

姓名	学号
崔璨明	20337025

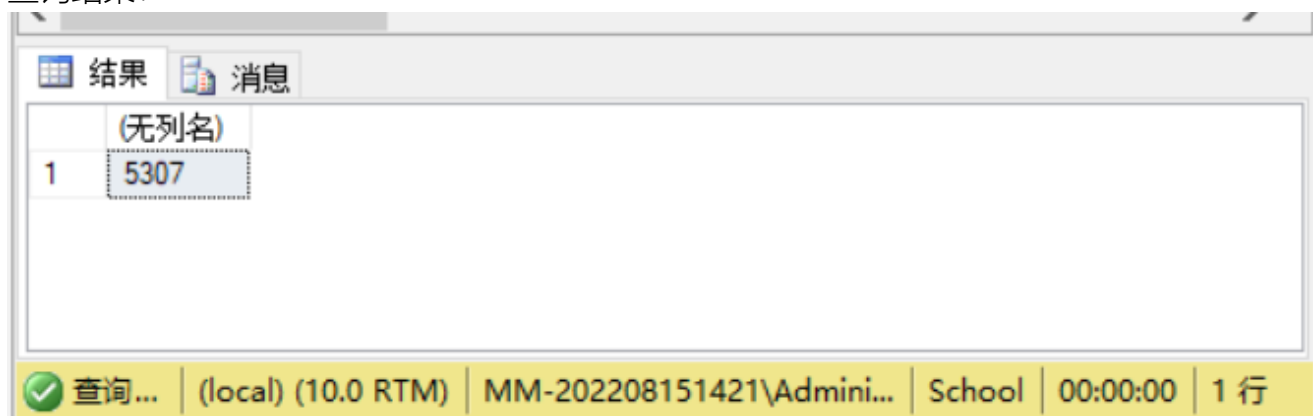
1.

通过查询选修课程C++的学生的人数，其中成绩合格的学生人数，不合格的学生人数，讨论NULL值的特殊含义。

选修课程C++的学生的人数：

```
select COUNT(*)  
FROM COURSES,CHOICES  
WHERE COURSES.cname='c++'  
and COURSES.cid=CHOICES.cid
```

查询结果：



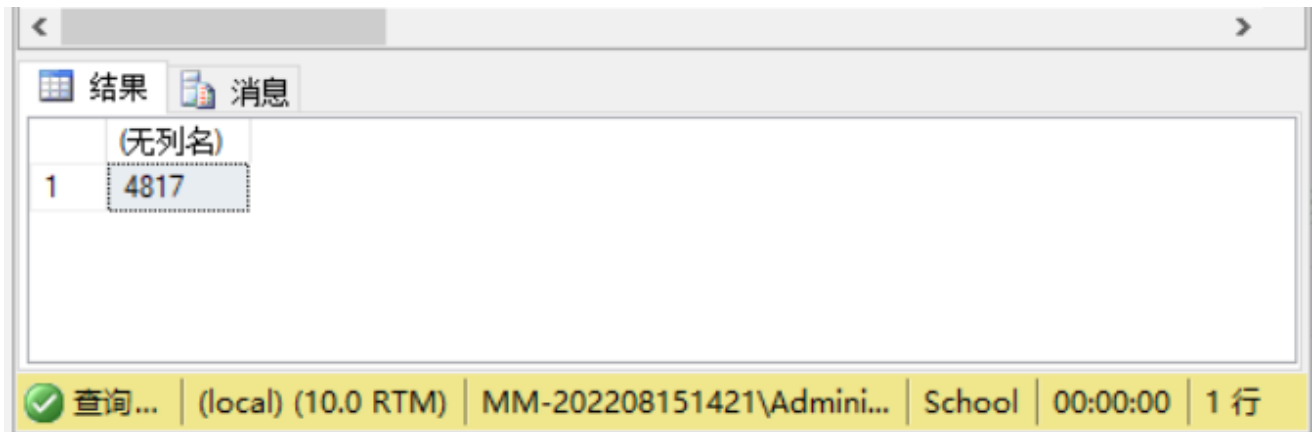
The screenshot shows a database query result window. At the top, there are two tabs: '结果' (Results) and '消息' (Messages). Below the tabs, there is a table with one column labeled '(无列名)' (No column name) and one row with the value '5307'. The status bar at the bottom indicates '查询...' (Query...), '(local) (10.0 RTM)', 'MM-202208151421\Admini...', 'School', '00:00:00', and '1 行' (1 row).

(无列名)
5307

成绩合格的学生人数：

```
SELECT count(CHOICES.score)  
FROM COURSES,CHOICES  
WHERE COURSES.cname='c++'  
and COURSES.cid=CHOICES.cid  
and CHOICES.score>=60
```

查询结果：



The screenshot shows a database query result window. At the top, there are tabs for '结果' (Results) and '消息' (Messages). Below the tabs is a table with one row and one column. The column header is '(无列名)' (No column name). The row contains the value '4817'. At the bottom of the window, there is a status bar with a green checkmark icon, the text '查询...' (Query...), and several details: '(local) (10.0 RTM)', 'MM-202208151421\Admini...', 'School', '00:00:00', and '1 行' (1 row).

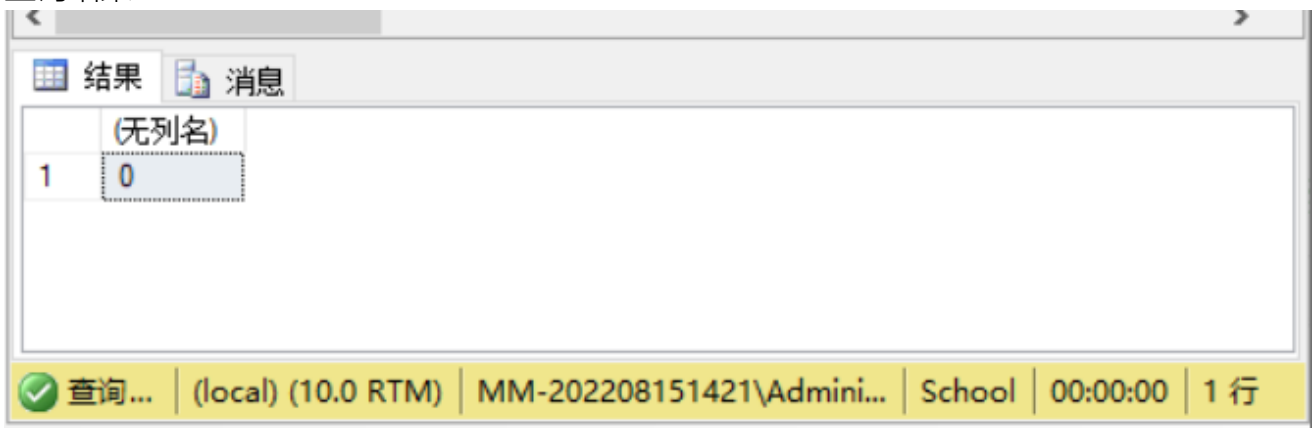
(无列名)
4817

查询... | (local) (10.0 RTM) | MM-202208151421\Admini... | School | 00:00:00 | 1 行

成绩不合格的学生人数：

```
SELECT count(CHOICES.score)
FROM COURSES,CHOICES
WHERE COURSES.cname='c++'
and COURSES.cid=CHOICES.cid
and CHOICES.score<60
```

查询结果：



The screenshot shows a database query result window. At the top, there are tabs for '结果' (Results) and '消息' (Messages). Below the tabs is a table with one row and one column. The column header is '(无列名)' (No column name). The row contains the value '0'. At the bottom of the window, there is a status bar with a green checkmark icon, the text '查询...' (Query...), and several details: '(local) (10.0 RTM)', 'MM-202208151421\Admini...', 'School', '00:00:00', and '1 行' (1 row).

(无列名)
0

查询... | (local) (10.0 RTM) | MM-202208151421\Admini... | School | 00:00:00 | 1 行

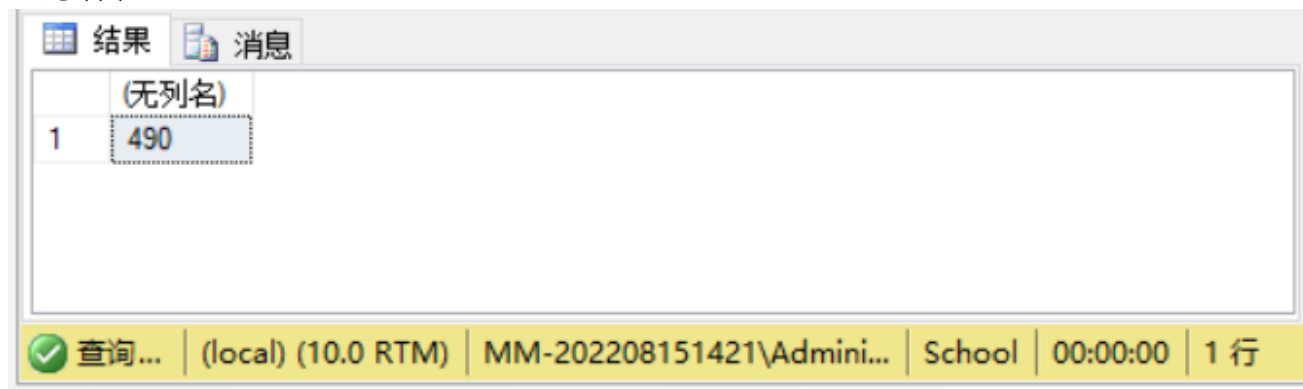
可以看见 $4817 + 0 \neq 5307$ ，这是因为选修课程C++的学生中成绩存在空值的原因。当score取NULL值时，NULL与所有的比较运算符都是不匹配的，所以都不会出现在统计结果中。

分析可知应该存在490个score为NULL的元组。验证如下：

查询选修课程C++的学生中成绩非空的人数：

```
select COUNT(*)
FROM COURSES,CHOICES
WHERE COURSES.cname='c++'
and COURSES.cid=CHOICES.cid
and CHOICES.score is NULL
```

查询结果：



The screenshot shows a database query result window with two tabs: '结果' (Results) and '消息' (Messages). The '结果' tab is active, displaying a table with one column labeled '(无列名)' (No column name) and one row containing the value '490'. The status bar at the bottom indicates the query was successful, the database is '(local) (10.0 RTM)', the user is 'MM-202208151421\Admini...', the database is 'School', the time is '00:00:00', and there is '1 行' (1 row).

(无列名)
490

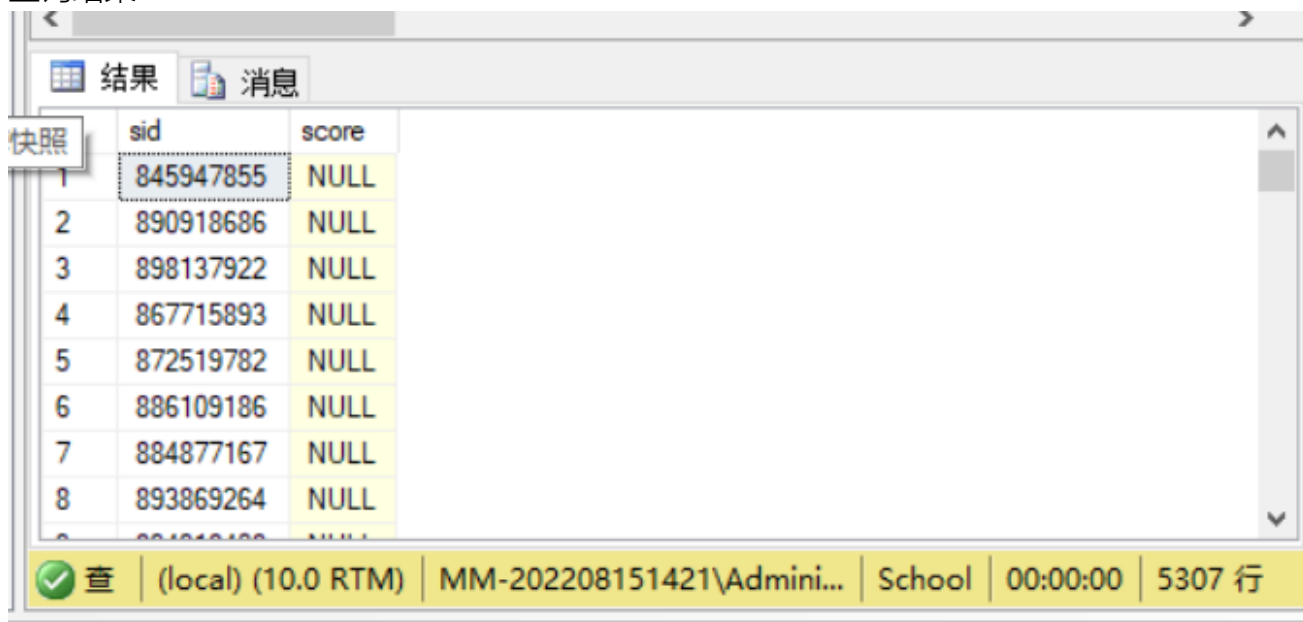
$4817 + 0 + 490 = 5307$ 。

2.

查询选修课程C++的学生的编号和成绩，使用 ORDER BY按成绩进行排序时,取NULL的项是否出现在结果中？如果有，在什么位置？

```
SELECT CHOICES.sid,CHOICES.score
FROM COURSES,CHOICES
WHERE COURSES.cname='c++'
and COURSES.cid=CHOICES.cid
ORDER BY score ASC
```

查询结果：



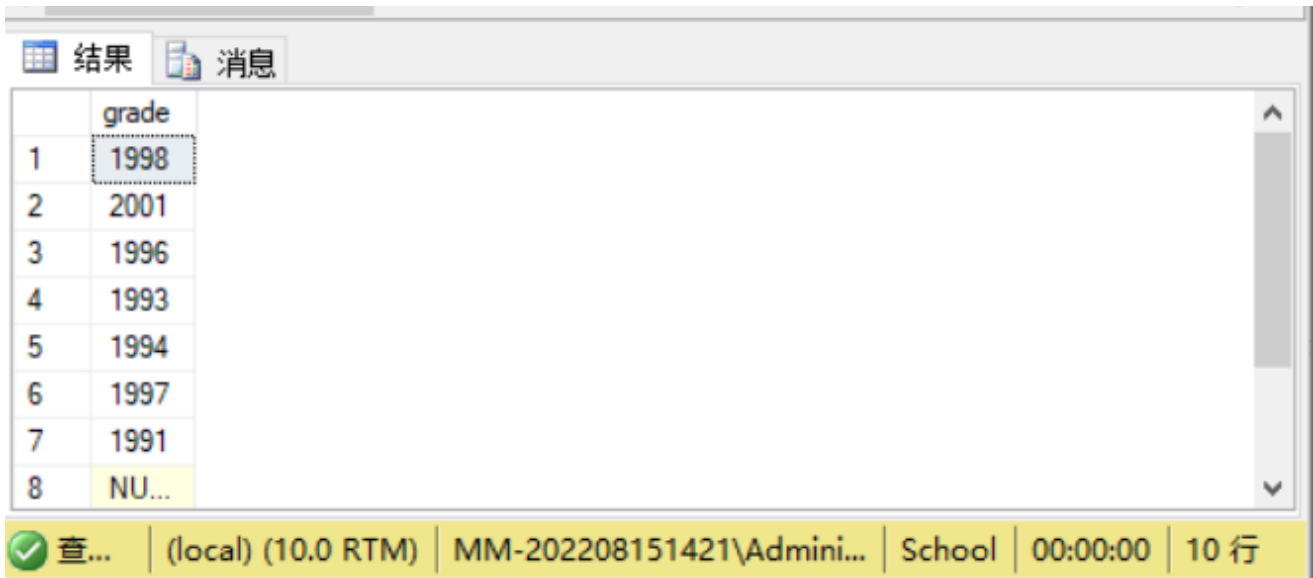
The screenshot shows a database query result window with two tabs: '结果' (Results) and '消息' (Messages). The '结果' tab is active, displaying a table with two columns: 'sid' and 'score'. The table contains 8 rows, all with NULL scores. The status bar at the bottom indicates the query was successful, the database is '(local) (10.0 RTM)', the user is 'MM-202208151421\Admini...', the database is 'School', the time is '00:00:00', and there are '5307 行' (5307 rows).

	sid	score
1	845947855	NULL
2	890918686	NULL
3	898137922	NULL
4	867715893	NULL
5	872519782	NULL
6	886109186	NULL
7	884877167	NULL
8	893869264	NULL

可见，取NULL的项会出现在查询结果中，当采用ASC升序排序时，出现在最前面。当采用DESC降序排序时，出现在最末尾：


```
select grade
from STUDENTS
GROUP BY grade
```

查询结果：



	grade
1	1998
2	2001
3	1996
4	1993
5	1994
6	1997
7	1991
8	NU...

查... | (local) (10.0 RTM) | MM-202208151421\Admini... | School | 00:00:00 | 10 行

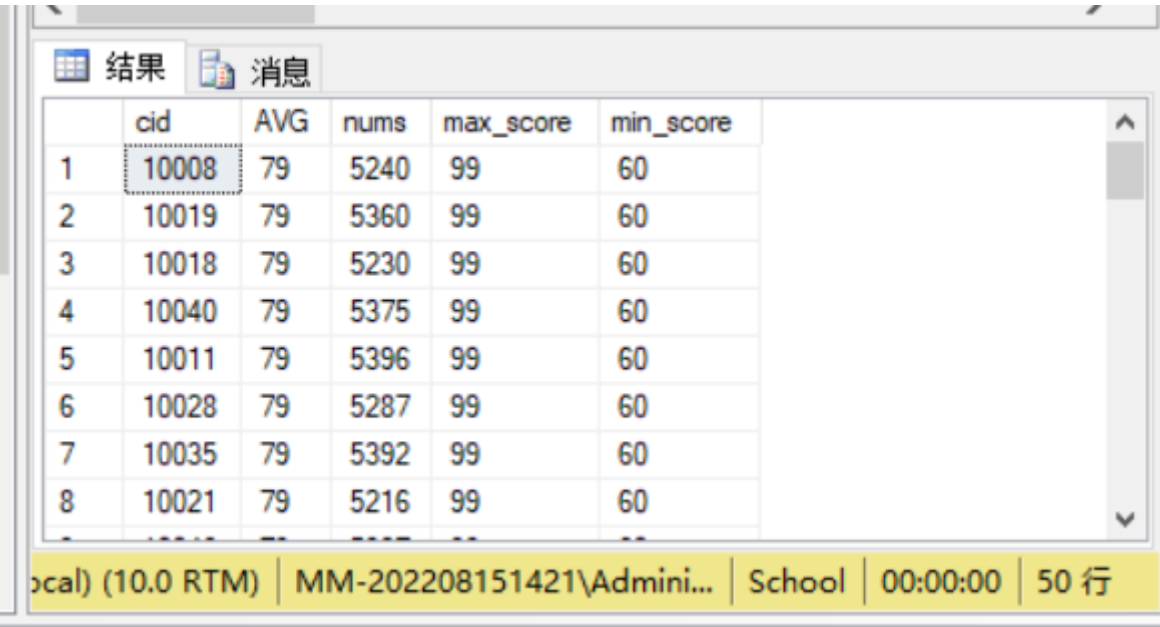
可以得到10个组，但有一个是NULL，和现实情况应该得到9个组不同。

5.

结合分组,使用集合函数求每个课程选修的学生的平均分,总的选课记录数,最高成绩,最低成绩,讨论考察取空值的项对集合函数的作用的影响。

```
select cid,AVG(score) AVG,COUNT(*) nums,MAX(score) max_score,MIN(score)
min_score
from CHOICES
group by cid
```

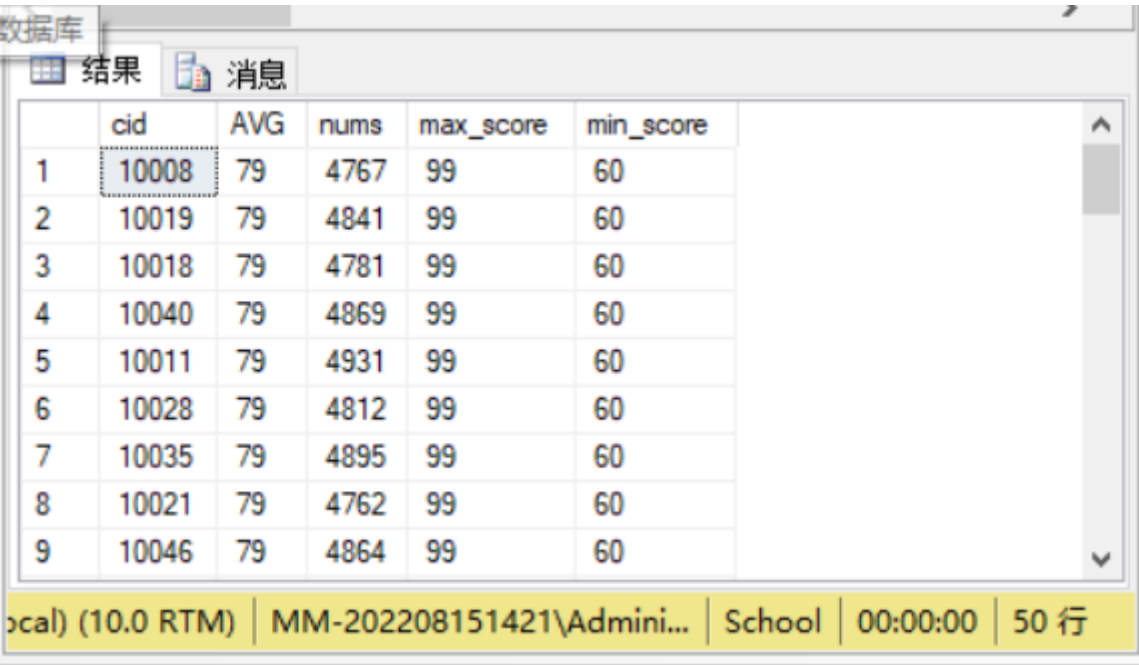
查询结果：



	cid	AVG	nums	max_score	min_score
1	10008	79	5240	99	60
2	10019	79	5360	99	60
3	10018	79	5230	99	60
4	10040	79	5375	99	60
5	10011	79	5396	99	60
6	10028	79	5287	99	60
7	10035	79	5392	99	60
8	10021	79	5216	99	60

ocal) (10.0 RTM) | MM-202208151421\Admini... | School | 00:00:00 | 50 行

此时cid为10008的学生的nums是5240，当将 COUNT(*) 改成 COUNT(SCORE) 后，得到的结果如下：



	cid	AVG	nums	max_score	min_score
1	10008	79	4767	99	60
2	10019	79	4841	99	60
3	10018	79	4781	99	60
4	10040	79	4869	99	60
5	10011	79	4931	99	60
6	10028	79	4812	99	60
7	10035	79	4895	99	60
8	10021	79	4762	99	60
9	10046	79	4864	99	60

ocal) (10.0 RTM) | MM-202208151421\Admini... | School | 00:00:00 | 50 行

此时cid为10008的学生的nums是4767，可见学生的选课数目nums发生了改变，这是因为在使用 COUNT(*) 时，虽然score是NULL，但该记录存在，所以也计算在内。而 COUNT(SCORE)，将取值为NULL的项忽略，导致数目减少。

6.

采用嵌套查询的方式,利用比较运算符和谓词ALL的结合来查询表 STUDENTS中最晚入学的学生年级。当存在 GRADE取空值的项时,考虑可能出现的情况,并解释。

当不去除NULL值时：

```

select distinct grade
from STUDENTS as s1
where s1.grade >= all(
    select s2.grade
    from STUDENTS as s2
)

```

查询结果：

结果		消息	
grade			
local) (10.0 RTM)		MM-202208151421\Admini...	School 00:00:00 0 行

当去除了NULL值时：

```

select distinct grade
from STUDENTS as s1
where s1.grade >= all(
    select s2.grade
    from STUDENTS as s2
    where s2.grade is not NULL
)

```

查询结果：

结果		消息	
grade			
1	2001		
local) (10.0 RTM)		MM-202208151421\Admini...	School 00:00:00 1 行

在嵌套的子查询中，如果存在 `GRADE`取空值的项时，那么`s1.grade>=all`无法成立，因此查询的结果为空，正如上面查询所示，要将`NULL`的值去掉才能查询到正确结果。
