INDEXES

Lesson objectives

- Explain how data entries are organized in order to support efficient retrieval of data

Definition of index

Indexes are special lookup tables that the database search engine can be used to speed up data retrieval.

Simply put an index is a pointer to a data in a table. An index in a database is very similar to an index in the back of a book. Example if you want to reference all pages in a book that discuss a certain topic, you first list the topics alphabetically and then referred to one or more numbers.

Indexes are automatically created when PRIMARY KEY and UNIQUE constraints are defined on table columns. For example, when you create a table and identify a particular column to be the primary key, the database engine automatically creates CONSTRAINTS and index on that column.

The concept of an index is to make it easier for the DBMS to relate data to an index to provide fast access to the rows in the table. Example a table called customers with the following columns name, address, city, state, phone number. The index could be applied to the column. This would result in an ordered list of keys (the column name) along with the locations of the associated rows in the customers table.

Index classification

1. Clustered vs. none clustered.
2. Dense vs sparse index
3. Primary and secondary index

- **Clustered vs. none clustered.**

Clustered: Clustered index sort and store the data rows in the table or view based on their key values. These are the columns included in the index definition. There can only be clustered index per table.

<u>Non clustered:</u> non clustered index contains the non-clustered index key values and each key value entry has a pointer to the data row that contains the key value.

- **Dense Vs. Sparse index**

<u>Dense index</u> a dense index in databases is a file with pairs of keys and pointers for every record in the data file. Every key is associated with a particular pointer to a record in the sorted data file.

<u>Sparse index</u>   A sparse index in database is a file with pairs of keys and pointers for every block of data file. Every key in the file is associated with a particular pointer to the block in the sorted data file.

- **Primary and secondary index**

  Primary index: the primary index contains the key fields of the table and a pointer to the non-key fields of the table. The primary index is created automatically when the table is created in the database.

  Secondary index: used to index field that are neither ordering fields nor key fields (there is no assurance that the file is organized on key field or primary key field). As an example of how secondary indexes might be used, consider a database containing a list of students at a college; each of whom has a unique students ID number.