

Ordered Indices

What is database index:

- A database index is a data structure that improves the speed of data retrieval operations on a database table at the cost of additional writes and storage space to maintain the index data structure.
- Indexing is a way of sorting a number of records on multiple fields.
- Indexing is a way to optimize performance of a database by minimizing the number of disk accesses required when a query is processed.

Uses of indices :

- Index is used to speeds up data retrieval from a database.
- Adding index to database increases the read performance.
- Using an index we can access a record in a file very quickly.
- Indexes can be created or dropped with no effect on the data.

How index works:

- If we run any query in database first database will check to see if there is an index on the column(s) being queried.
- If the index exists in table then index is used to fetch a row, the index is walked until it finds the row(s) of interest, and the base table is then looked up to fetch the actual row data otherwise it will do complete table scan from first row of table to end.
- When a row is inserted, a corresponding row is written to the index, and when a row is deleted, its index row is taken out.

Types of Indices :

A table or view can contain the following types of indexes:

- Primary Index
 - Dense Index
 - Sparse Index
- Secondary Indices

Primary Index:

- A primary index is an index whose search key also defines the sequential order of the file.
- It is also Called as a **Clustered index**.
- Clustered indexes 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 be only one clustered index per table, because the data rows themselves can be stored in only one order
- The only time the data rows in a table are stored in sorted order is when the table contains a clustered index. When a table has a clustered index, the table is called a clustered table. If a table has no clustered index, its data rows are stored in an unordered structure called a heap.
- The search key of a clustering index is often the primary key, although that is not necessarily.
- This helps improving the performance of queries using *Order By* and *Group By*, since the data rows are already sorted in the order of the indexes.

Secondary Index:

- Indices whose search key specifies an order different from the sequential order of the file are called Secondary indices.
- Also called a nonclustering indices.
- Nonclustered indexes have a structure separate from the data rows. A nonclustered index contains the nonclustered index key values and each key value entry has a pointer to the data row that contains the key value.
- The pointer from an index row in a nonclustered index to a data row is called a row locator. The structure of the row locator depends on whether the data pages are stored in a heap or a clustered table. For a heap, a row locator is a pointer to the row. For a clustered table, the row locator is the clustered index key.
- In non clustered index index record points to a bucket that contains pointers to all the actual records with that particular search-key value.
- This type of index does not alter and/or prescribe any sort of ordering of rows in the table.

Data Set:

In this project we are using sales schema in adventure works (Sample data provided by Microsoft) database 2017 in SQL server.

The Sales schema is as follows :



