**Indexes**

➔ Two main types:

➔ B-tree indexes: use a tree-like data structure that maintains data sorted and allow for search, order, range search in log time.

➔ Hash indexes: use a hash-function to map keys to values; are only considered when an equality operator is used (no sorting or ranges).

**Clustering**

➔ Physical re-ordering of data in disk based on the index information.

➔ To cluster a table, an index must already be defined.

➔ Clustering is a one-time operation: when the table is subsequently updated, the changes are not clustered.

➔ Clustering will help when multiple records are read together and an index can group them.

**Cardinality**

➔ The uniqueness of data values contained in a particular column.

➔ The lower the cardinality, the more duplicate values in the column.

**Transactions**

➔ A transactions is a set of database operations that is considered as a single unit.

➔ A transaction either succeeds or fails in its entirety.

➔ Concurrency Problems

➔ **Dirty reads:** transaction reads data written by uncommitted transaction.

➔ **Non-repeatable reads:** transaction re-reads data and finds that data has been modified by another transaction.

➔ **Phantom reads:** transaction re-executes query and different rows are being returned.

➔ **Serialization anomaly:**