



本科生实验报告

学生姓名： 丁晓琪

学生学号： 22336057

专业名称： 计科

1: 实验任务

- (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。
- (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？
- (3)在上面的查询的过程中，如果加上保留字 DISTINCT会有什么效果呢？
- (4)按年级对所有的学生进行分组，能得到多少个组？与现实的情况有什么不同？
- (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。
- (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。

2: 实验过程

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：

- 查询选修C++的学生的人数：（补充）其实直接选择 `count(*)` 看选修的学生数不太正确，因为一个学生可能选修多次，但是此处还是延用ppt上的查询方法

```
use School
select count(*)
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
          )
```

99 %

结果 消息

	(无列名)
1	6031

- 查询成绩合格的学生数：加上条件 `score>=60`

```

use School
select count(*)
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
            ) and score>=60

```

99 %

结果 消息

	(无列名)
1	4817

- 查询成绩不合格的学生数：将条件改为 `score<60`

```

use School
select count(*)
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
            ) and score<60

```

99 %

结果 消息

	(无列名)
1	724

- NULL 的含义：6031>4817+724，这是由于 NULL 无论在什么比较中得到的布尔值都为假，因为 NULL 代表的意思是空值，未知的值，未知怎么比较都是不匹配的

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？

- 用 order by 对C++的成绩排序：默认按照升序排列，成绩取 NULL 出现在结果的开头位置

SQLQuery1.sql...琪的电脑\丁晓琪 (70))* 对象资源管理器详细

```

use School
select sid,score
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
            )
order by score

```

99 %

结果 消息

	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
9	834918408	NULL
10	873557912	NULL
11	802605966	NULL

3. 在上面的查询的过程中，如果加上保留字 DISTINCT会有什么效果呢？

- 只对 score 使用 DISTINCT 时：对所有重复的 score 只保留一个，包括 NULL 值

SQLQuery1.sql...琪的电脑\丁晓琪 (70))* 对象资源管理器

```
use school
select distinct score
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
          )
order by score
```

99 %

结果 消息

	score
1	NULL
2	50
3	51
4	52
5	53
6	54
7	55
8	60
9	61
10	62
11	62

- 对 score, sid 都使用 distinct: 对相同的 score, sid 元组只保留一个结果

SQLQuery1.sql... 填的电脑\J 晓琪 (/0))* 对象资源管

```
use School
select distinct score, sid
from CHOICES
where cid=(select cid
            from COURSES
            where cname='c++'
          )
order by score
```

99 %

结果 消息

	score	sid
1	NULL	800554079
2	NULL	800575453
3	NULL	800579690
4	NULL	800758208
5	NULL	800895421
6	NULL	800989091
7	NULL	801045247
8	NULL	801095737
9	NULL	801197660
10	NULL	801259696
11	NULL	801863111
12	NULL	802566342
13	NULL	802605966
14	NULL	802777255
15	NULL	802967609
16	NULL	803258105
17	NULL	803285444
18	NULL	803478243
19	NULL	804102167
20	NULL	804264839
21	NULL	804667865
22	NULL	804921476

4. 按年级对所有的学生进行分组, 能得到多少个组?与现实的情况有什么不同?

- 按年级对所有学生分组: 一共得到15个组, 且有一个年级为 NULL 的组

SQLQuery2.sql...琪的电脑\丁晓琪 (51))* X

```

use School
select grade
from STUDENTS
group by grade

```

99 %

结果 消息

	grade
1	1998
2	2004
3	2001
4	1996
5	2002
6	1999
7	1993
8	2000
9	1994
10	1997
11	1991
12	NULL
13	2003
14	1992
15	1995

- 再对 grade 为 NULL 的学生元组进行查询：确实存在多名学生的年级未知，但是现实中不存在年级 NULL,且学生的年级应该是明确的

SQLQuery2.sql...琪的电脑\丁晓琪 (51))* X

```

use School
select *
from STUDENTS
where grade is NULL

```

99 %

结果 消息

	sid	sname	email	grade
1	800032434	mglhxm	k5a1wz8@pvq.net	NULL
2	800035156	nsyxl	zjsc2@xui.gov	NULL
3	800052159	zvjteiv	41ku@uuyqx.gov	NULL
4	800079081	yecowu	ofwb1j@gpv.edu	NULL
5	800097178	bfqul	bjel4@zqqvp.edu	NULL
6	800110485	pmkfspshz	fi7b@cyxri.org	NULL
7	800110633	cekbmkhgc	pli2@rhn.org	NULL
8	800111391	pwnxku	wsri@yjbnet.net	NULL
9	800122464	donvolh	p4d6e@nezo.com	NULL
10	800124722	szcuo	7r41zv@nvdo.com	NULL
11	800139095	zuehilkka	ezsk@hrnce.edu	NULL
12	800141278	tljme	z8vi@yzs.com	NULL
13	800151262	oxwgyhprt	kyti@hxchk.net	NULL
14	800171998	dcidv	xvbrfdf@ehpzu.net	NULL
15	800178280	teyqkp	vp9fnuc@krfkm.org	NULL
16	800184238	sxkxyqqo	s703sw6@swvjin.edu	NULL
17	800209484	jkwuunn	nnhe6@meo.gov	NULL
18	800211044	lednohj	y_gmm6x@gnn.gov	NULL

- 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。

- 求每个课程选修的学生平均分(AVG(score)), 总的选课记录(count(*)), 最高成绩(MAX(score)), 最低成绩(MIN(score))

SQLQuery1.sql...琪的电脑\丁晓琪 (70))*

```

use School
select cid, AVG(score) as [avg], COUNT(*) as [count],
       MAX(score) as [max], MIN(score) as [min]
from CHOICES
group by cid

```

100 %

结果 消息

	cid	avg	count	max	min
15	10036	75	6043	99	50
16	10045	76	6041	99	50
17	10025	75	5955	99	50
18	10030	76	5985	99	50
19	10049	76	5945	99	50
20	10007	75	5965	99	50
21	10026	75	5881	99	50
22	10010	75	6027	99	50
23	10024	75	6013	99	50
24	10033	75	6039	99	50
25	10041	76	6027	99	50
26	10002	75	6013	99	50
27	10020	76	6027	99	50
28	10037	75	5909	99	50
29	10048	76	5927	99	50
30	10016	76	5867	99	50
31	10039	76	5951	99	50
32	10004	76	6110	99	50
33	10003	75	5975	99	50
34	10006	76	6090	99	50
35	10022	75	5818	99	50
36	10044	75	5910	99	50

- 对 count() 的影响: 将 count() 换为 count(score) 进行查询, 发现 count 列有一些行小了, 这是因为 count(*) 会将所有计数所有行 (包括 score=NULL 的), 但是 count(score) 会忽略 score=NULL 的行

SQLQuery1.sql...琪的电脑\丁晓琪 (70))*

```

use School
select cid,AVG(score) as [avg],COUNT(score) as [count],
MAX(score) as [max],MIN(score) as [min]
from CHOICES
group by cid

```

100 %

结果 消息

	cid	avg	count	max	min
1	10008	75	5512	99	50
2	10019	75	5555	99	50
3	10018	76	5520	99	50
4	10040	75	5596	99	50
5	10011	76	5621	99	50
6	10028	75	5567	99	50
7	10035	76	5607	99	50
8	10021	76	5462	99	50
9	10046	76	5606	99	50
10	10032	75	5614	99	50
11	10005	76	5541	99	50
12	10050	75	5609	99	50
13	10034	76	5506	99	50

- 其他 AVG,MAX,MIN(score) 都会忽略score=NULL的行, 任何课程的 max,min 都不为 NULL

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。

- 查询最晚入学的学生的年级:
 - 在 where 子句的子查询中查询出学生所有入学年级 (不包括 NULL)
 - 在 where 子句中用 \geq ALL(子查询结果), 查询出大于等于所有不为 NULL 的年级的入学年级, 也即选出最晚入学年级
 - 存在 grade=NULL 的元组时, NULL 和子查询结果中所有grade的比较布尔值都为0, 不满足 where 子句的条件

SQLQuery2.sql...琪的电脑\丁晓琪 (51))*

```

select distinct grade
from STUDENTS
where grade >=ALL(select distinct grade
                    from STUDENTS
                    where grade is NOT NULL
                  )

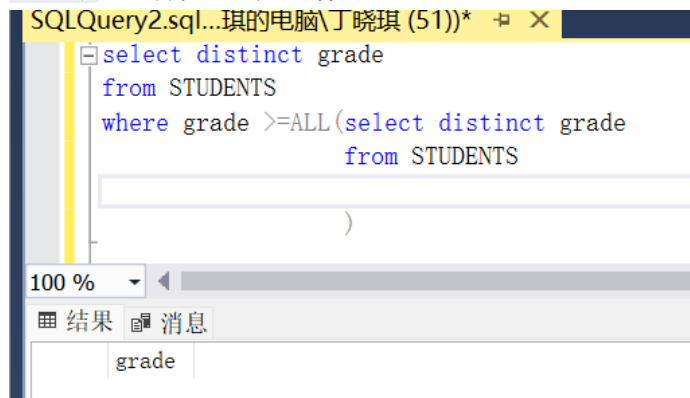
```

100 %

结果 消息

	grade
1	2004

- 如果子查询包含 `grade=NULL`，此时整个查询结果为空：因为 `STUDENTS` 任何元组的 `grade` 和子查询结果中的 `NULL` 比较时，布尔值结果都为0,即不存在元组满足 `where` 子句中 `>=ALL` (子查询结果) 的条件，查询方法错误



3: 实验总结

- `NULL` 和所有数值的所有的比较运算符不匹配
- 按序排列时，`NULL` 没有被忽略而是当作最小值处理
- `distinct` 会把所有的 `NULL` 当成一个值
- `Group by` 会把所有 `NULL` 值当成一个分组
- 集合函数中，除了 `COUNT(*)` 计算元组时把 `NULL` 计算，其他集合函数忽略取空值的项
- 对 `NULL` 求集合函数，除了 `count(*)` 为0，其他都为 `NULL`
- 其他：

空集。SQL对于“ `HOUR<=ALL(空集)` ”返回 **TRUE**，所以上述语句也能成功查询到这唯一的元组。

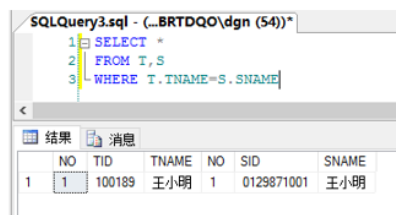
8.对于下面学生表和教师表，对这两个表作对姓名的等值连接运算,找出既是老师又是学生的人员的学生编号和教师编号.

学生表

	NO	SID	SNAME
1	1	0129871001	王小明
2	2	0129871002	李兰
3	3	0129871005	NULL
4	4	0129871004	关红

教师表

	NO	TID	TNAME
1	1	100189	王小明
2	2	1001801	李晓
3	3	100121	NULL
4	4	100128	NULL



等值连接后，姓名为**NULL**的项都被忽略了。即，两个属于不同表，并且对应列都为 `NULL` 的列是不会发生连接的，也不会出现在结果中。