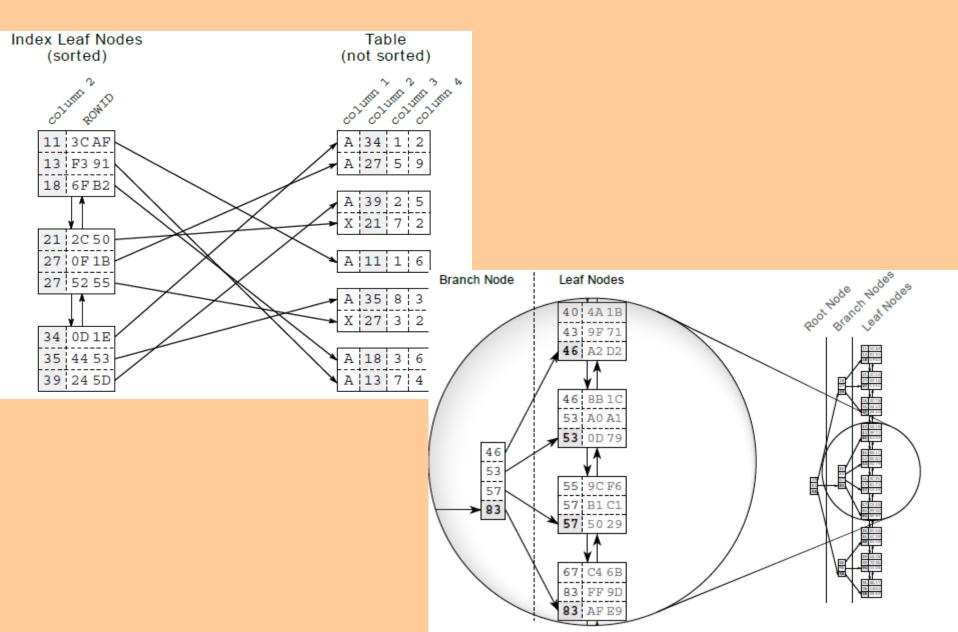
# Оптимизация запросов

## Индексы

• Количество элементов в индексе в зависимости от глубины дерева при списке в одном узле 4

Глубина	Кол-во
	элементов
3	64
4	256
5	1,024
6	4,096
7	16,384
8	65,536
9	262,144
10	1,048,576

#### Индексы



# Типы операций поиска в индексе

INDEX UNIQUE SCAN

Только проход по дереву

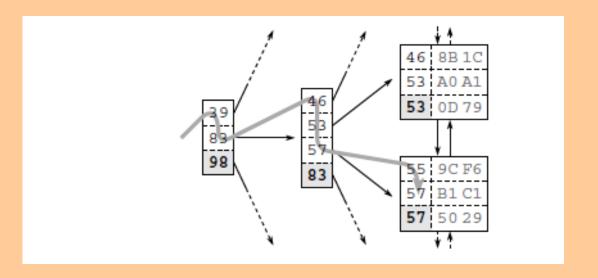
INDEX RANGE SCAN

Проход по дереву и по списку листа

TABLE ACCESS BY INDEX ROWID

После поиска (по индексу ) возврат одной строки

## Поиск в неуникальном индексе



- CREATE TABLE employees (
- employee\_id NUMBER NOT NULL,
- first\_name VARCHAR2(1000) NOT NULL,
- last\_name VARCHAR2(1000) NOT NULL,
- date\_of\_birth DATE NOT NULL,
- phone\_number VARCHAR2(1000) NOT NULL,
- CONSTRAINT employees\_pk PRIMARY KEY (employee\_id)
- );
- SELECT first\_name, last\_name
   FROM employees
   WHERE employee\_id = 123

SELECT first\_name, last\_name
 FROM employees
 WHERE employee\_id = 123

- CREATE TABLE employees (
- employee\_id NUMBER NOT NULL,
- first\_name VARCHAR2(1000) NOT NULL,
- last\_name VARCHAR2(1000) NOT NULL,
- date\_of\_birth DATE NOT NULL,
- phone\_number VARCHAR2(1000) NOT NULL,
- CONSTRAINT employees\_pk PRIMARY KEY (employee\_id)
- );
- Alter table employees add SUBSIDIARY\_ID NUMBER;
- Alter table employees drop primary key;
- Alter table employees add primary key (employee\_id ,SUBSIDIARY\_ID);
- CREATE UNIQUE INDEX employee\_pk
   ON employees (employee\_id, subsidiary\_id);

SELECT first\_name, last\_name FROM employees WHERE employee\_id = 123 AND subsidiary\_id = 30;

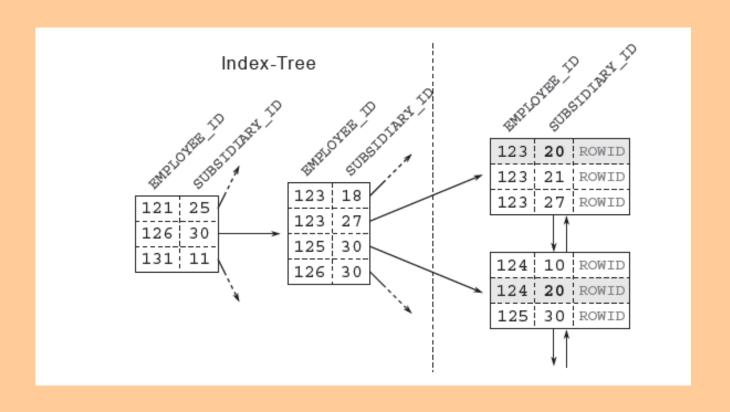
SELECT first\_name, last\_name FROM employees WHERE subsidiary id = 20;

```
| Id | Operation | Name | Rows | Cost |
| 0 | SELECT STATEMENT | | 106 | 478 |
|* 1 | TABLE ACCESS FULL | EMPLOYEES | 106 | 478 |

Predicate Information (identified by operation id):

1 - filter("SUBSIDIARY_ID"=20)
```

# Составной индекс



#### Разница между индексами

SELECT first\_name, last\_name FROM employees WHERE subsidiary\_id = 20;

#### CREATE UNIQUE INDEX employee\_pk

ON employees (employee\_id, subsidiary\_id);

# CREATE UNIQUE INDEX EMPLOYEES\_PK ON EMPLOYEES (SUBSIDIARY\_ID, EMPLOYEE\_ID);

Id  Operation	Name	Rows   Cost
0  SELECT STATEMENT   1   TABLE ACCESS BY INDEX ROWID   *2   INDEX RANGE SCAN	EMPLOYEES EMPLOYEE_PK	106   75     106   75     106   2
Predicate Information (identified 2 - access("SUBSIDIARY_ID"=20)	by operation	id):

### Влияния на другие запросы

SELECT first\_name, last\_name, subsidiary\_id, phone\_number
 FROM employees
 WHERE last\_name = 'Иванов'
 AND subsidiary\_id = 30;

- SELECT /\*+ NO\_INDEX(EMPLOYEES EMPLOYEE\_PK) \*/
- first\_name, last\_name, subsidiary\_id, phone\_number
- FROM employees
- WHERE last name = 'Иванов'
- AND subsidiary\_id = 30;

Скан полный, но выполняется быстрее. Почему?

#### После обновления статистики

Id  Operation   Name   Rows   Cost	
0   SELECT STATEMENT	Без
Predicate Information (identified by operation id):	
1 - filter("LAST_NAME"=' <b>Иванов</b> ) 2 - access("SUBSIDIARY_ID"=∰0)	

Без Индекса

#### CREATE INDEX emp\_name ON employees (last\_name);

Id   Operation	Name   Rows   Cost
0   SELECT STATEMENT	1   3
* 1   TABLE ACCESS BY INDEX ROWID    * 2   INDEX RANGE SCAN	EMPLOYEES   1   3   EMP_NAME   1   1
Predicate Information (identified by	operation id):
1 - filter("SUBSIDIARY_ID"=30) 2 - access("LAST_NAME"= 'Иванов'	

С Индексом

#### Функции и арифметика

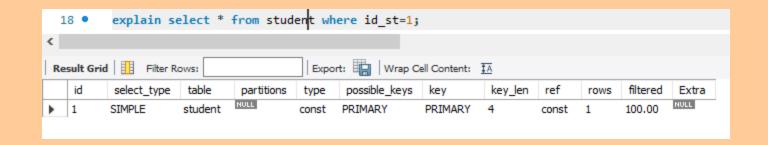
SELECT first\_name, last\_name, phone\_number FROM employees WHERE UPPER(last\_name) = UPPER('Иванов');

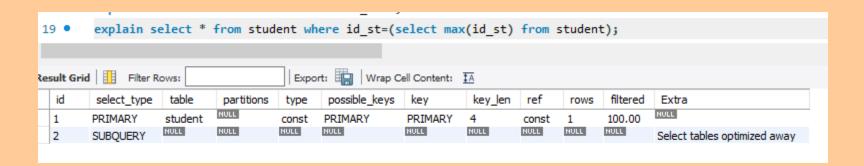
CREATE INDEX emp\_up\_name ON employees (UPPER(last\_name));

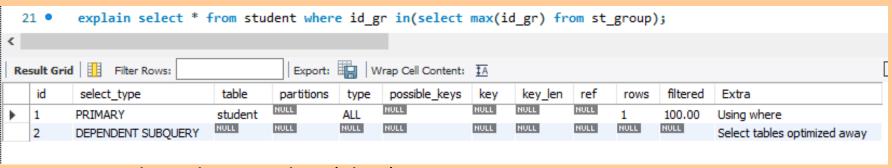
#### Как посмотреть план выполнения?

- {EXPLAIN | DESCRIBE | DESC}
   tbl\_name [col\_name | wild]
- {EXPLAIN | DESCRIBE | DESC} [explain\_type] {explainable\_stmt | FOR CONNECTION connection\_id}
- explain\_type: { FORMAT = format\_name }
- format\_name: { TRADITIONAL | JSON | TREE }

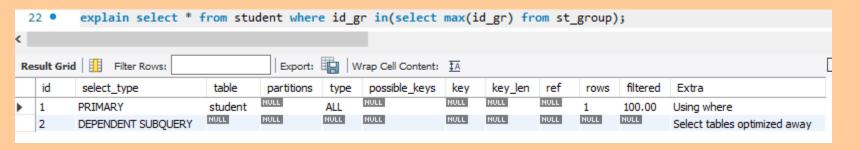
## Пример explain

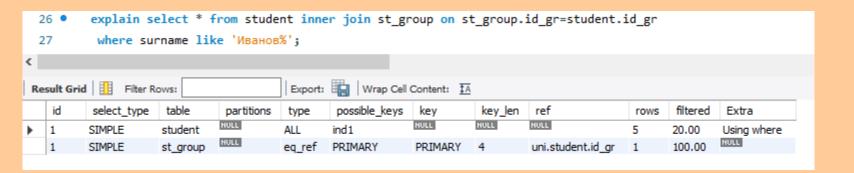






#### create index ind1 on student (id\_gr);





# EXPLAIN Выходные столбцы

Столбец	Имя в JSON	Назначение
id	select_id	Идентификатор SELECT
select type	None	Тип SELECTa
<u>table</u>	table_name	Таблица выходной строки
<u>partitions</u>	partitions	The matching partitions
<u>type</u>	access_type	Тип соединения
possible keys	possible_keys	Возможные индексы
<u>key</u>	key	Выбранный индекс
key len	key_length	Длина выбранного ключа
<u>ref</u>	ref	Столбец сравниваемый с индексом
rows	rows	Оценочное количество возвращенных строк
<u>filtered</u>	filtered	Процент строк отфильтрованных по условию
<u>Extra</u>	None	Дополнительная информация

#### type

eq\_ref

Проход только по В-дереву (только первичный уник)

ref, range

Проход по В-дереву и по списку в листе (диапазон)

index

Просматривает весь индекс

ALL

Просматривает всю таблицу на диске

# Другие столбцы

Using Index ("Extra")

Не просматривается таблица, т.к. все необходимое есть в индексе

PRIMARY ("key" / "possible\_keys")

PRIMARY – имя автоматически сгенерированного индекса на первичный ключ

• using filesort ("Extra")

Явная операция сортировки на диске или в памяти – ресурсоемкое т.к материлизация

implicit: no "using filesort" ("Extra")

При использовании limit если нет using filesort- то pipeline

# Подробности explain

explain select \* from student where id\_st=(select max(id\_st) from student);
SHOW WARNINGS;

						_							
	id	select	_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
•	1	PRIMA	RY	student	NULL	const	PRIMARY	PRIMARY	4	const	1	100.00	NULL
	2	2 SUBQUERY NULL		NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	Select tables optimized away
						,							
	Level	Code	Messa	age									
•	Note	1003	/* sele	ct#1 */ sel	ect '7' AS `id	_st`,'gr',	AS `surname`,'gr	1' AS `name`	','gr' AS `pa	atronym`	,'7' AS `	id_gr`fron	m `uni`.`student` where 1

#### **SHOW WARNINGS**

- <auto\_key> Автоматически сгенерированный ключ для временной таблицы.
- <cache>(expr) Выражение (например, скалярный подзапрос) выполняется один раз, и полученное значение сохраняется в памяти для последующего использования. Для результатов, состоящих из нескольких значений, можно создать временную таблицу, и вместо этого будет <temporary table>.
- <exists> (фрагмент запроса) Предикат подзапроса преобразуется в предикат EXISTS, а подзапрос преобразуется, чтобы его можно было использовать вместе с предикатом EXISTS.
- <in\_optimizer> (фрагмент запроса) Это внутренний объект оптимизатора без значения пользователя.
- <index\_lookup> (фрагмент запроса) Фрагмент запроса обрабатывается с использованием поиска по индексу для поиска подходящих строк.
- <if>(условие, expr1, expr2) Если условие истинно, оцените как expr1, иначе expr2.
- Is\_not\_null\_test> (выражение) Тест для проверки того, что выражение не оценивается как NULL.
- <materialize> (фрагмент запроса) используется материализация подзапроса. `Материализовалсяsubquery`.col\_name Ссылка на столбец col\_name во внутренней временной таблице, материализованной для хранения результата оценки подзапроса.
- <primary\_index\_lookup> (фрагмент запроса) Фрагмент запроса обрабатывается с использованием поиска первичного ключа для поиска подходящих строк.
- <Ref\_null\_helper> (expr1) Это внутренний объект оптимизатора без значения пользователя.
- /\* select # N \* / select\_stmt SELECT связан со строкой в расширенном выводе EXPLAIN со значением id, равным N.
- outer\_tables semi join (inner\_tables) Операция полусоединения. inner\_tables показывает таблицы, которые не были извлечены.
- **<temporary table>** Это представляет собой внутреннюю временную таблицу, созданную для кэширования промежуточного результата.

#### Анализ статистики таблицы

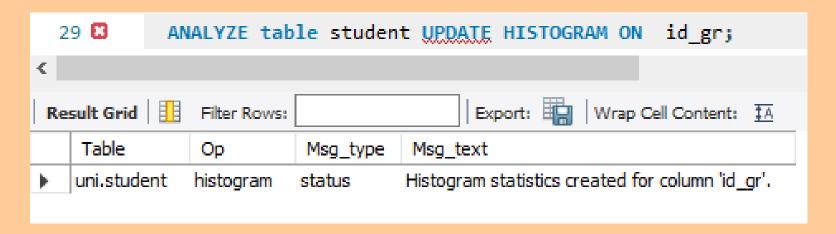
- ANALYZE [NO\_WRITE\_TO\_BINLOG | LOCAL]
   TABLE tbl\_name [, tbl\_name] ...
- ANALYZE [NO\_WRITE\_TO\_BINLOG | LOCAL]
   TABLE tbl\_name UPDATE HISTOGRAM ON
   col\_name [, col\_name] ... [WITH N BUCKETS]

•

 ANALYZE [NO\_WRITE\_TO\_BINLOG | LOCAL]
 TABLE tbl\_name DROP HISTOGRAM ON col\_name
 [, col\_name] ...

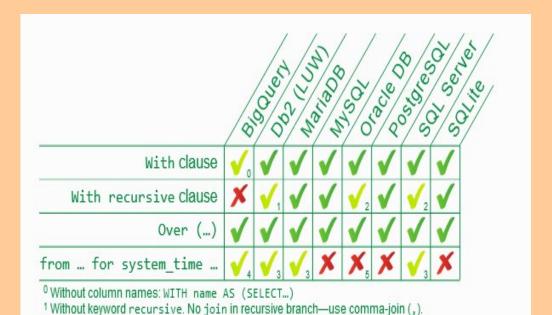
# Пример анализа статистики таблицы

2	28 • AN/	ALYZE ta	ble stude	n <b>t</b> ;									
<													
Result Grid Filter Rows: Export:													
	Table	Op	Msg_type	Msg_text									
<b>&gt;</b>	uni.student	analyze	status	OK	_								



# Оператор/ компонент WITH

## Поддержка оператора With



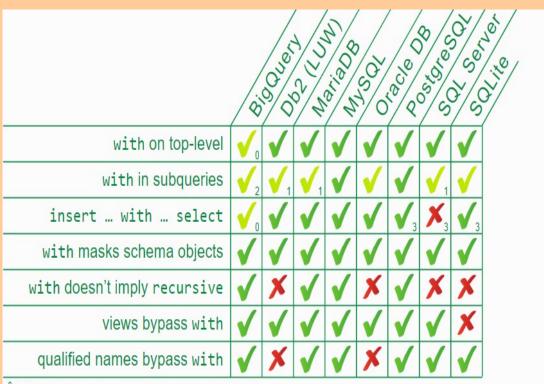
<sup>2</sup> Without keyword recursive.

3 Some minor omissions and variations.

<sup>5</sup> Alternative syntax. E.g. no for system\_time.

4 Without DDL (automatically). Only FOR SYSTEM\_TIME AS OF.

# Поддержка различных особенностей



Without column names: WITH name AS (SELECT...)

<sup>&</sup>lt;sup>1</sup> CTE in subquery cannot see global CTEs.

<sup>&</sup>lt;sup>2</sup> Without column names: WITH name AS (SELECT...) CTE in subquery cannot see global CTEs.

 $<sup>^3</sup>$  Supports proprietary syntax with  $\dots$  insert  $\dots$  select

#### Синтаксис

```
    WITH query_name (column_name1, ...) AS
```

```
• (SELECT ...)
```

•

• SELECT ...

```
    WITH query_name1 AS (

    SELECT ...
   , query_name2 AS (
    SELECT ...
      FROM query_name1
• SELECT ...
```

#### Пример использования with

select st\_group.id\_gr, number\_gr, count(id\_st) as cnt from st\_group join student on student.id\_gr=st\_group.id\_gr group by st group.id gr, number gr having count(id st)= (select max(cnt) from (select st group.id gr, count(id\_st) as cnt from st\_group join student on student.id gr=st group.id gr group by st group.id gr )q1);

• with q1 as (select st\_group.id\_gr, number\_gr, count(id\_st) as cnt from st\_group join student on student.id\_gr=st\_group.id\_gr group by st\_group.id\_gr, number\_gr) select q1.id\_gr, q1.number\_gr, cnt from q1 where cnt= (select max(cnt) from q1);

# Рекурсивный with

Recursive Common Table Expressions CTE

```
SELECT ... -- return initial row set
UNION ALL
SELECT ... -- return additional row sets
```

```
WITH RECURSIVE cte (n) AS

(
SELECT 1
UNION ALL
SELECT n + 1 FROM cte WHERE n < 5
)
SELECT * FROM cte;</pre>
```

# Рекурсивный with

```
WITH RECURSIVE cte AS

(

SELECT 1 AS n, CAST('abc' AS CHAR(20)) AS str

UNION ALL

SELECT n + 1, CONCAT(str, str) FROM cte WHERE n < 3

)

SELECT * FROM cte;
```

Re	sult G	rid 📙	Filter Rows
	n	str	
<b>•</b>	1	abc	
	2	abcab	С
	3	abcab	cabcabc

#### Ограничения врекурсивных

- ПОДЗапросах применяются некоторые синтаксические ограничения:
- Рекурсивная часть SELECT не должна содержать этих конструкций:
- Агрегатные функции, такие как SUM ()
- Оконные функции
- **GROUP BY**
- **ORDER BY**
- DISTINCT
- До MySQL 8.0.19 рекурсивная часть SELECT рекурсивного СТЕ также не могла использовать предложение LIMIT. Это ограничение снято в MySQL 8.0.19, и теперь в таких случаях поддерживается LIMIT вместе с дополнительным предложением OFFSET. Эффект на результирующий набор такой же, как при использовании LIMIT во внешнем SELECT, но также более эффективен, поскольку его использование с рекурсивным SELECT останавливает генерацию строк, как только будет создано их запрошенное количество.
- Эти ограничения не применяются к нерекурсивной части SELECT рекурсивного СТЕ. Запрет на DISTINCT распространяется только на членов UNION; UNION DISTINCT разрешен.
- Рекурсивная часть SELECT должна ссылаться на СТЕ только один раз и только в своем предложении FROM, а не в каком-либо подзапросе. Он может ссылаться на таблицы, отличные от СТЕ, и объединять их с

#### WITH

```
explain
with q1 as(select st_group.id_gr,number_gr,
count(id_st) as cnt
from st_group
join student on student.id_gr=st_group.id_gr
group by st_group.id_gr,number_gr)
select q1.id_gr,q1.number_gr,cnt from q1
where cnt=(select max(cnt) from q1);
```

	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
Þ	1	PRIMARY	<derived2></derived2>	NULL	ref	<auto_key0></auto_key0>	<auto_key0></auto_key0>	8	const	1	100.00	Using where
	3	SUBQUERY	<derived2></derived2>	NULL	ALL	NULL	NULL	NULL	NULL	3	100.00	HULL
	2	DERIVED	st_group	NULL	ALL	PRIMARY	NULL	NULL	NULL	3	100.00	Using temporary
	2	DERIVED	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index

### Подзапрос

```
explain select st_group.id_gr,number_gr,
  count(id_st) as cnt

from st_group join student
  on student.id_gr=st_group.id_gr
group by st_group.id_gr,number_gr
having count(id_st)=
(select max(cnt) from
(select st_group.id_gr,number_gr,count(id_st) as cnt
from st_group join student
  on student.id_gr=st_group.id_gr
group by st_group.id_gr,number_gr)q1);
```

Result Grid Filter Rows: Export: Wrap Cell Content: 🖽												
	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
•	1	PRIMARY	st_group	NULL	ALL	PRIMARY	NULL	NULL	HULL	3	100.00	Using temporary
	1	PRIMARY	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index
	2	SUBQUERY	<derived3></derived3>	NULL	ALL	NULL	NULL	NULL	NULL	3	100.00	NULL
	3	DERIVED	st_group	NULL	ALL	PRIMARY	NULL	NULL	NULL	3	100.00	Using temporary
	3	DERIVED	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index

	id	d	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
•	1		PRIMARY	<derived2></derived2>	NULL	ref	<auto_key0></auto_key0>	<auto_key0></auto_key0>	8	const	1	100.00	Using where
	3		SUBQUERY	<derived2></derived2>	NULL	ALL	NULL	NULL	NULL	NULL	3	100.00	HULL
	2		DERIVED	st_group	NULL	ALL	PRIMARY	NULL	NULL	NULL	3	100.00	Using temporary
	2		DERIVED	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index

Re	sult Gri	d 📗 Filter R	ows:		Export: Wrap Cell Content: IA							
	id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
•	1	PRIMARY	st_group	NULL	ALL	PRIMARY	NULL	NULL	NULL	3	100.00	Using temporary
	1	PRIMARY	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index
	2	SUBQUERY	<derived3></derived3>	NULL	ALL	NULL	NULL	NULL	NULL	3	100.00	NULL
	3	DERIVED	st_group	NULL	ALL	PRIMARY	NULL	NULL	NULL	3	100.00	Using temporary
	3	DERIVED	student	NULL	ref	id_gr_st_1	id_gr_st_1	5	uni.st_group.id_gr	1	100.00	Using index

#### Литература

