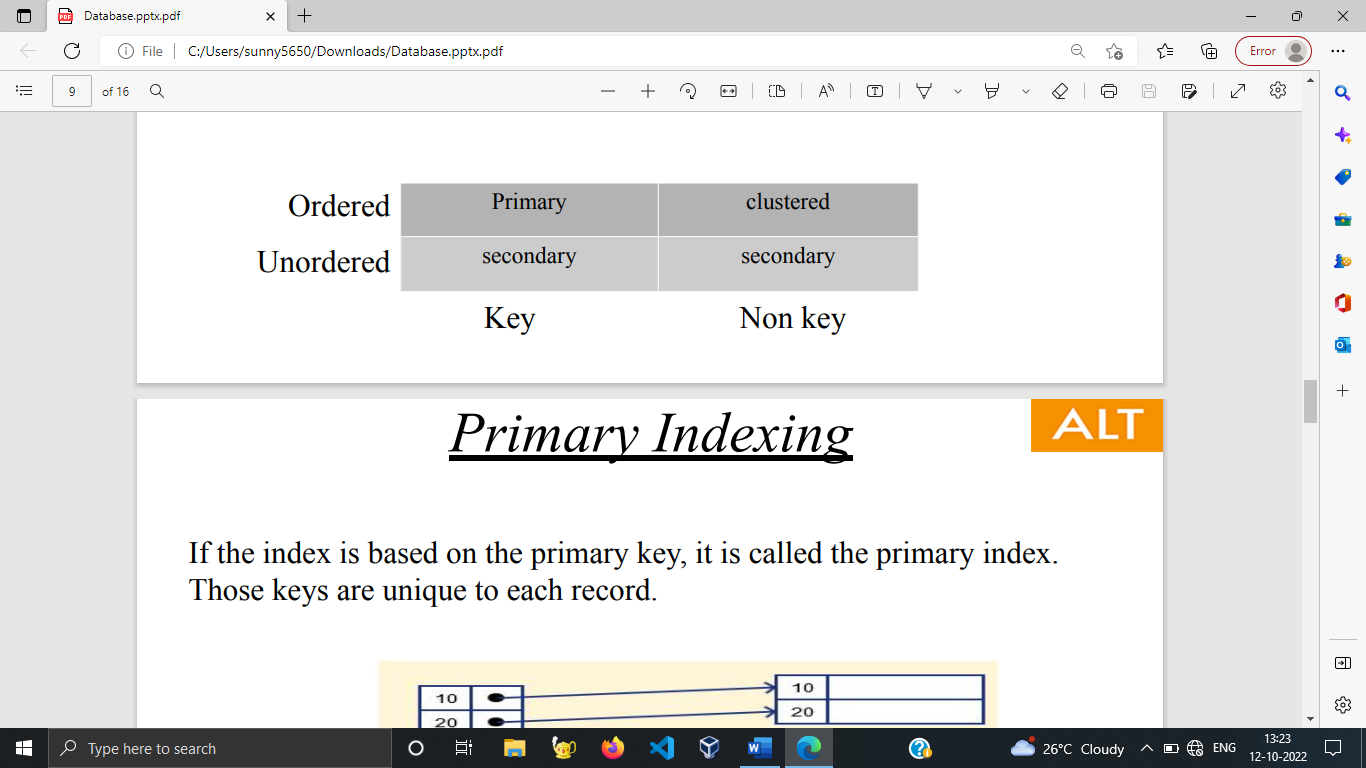
* Read -> db.collection.find()
* Update -> db.collection.updateOne() db.collection.updateMany()
* Delete -> db.collection.deleteOne() db.collection.deleteMany()



**INDEXING**

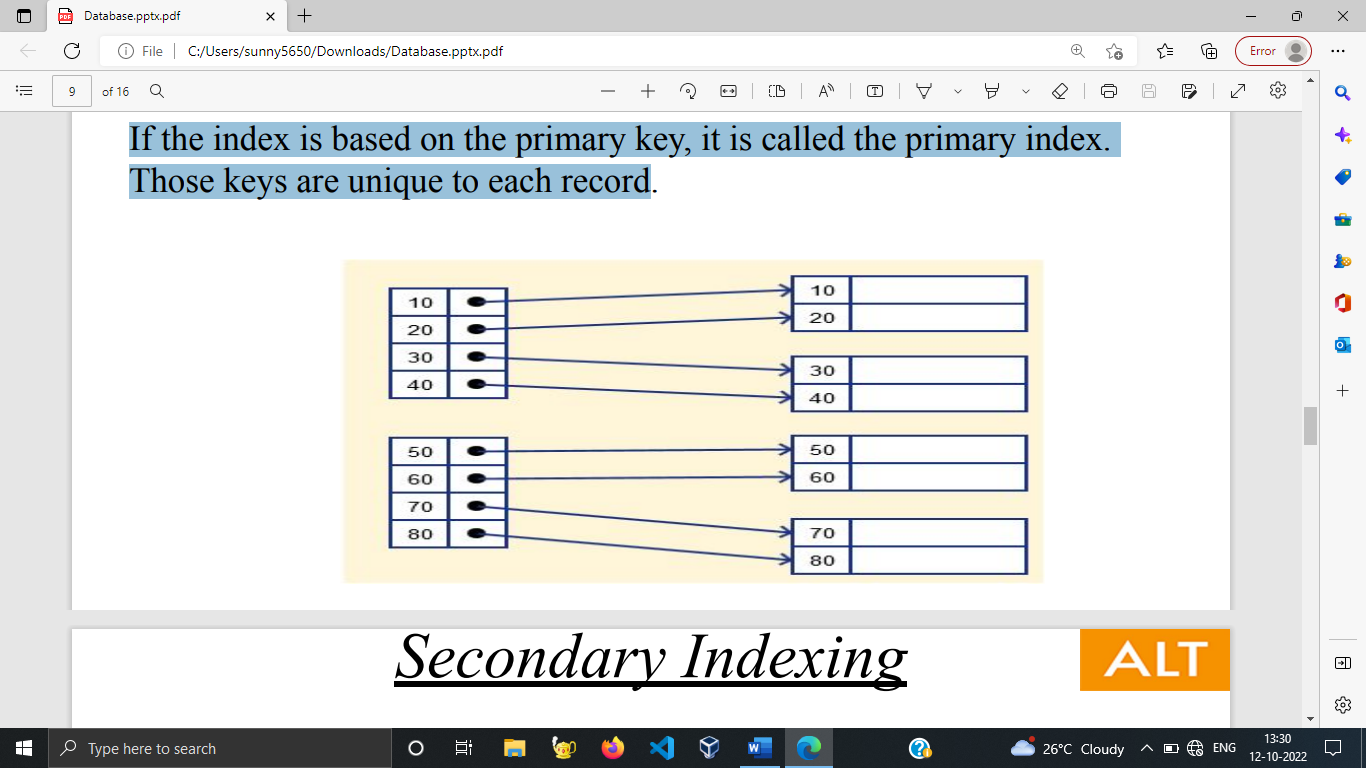


Indexing is a way to optimize the performance of a database by minimizing the number of disk accesses required when a query is processed. It is a data structure technique which is used to quickly locate and access the data in a database.



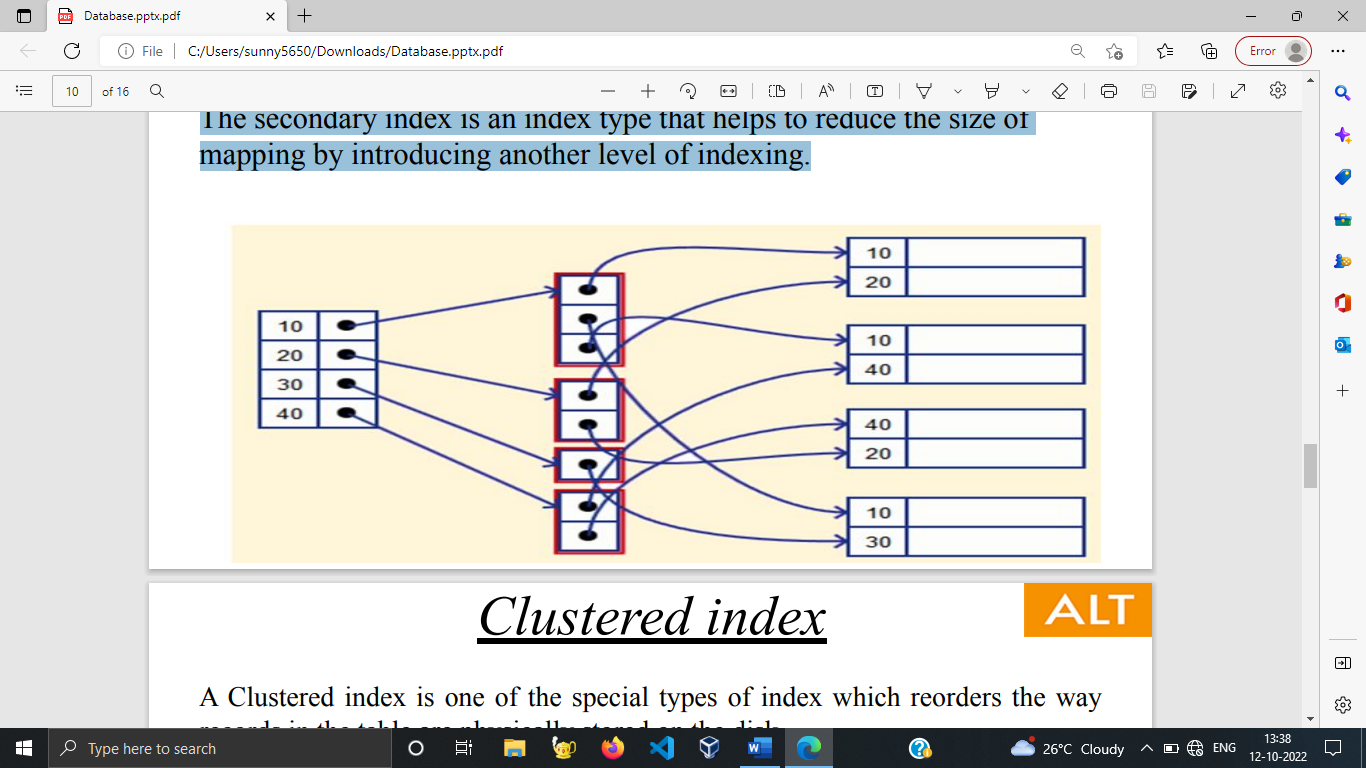
**PRIMARY INDEXING**

If the index is based on the primary key, it is called the primary index. Those keys are unique to each record

****

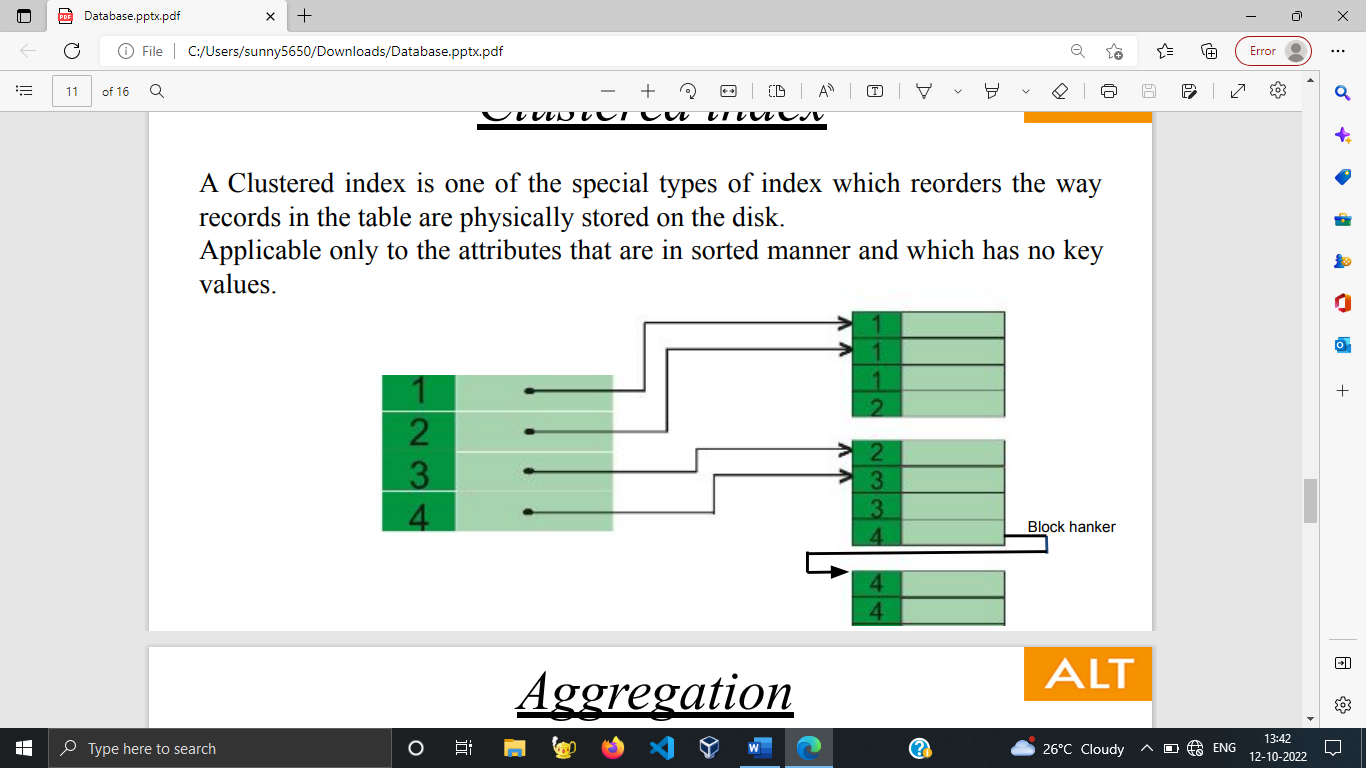
**SECONDARY INDEXING**

The secondary index is an index type that helps to reduce the size of mapping by introducing another level of indexing.

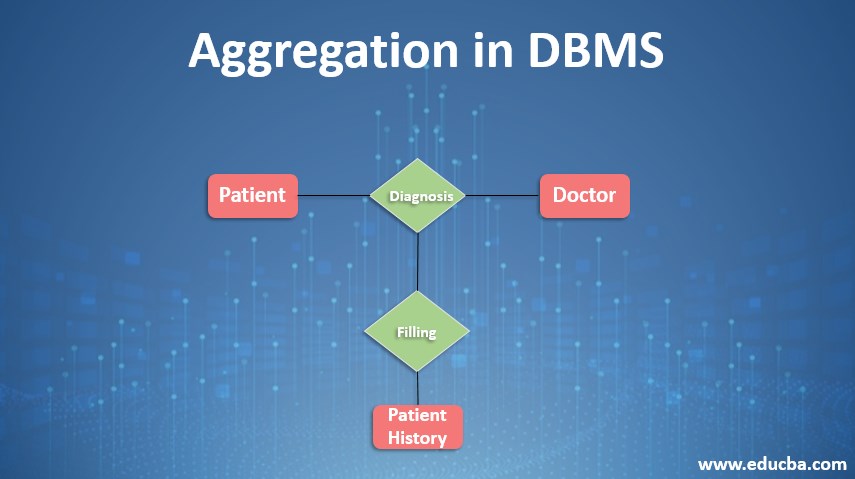
****

**CLUSTERED INDEXING**

A Clustered index is one of the special types of index which reorders the way records in the table are physically stored on the disk. Applicable only to the attributes that are in sorted manner and which has no key values.

****

**AGGREGATION**

****A Clustered index is one of the special types of index which reorders the way records in the table are physically stored on the disk. Applicable only to the attributes that are in sorted manner and which has no key values.

Aggregation Functions :

● Min

● Max

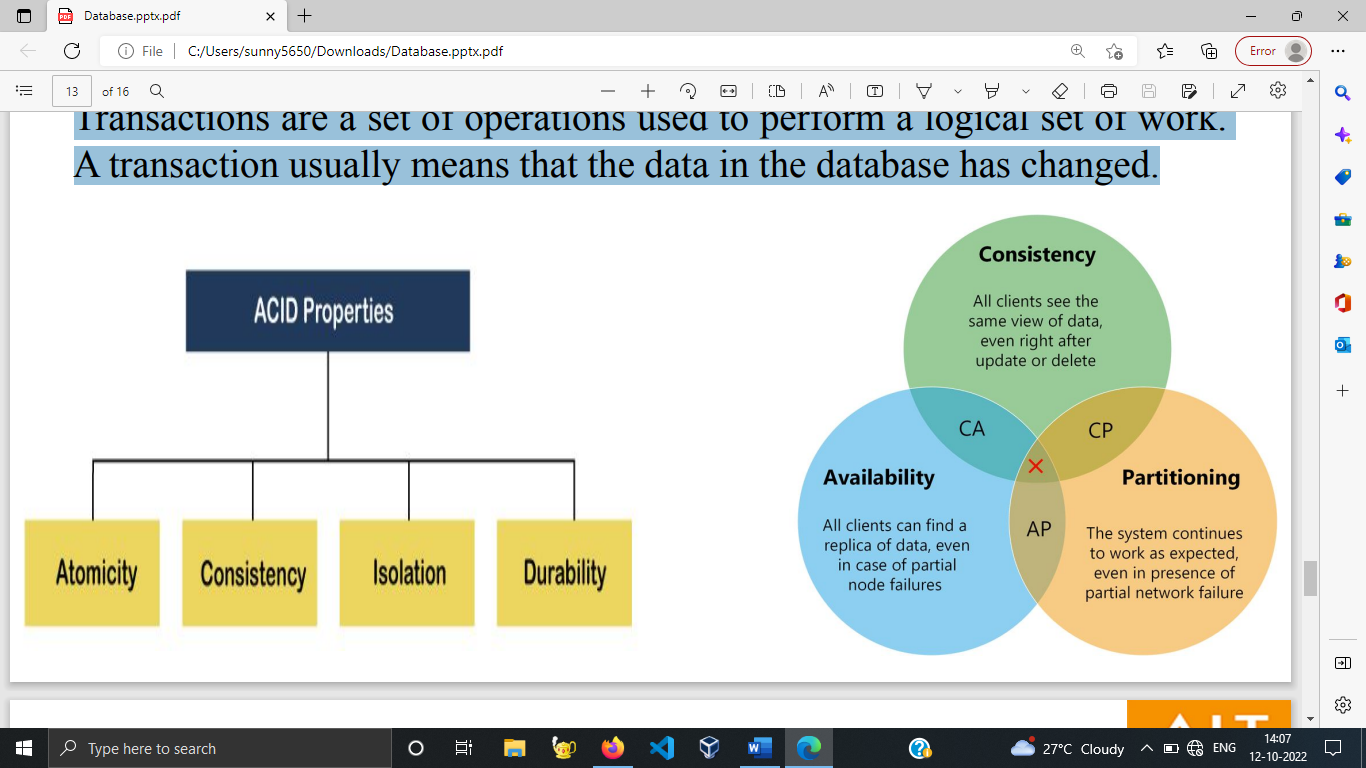
● Sum

● Avg

● Count

**ACID PROPERTES AND CAP THEOREM**

Transactions are a set of operations used to perform a logical set of work. A transaction usually means that the data in the database has changed.

****

**Atomicity :**

The term atomicity defines that the data remains atomic. It means if any operation is performed on the data, either it should be performed or executed completely or should not be executed at all.

**Consistency :**

The word consistency means that the value should remain preserved always. In DBMS, the integrity of the data should be maintained, which means if a change in the database is made, it should remain preserved always.

**Isolation :**

The term 'isolation' means separation. In DBMS, Isolation is the property of a database where no data should affect the other one and may occur concurrently.

**Durability :**

Durability ensures the permanency of something. In DBMS, the term durability ensures that the data after the successful execution of the operation becomes permanent in the database.

**Consistency :**

Means that the nodes will have the same copies of a replicated data item visible for various transactions. A guarantee that every node in a distributed cluster returns the same, most recent and a successful write.

**Availability :**

Each read or write request for a data item will either be processed successfully or will receive a message that the operation cannot be completed.

**Partition Tolerance :**

It’s a guarantee that the system continues to operate despite arbitrary message loss or failure of part of the system. In other words, even if there is a network outage in the data center and some of the computers are unreachable, still the system continues to perform.

