

ষষ্ঠ অধ্যায় পাঠ-৬: ডেটাবেজ সটিং এবং ইনডেক্সিং।

এই পাঠ শেষে যা যা শিখতে পারবে-

- ১। ডেটাবেজ সটিং ব্যাখ্যা করতে পারবে।
- ২। ডেটাবেজ ইনডেক্সিং ব্যাখ্যা করতে পারবে।
- ৩। ইনডেক্সিং এর সুবিধা এবং অসুবিধা বর্ণনা করতে পারবে।
- ৪। সটিং এবং ইনডেক্সিং এর মধ্যে পার্থক্য করতে পারবে।

ডেটাবেজ সর্টিং: এক বা একাধিক ফিল্ড এর মানের উপর ভিত্তি করে ডেটাবেজের রেকর্ডগুলোকে উর্ধ্বক্রম বা নিম্নক্রমে সাজানোর প্রক্রিয়া হচ্ছে সর্টিং। একটি ডেটা টেবিল সর্টিং করার ফলে নতুন একটি টেবিল তৈরি হয় যেখানে টেবিলের রেকর্ডগুলো সর্টিং করা থাকে। একটি টেবিল সর্টিং করার পর টেবিলে নতুন কোন ডেটা ইনপুট দিলে সর্টেড টেবিলটির কোন পরিবর্তন হয় না। এই ক্ষেত্রে পুনরায় সর্টিং করতে হয়। এছাড়া কোন টেবিলের ডেটা সর্টিং করলে যেহেতু টেবিলের নতুন একটি কপি তৈরি হয় তাই অতিরিক্ত মেমোরির প্রয়োজন হয়। বাস্তবক্ষেত্রে সর্টিং এর চেয়ে ইনডেক্সিং বেশি ব্যবহৃত হয়।

সর্টিং দু-প্রকার হয়ে থাকে-

১। উচ্চক্রম/উর্ধ্বক্রম/Ascending order

২। নিম্নক্রম/অবরোহী/Descending order

টেবিলের রেকর্ডগুলো GPA ফিল্ডের এর ভিত্তিতে নিম্নক্রমে সর্টিং করা হয়েছে-

student_info

Roll	Name	GPA
1	Ashek	3.50
2	Amirul	5.00
3	Sajid	4.50
4	Sondha	5.00
5	Shanta	4.90

Before sorting

Roll	Name	GPA
2	Amirul	5.00
4	Sondha	5.00
5	Shanta	4.90
3	Sajid	4.50
1	Ashek	3.50

After sorting

ডেটাবেজ সটিং এর উদ্দেশ্যঃ

১। ডেটাবেজ সটিং এর উদ্দেশ্য হলো কোন ডেটাবেজ টেবিল থেকে কুয়েরির মাধ্যমে প্রাপ্ত আউটপুট ডেটাকে সাজিয়ে উপস্থাপন করা।

টেবিল থেকে ডেটা শর্তসাপেক্ষে সিলেক্ট করে কোন এক বা একাধিক ফিল্ডের উপর ভিত্তি করে রেকর্ড গুলোকে উর্ধক্রম অনুসারে সাজিয়ে দেখানোর **SQL** কমান্ডের সিনট্যাক্সঃ

```
SELECT */ column_name_1,column_name_2,.. FROM table_name  
WHERE condition ORDER BY column_name ASC;
```

student_info

Id	Name	Section	GPA	City
2	Ashek	A	5.00	Dhaka
3	Amirul	B	4.00	Jessore
1	Shanta	A	5.00	Dhaka
5	Sondha	C	4.50	Jamalpur
4	Sajid	A	5.00	Khulna

student_info নামে উপরের টেবিলের ডেটা সিলেক্ট করে **Id** ফিল্ডের উপর ভিত্তি করে রেকর্ড গুলোকে উর্ধক্রম অনুসারে সাজিয়ে দেখানোর **SQL** কমান্ড:

```
SELECT * FROM student_info ORDER BY Id ASC;
```

SQL কমান্ডটি রান করলে আউটপুট টেবিলটি দেখতে নিম্নরূপ হবে-

Id	Name	Section	GPA	City
1	Shanta	A	5.00	Dhaka
2	Ashek	A	5.00	Dhaka
3	Amirul	B	4.00	Jessore
4	Sajid	A	5.00	Khulna
5	Sondha	C	4.50	Jamalpur

টেবিল থেকে ডেটা সিলেক্ট করে অধঃক্রম অনুসারে সাজিয়ে দেখানোর **SQL** কমান্ডের সিনট্যাক্সঃ

```
SELECT */ column_name_1,column_name_2,.. FROM table_name
ORDER BY column_name DESC;
```

student_info

Id	Name	Section	GPA	City
2	Ashek	A	5.00	Dhaka
3	Amirul	B	4.00	Jessore
1	Shanta	A	5.00	Dhaka
5	Sondha	C	4.50	Jamalpur
4	Sajid	A	5.00	Khulna

student_info নামে উপরের টেবিলের ডেটা সিলেক্ট করে **Id** ফিল্ডের উপর ভিত্তি করে রেকর্ড গুলোকে অধঃক্রম অনুসারে সাজিয়ে দেখানোর **SQL** কমান্ড:

```
SELECT * FROM student_info ORDER BY Id DESC;
```

SQL কমান্ডটি রান করলে আউটপুট টেবিলটি দেখতে নিম্নরূপ হবে-

Id	Name	Section	GPA	City
5	Sondha	C	4.50	Jamalpur
4	Sajid	A	5.00	Khulna
3	Amirul	B	4.00	Jessore
2	Ashek	A	5.00	Dhaka
1	Shanta	A	5.00	Dhaka

ইনডেক্সিংঃ ইনডেক্সিং হচ্ছে সুসজ্জিতভাবে বা সুবিন্যস্তভাবে তথ্যাবলির সূচি প্রণয়ন করা। ডেটাবেজ থেকে ব্যবহারকারি কোনো ডেটা যাতে দ্রুত খুঁজে বের করতে পারে সেজন্য ডেটাকে একটি বিশেষ অর্ডারে সাজিয়ে ডেটাগুলোর একটা সূচি প্রণয়ন করা হয়। ডেটাবেজ টেবিলের রেকর্ড সমূহকে এরূপ কোনো লজিক্যাল অর্ডারে সাজিয়ে রাখাকেই ইনডেক্স বলায় ডেটাবেজ টেবিলের এক বা একাধিক ফিল্ডের উপর ইনডেক্স করে **Alphabetically** বা **Numerically** সাজানো যায়। ইনডেক্স ফাইল মূল ডেটাবেজ ফাইলের কোনরূপ পরিবর্তন না করে বিভিন্নভাবে সাজাতে পারে।

যেমনঃ নিচের **Fact** টেবিল থেকে যদি ২ রোল ধারীর **GPA** জানতে চাওয়া হয়। তাহলে কিন্তু একটু বেশি টাইম লাগবে। কারণ রোল ফিল্ডের মান গুলো সাজানো নেই। কিন্তু **Fact** টেবিলের ইনডেক্স থেকে ২ রোল ধারীর **GPA** খুব সহজেই পাওয়া যায়। কারণ রোল ফিল্ডের মান গুলো সাজানো আছে।

Fact Table

Roll	Name	Section	GPA
3	Rafid	C	4.50
6	Akash	A	3.50
11	Monir	C	5.00
5	Ridoy	B	4.80
2	Babul	B	3.50
4	Nirob	A	4.00
7	Ashek	B	5.00
9	Obama	C	2.60
8	Fahim	A	3.10
1	Razu	A	5.00
10	Kamal	B	4.00

Index Table

Roll	GPA
1	5.00
2	3.50
3	4.50
4	4.00
5	4.80
6	3.50
7	5.00
8	3.10
9	2.60
10	4.00
11	5.00

ইনডেক্স করার সময় বিবেচ্য বিষয় সমূহঃ

- ১। সাধারণত কী ফিল্ডের উপর ইনডেক্স করতে হয় এবং ইনডেক্স এর একটি নাম দিতে হয়।
- ২। যে ফিল্ডের উপর ভিত্তি করে ইনডেক্স করা হয় সেই নামের অনুরূপ নাম নির্বাচন করতে হয়। এতে ইনডেক্স সমূহ মনে রাখতে সুবিধা হয়।
- ৩। এক বা একাধিক ফিল্ডের উপর ভিত্তি করে ইনডেক্স করা যায়। কোনো ডেটা টেবিলে এক বা একাধিক ইনডেক্স থাকতে পারে বা একই সময়ে খোলা থাকতে পারে। কিন্তু একই সময়ে কেবল একটি ইনডেক্স সক্রিয় থাকবে এবং রেকর্ডসমূহ প্রদর্শনের অর্ডার নিয়ন্ত্রণ করবে।

ইনডেক্সিং এর সুবিধা সমূহঃ

- ১। ইনডেক্সিং এর বড় সুবিধা হলো ইনডেক্স তৈরি করার ফলে ফাইলে সহজে ডেটা খুঁজে বের করা যায়।
- ২। ইনডেক্স তৈরি করার ফলে ডেটাবেজ টেবিলে নতুন কোনো রেকর্ড ইনপুট করা হলেও ইনডেক্স ফাইলগুলো স্বয়ংক্রিয়ভাবে আপডেট হয়ে যায়।
- ৩। ডেটাসমূহের ইনডেক্স তৈরি করলে পারফরমেন্স ভাল পাওয়া যায়।
- ৪। ইনডেক্স ফাইল মূল ডেটাবেজ ফাইলের কোনোরূপ পরিবর্তন না করে বিভিন্নভাবে সাজাতে পারে।

ইনডেক্সিং এর অসুবিধা সমূহঃ

- ১। যদি একাধিক ফিল্ডের উপর ইনডেক্সিং করা থাকে তাহলে কোনো ডেটা এডিট করলে ইনডেক্স ফাইল আপডেট করার জন্য দীর্ঘ সময় অপেক্ষা করতে হয়।
- ২। অনেক রেকর্ডের জন্য ইন্ডেক্স করা হলে অপেক্ষাকৃত বেশি মেমোরির প্রয়োজন হয়। তাছাড়া ইনডেক্স সংরক্ষণের জন্যও কিছু জায়গা লাগে।
- ৩। ডেটা এন্ট্রির ক্ষেত্রে ইনডেক্স ফাইলের রেফারেন্সসমূহ আপডেট হতে বেশ সময়ের প্রয়োজন হয়। সেজন্য ডেটা এন্ট্রি করতেও বেশি সময় লাগে।

সটিং এবং ইনডেক্সিং এর মধ্যে পার্থক্য আলোচনাঃ

টেবিলের ডেটাগুলোর একটি ক্রম তৈরি করতে ইনডেক্সিং এবং সটিং উভয় পদ্ধতিই ব্যবহৃত হয়। ইনডেক্সিং একটি ইনডেক্স ফাইল তৈরি করে যা টেবিলের রেকর্ডগুলোর ফিজিক্যাল অবস্থানের পাশাপাশি সারিগুলোর লজিক্যাল ক্রম ধারণ করে, অপরদিকে একটি টেবিল সটিং করার সাথে সাথে টেবিলটির একটি স্টেড অনুলিপি তৈরি হয়। সাধারণত, ইনডেক্স ফাইলটি একটি স্টেড টেবিল সংরক্ষণের চেয়ে কম মেমোরির প্রয়োজন হয়। ইনডেক্সিং রেকর্ডগুলোর মূল ক্রম পরিবর্তন করে না, যেখানে সটিং রেকর্ডগুলোর মূল ক্রম পরিবর্তন করে। বাস্তবক্ষেত্রে সটিং এর চেয়ে ইনডেক্সিং বেশি ব্যবহৃত হয়।

ইনডেক্স তৈরি করার SQL কমান্ডের সিনট্যাক্স:

একটি টেবিলের উপর ইনডেক্স তৈরি করবে, যেখানে ডুপ্লিকেট ডেটা থাকতে পারেঃ

```
CREATE INDEX index_name  
ON table_name (column1, column2, ...);
```

একটি টেবিলের উপর ইনডেক্স তৈরি করবে, যেখানে ডুপ্লিকেট ডেটা থাকবে নাঃ

```
CREATE UNIQUE INDEX index_name  
ON table_name (column1, column2, ...);
```

ইনডেক্স ডিলিট করার **SQL** কমান্ডের সিনট্যাক্স:

```
DROP INDEX index_name ON table_name;
```

পাঠ মূল্যায়ন-

জ্ঞানমূলক প্রশ্নসমূহঃ

- কা ডেটাবেজ সটিং কী?
- কা ইনডেক্সিং কী?

অনুধাবনমূলক প্রশ্নসমূহঃ

- খা সটিং ও ইনডেক্সিং এক নয়-ব্যাখ্যা কর।
- খা “OLE object এর উপর সটিং সম্ভব নয়”-ব্যাখ্যা কর।
- খা কোন কোন ডেটা টাইপের উপর সটিং করা যায় না? ব্যাখ্যা কর।
- খা ইনডেক্স করা ফাইলে ডেটা এন্ট্রি করা হলে স্বয়ংক্রিয়ভাবে আপডেট হয় —বুঝিয়ে লেখ।
- খা ইনডেক্সিং ডেটাবেজ সিস্টেমের কাজের গতি বৃদ্ধি করে- ব্যাখ্যা কর।
- খা ইনডেক্সিং মূল ফাইলের পরিবর্তন করে না-বুঝিয়ে লেখ।
- খা ইনডেক্সিং এর তুলনায় সটিং এ বেশি মেমরি প্রয়োজন হয় কেন? ব্যাখ্যা কর।

সৃজনশীল প্রশ্নসমূহঃ

উদ্দীপক পড়ে প্রশ্নগুলোর উত্তর দাও-

একটি কলেজের ফলাফলের ডেটাবেজ থেকে একজন শিক্ষার্থীর তথ্য খোঁজার জন্য তিনজন ছাত্রকে নির্দেশ দেওয়া হলো। ১ম ছাত্র শর্ত সাপেক্ষে কমান্ড দিয়ে, ২য় ছাত্র ডেটাবেজের টেবিলের তথ্য সাজিয়ে এবং ৩য় ছাত্র ২য় ছাত্রের চেয়ে দ্রুততর কৌশল প্রয়োগে তথ্য খুঁজে বের করে।

গা তথ্য খোঁজার ক্ষেত্রে ২য় ছাত্রটির কৌশল বর্ণনা কর।

ঘা ১ম ও ৩য় ছাত্রের কৌশল দুটির মধ্যে কোনটি উত্তম? বিশ্লেষণপূর্বক মতামত দাও।

বহুনির্বাচনি প্রশ্নসমূহঃ

১। যদি ডেটাবেজে প্রচুর তথ্য সংরক্ষিত থাকে তবে-

- i. ডেটা দ্রুত খুঁজে পেতে ইন্ডেক্সিং প্রয়োজন
- ii. এমনভাবেই ডেটা দ্রুত খুঁজে পাওয়া যায়
- iii. ডেটা খুঁজে বের করা সময় সাপেক্ষ

নিচের কোনটি সঠিক?

- ক) i ও ii খ) i ও iii গ) ii ও iii ঘ) i, ii ও iii

২। কোন সম্পর্কটি সঠিক?

- ক) কুয়েরিং-বাছাই খ) সটিং-খুঁজা গ) ইন্ডেক্সিং-সাজানো ঘ) সার্চিং-শনাক্ত

৩। সটিং হচ্ছে-

- i. ডেটাকে মানের উর্ধ্বক্রমে সাজানো
- ii. ডেটাকে মানের নিম্নক্রমে সাজানো
- iii. ডেটাকে দৈর্ঘ্যের ভিত্তিতে সাজানো

নিচের কোনটি সঠিক?

- ক) i ও ii খ) i ও iii গ) ii ও iii ঘ) i, ii ও iii

৪। ইন্ডেক্স করা যায়-

- i. একটি ফিল্ডের উপর
- ii. দুইটি ফিল্ডের উপর
- iii. একাধিক ফিল্ডের উপর

নিচের কোনটি সঠিক?

- ক) i ও ii খ) i ও iii গ) ii ও iii ঘ) i, ii ও iii

নিচের উদ্দীপক অনুসারে ৫ নং প্রশ্নের উত্তর দাও:

Roll	Name	Dist
1	Zafar	Dhaka
2	Rana	Comilla
3	Kamal	Rangpur
4	Baki	Chittagong

৫। উদ্দীপক টেবিলের **Dist** ফিল্ডের উপর **A→Z** সর্টিং করলে **Roll** ফিল্ডের মানের ক্রম হবে-

ক) ২, ৩, ৪, ১ খ) ৪, ২, ১, ৩ গ) ৩, ১, ২, ৪ ঘ) ৪, ৩, ২, ১

৬। ডেটাবেজ সাজানো প্রক্রিয়া হলো-

i. সর্টিং ii. ইনডেক্সিং iii. কুয়েরি

নিচের কোনটি সঠিক?

ক) i ও ii খ) i ও iii গ) ii ও iii ঘ) i, ii ও iii