

# 数据高级查询

Email: 18442056@QQ.com

# 学习目标

- TOP 限制行数
- ORDER BY 结果排序
- DISTINCT 关键字
- 聚合函数 —— AVG(), COUNT(), MAX(), MIN(), SUM()
- GROUP BY 分组结果集

# 问题1:

- 教务学生管理系统的信息显示页面大小有限，不能一次性展现所有的学生信息，能否按分页进行显示，比如一次一个页面只显示15个学生信息。

# 数据查询-限制行数

限制固定行数 第一页

```
SELECT TOP 15 stuid,stuName, sex  
FROM T_StudentInfo
```

第二页

```
SELECT TOP 15 stuid,stuName, sex  
FROM T_StudentInfo  
where stuid > '130015'
```

# 数据查询-限制行数

使用百分比限制行数

```
SELECT TOP 20 PERCENT *  
FROM T_StudentInfo  
WHERE clsid = 4
```

# 数据查询-排序1

## 升序排列

```
select stuid as 学号, stuname as 姓名, sex as 性别  
from T_StudentInfo  
order by stuid asc      --升序
```

## 降序排列

```
select stuid as 学号, stuname as 姓名, sex as 性别  
from T_StudentInfo  
WHERE clsid= 3  
ORDER BY firstScore DESC  -- 降序
```

# 数据查询-排序2

完整的按学号排序的分页查询

```
select top 15 *  
from T_studentInfo  
order by stuid asc
```

查询14软件1班完整的按学号排序的分页查询

```
select top 15 *  
from T_studentInfo  
where clsid = 3  
order by stuid asc
```

# 数据查询-排序 2

## 按多列排序

```
select * from T_StudentInfo  
order by clsid asc, stuname asc
```

特点：先按clsid中的值进行升序排序，再在每个clsid的基础上使用stuname进行二次升序排序



# 问题2

- 如何从数以千计的班级信息记录中分析出学校有哪些学院，另外再分析出哪些学院的哪些专业？

# DISTINCT 消除重复项

●消除查询列中重复的信息项

	classid	classname	creategrade	stucount	college	speciality	teacher
1	3	1班	2014	7	信息工程学院	软件技术	谭小文
2	4	1班	2014	8	信息工程学院	网络技术	赵应先
3	5	1班	2013	0	信息工程学院	网络技术	赵应先
4	6	1班	2013	0	信息工程学院	软件技术	谭小文
5	7	1班	2014	8	交通机电学院	电气自动化	李明亮
6	8	2班	2014	0	交通机电学院	电气自动化	赵应先
7	10	1班	2014	0	信息工程学院	电子商务	卢兰江
8	11	1班	2013	0	信息工程学院	电子商务	俞铮
9	12	1班	2014	0	会计学院	会计电算化	龙可菲
10	13	2班	2014	0	会计学院	会计电算化	龙可菲
11	14	3班	2014	0	会计学院	会计电算化	吴圆圆
12	16	4班	2014	0	会计学院	会计电算化	吴圆圆
13	17	5班	2014	0	会计学院	会计电算化	林涛

重复的专业

重复的学院

# DISTINCT 格式

```
SELECT DISTINCT 列 FROM TABLE
```

- 将指定列中项的重复数据消除

```
select distinct college  
from T_Classes
```

## DISTINCT 格式 (2)

**SELECT DISTINCT** 列1, 列2 **FROM** TABLE

- 将指定多个列中的组合项的重复数据消除
- 例如：信息工程系 软件技术
- 信息工程系 网络技术
- 是属于不同值项的组合列

college	speciality
会计学院	会计电算化
建工学院	工程造价
建工学院	装潢设计
交通机电学院	电气自动化
信息工程学院	电子商务
信息工程学院	软件技术
信息工程学院	网络技术

```
select distinct college,speciality  
from T_Classes
```

随着选择列的增多，  
**DISTINCT**将失去意义

# 问题3

- 如何统计数以万计的学生信息，比如学生的总人数，各班的学生人数。某门课程学生的考试平均分，最高分，最低分。

# 聚合函数

- 实现查询列中项的统计、汇总功能的函数
- **AVG**(列) 返回所有列项的平均值
- **COUNT**(列) 返回所有列项的数目
- **MAX**(列) 返回所有列项中的最大值
- **MIN**(列) 返回所有列项中的最小值
- **SUM**(列) 返回所有列项的总值

# 聚合函数-1

**COUNT** 返回所有查询出列项的总数

```
SELECT COUNT ( 列 / 表达式)  
FROM TABLE  
WHERE 条件
```

统计学校所有学生的人数

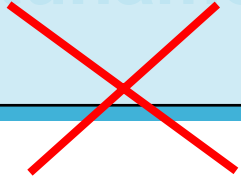
```
select COUNT(stuid) as 学生数量 from  
T_StudentInfo
```

```
select COUNT(*) as 学生数量 from T_StudentInfo
```

## 聚合函数与一般函数

- 作为统计结果的聚合函数列与一般结果的普通列不能同时出现在查询中。

```
select COUNT(stuid) as 学生数量, stuname  
as 姓名 from T_StudentInfo
```



- 聚合函数可以在条件查询的基础上进行统计

```
select COUNT(*) as '14软件班学生数量' from  
T_StudentInfo  
where clsid = 3
```



# 聚合函数-1

**SUM** 返回所有查询出列项的总值

```
SELECT SUM(列 / 表达式)  
FROM TABLE  
WHERE 条件
```

学校教务要求考试科给出关于  
**2014**软件技术**1**班 《工程数  
据库》课程的所有参加考试学  
生的总成绩

# 聚合函数- SUM

```
select plancode as 排课编号, SUM(score) as 成绩  
总分  
from T_StuScore  
where plancode = 1
```

```
select  
'2014软件技术班工程数据库' as 考试科目,  
SUM(score) as 成绩总分  
from T_StuScore  
where plancode = 1
```

# 聚合函数-1

**AVG** 返回所有查询出列项的平均值

```
SELECT AVG( 列 / 表达式)  
FROM TABLE  
WHERE 条件
```

学校教务要求考试科给出关于  
**2014**软件技术**1**班 《工程数  
据库》课程的所有参加考试学  
生的平均分

# 聚合函数- **AVG**

```
select '2014软件技术班工程数据库' as 考试科目,  
AVG(score) as 成绩总分  
from T_StuScore  
where plancode = 1
```

	考试科目	成绩总分
1	2014软件技术1班工程数据库	59.428571

## 聚合函数- **AVG** (续)

统计前将非正常考试的成绩进行去除

```
select '2014软件技术班工程数据库' as 考试科目,  
AVG(score) as 成绩总分  
from T_StuScore  
where plancode = 1 and examtype = '正常'
```

	考试科目	成绩总分
1	2014软件技术1班工程数据库	69.333333

# 聚合函数-3

**MAX、MIN** 返回所查询出的列值中的最大值、最小值

**2014软件技术1班 《工程数据库》 考试成绩的平均分、最高分和最低分**

```
SELECT '2014软件技术班工程数据库' as 考试科目,  
AVG(score) AS 平均成绩,  
MAX (score ) AS 最高分,  
MIN (score ) AS 最低分  
From T_StuScore  
where plancode = 1 and examtype = '正常'
```

	考试科目	平均成绩	最高分	最低分
1	2014软件技术1班工程数据库	69.333333	91.0	27.0

# 聚合函数的错误用法

## 1、SUM、AVG、MIN、MAX统计的列非数字类型

```
select SUM(stuid) from T_StuScore  
where plancode = 1 and examtype = '正常'
```

## 2、SUM、AVG、MIN、MAX统计的列非单个列

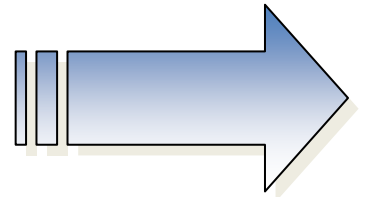
```
select SUM(*) from T_StuScore  
where plancode = 1 and examtype = '正常'
```

## 3、select 后跟了非聚合列

```
select plancode, examtype, SUM(score) from  
T_StuScore  
where plancode = 1 and examtype = '正常'
```

# 问题4

如果我不想一个一个统计每个班的每门课程的成绩信息，而是一次性希望用列表方式**同时统计**出各个班级的各门课程的成绩信息，怎么办呢？





# 分组汇总，一次对同一列上的多个相同值做分组再汇总统计

	scorecode	plancode	stuid	score	examtype
1	1	1	130001	80.0	正常
2	2	1	130002	52.0	正常
3	3	1	130003	0.0	缓考
4	4	1	130004	91.0	正常
5	5	1	130005	88.0	正常
6	6	1	130006	27.0	正常
7	7	1	130007	78.0	正常
8	9	2	130001	75.0	正常
9	13	2	130002	60.0	补考
10	14	2	130003	82.0	正常
11	17	2	130004	92.0	正常
12	18	2	130005	97.0	正常
13	20	2	130006	76.0	正常
14	21	2	130007	67.0	正常
15	22	3	130008	60.0	补考
16	23	3	130009	71.0	正常
17	24	3	130010	0.0	缺考

统计数量

统计数量

统计数量

# 以课程号作为分组的依据

```
select plancode as 班级课程编号 from  
T_StuSore group by plancode
```

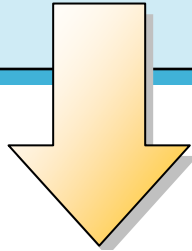


	班级课程编号
1	1
2	2
3	3
4	6

在分组的基础上，对分组记录进行数据的汇总

## 分组查询—GROUP BY （2）

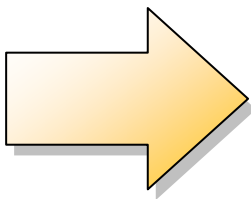
```
select  
plancode as 排课编号,  
AVG(score) as 考试平均分  
from T_StuScore  
group by plancode
```



	排课编号	考试平均分
1	1	59.428571
2	2	78.428571
3	3	63.142857
4	6	74.571428

# 分组查询—GROUP BY(续)

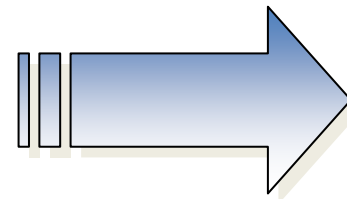
```
select plancode as 排课号, COUNT(*) as 参考人数, AVG(Score) as 平均分, MAX(Score) as 最高分, MIN(Score) as 最低分
from T_StuScore
group by plancode
```



	排课编号	参考人数	考试平均分	最低分	最高分
1	1	7	59.428571	0.0	91.0
2	2	7	78.428571	60.0	97.0
3	3	7	63.142857	0.0	94.0
4	6	7	74.571428	43.0	98.0

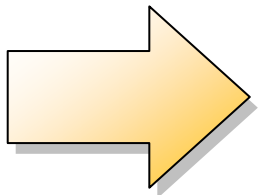
## 问题2: group by 可以对多列进行分组吗?

- 分组group by 不仅可以基于一列进行分组，还可以根据多列进行分组。



# 多列分组查询—GROUP BY(续)

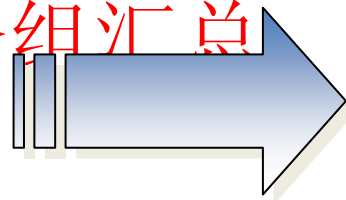
```
select  
plancode as 排课编号,examtype as 考试状态,  
COUNT(*) as 参考人数, AVG(score) as 考试平均分,  
MIN(score) as 最低分, MAX(score) as 最高分  
from T_StuScore  
group by plancode,examtype  
order by plancode,examtype
```



	排课编号	考试状态	参考人数	考试平均分	最低分	最高分
1	1	缓考	1	0.000000	0.0	0.0
2	1	正常	6	69.333333	27.0	91.0
3	2	补考	1	60.000000	60.0	60.0
4	2	正常	6	81.500000	67.0	97.0
5	3	补考	1	60.000000	60.0	60.0
6	3	缺考	1	0.000000	0.0	0.0
7	3	正常	5	76.400000	54.0	94.0
8	6	正常	7	74.571428	43.0	98.0

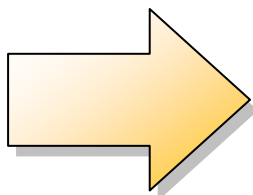
# 问题3: group by可以拥有where条件吗?

- 如何不将补考、缓考、缺考的成绩统计出来?
- 关键词: 先用where将正常考试的记录筛选出来, 再对选出的记录进行分组汇总



# 分组查询—GROUP BY(续)

```
select  
plancode as 排课编号, COUNT(*) as 参考人数,  
AVG(score) as 考试平均分, MIN(score) as 最低分,  
MAX(score) as 最高分  
from T_StuScore  
where examtype='正常'  
group by plancode
```



	排课编号	参考人数	考试平均分	最低分	最高分
1	1	6	69.333333	27.0	91.0
2	2	6	81.500000	67.0	97.0
3	3	5	76.400000	54.0	94.0
4	6	7	74.571428	43.0	98.0



## 问题3： 可以对group by分组后的信息进行再次筛选吗？

- 输出按班级排课课程编号分组后，考试成绩平均分小于70分的班级
- 关键词：先用group by分组并汇总出每门课程的考试平均分，再使用考试平均分来作为条件进行筛选

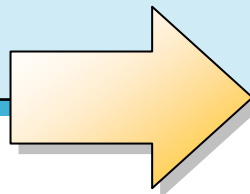
# 分组查询—GROUP BY(续)

按排课号分组统计出所有课程的平均分

```
select plancode as 排课编号,AVG(score) as 考试平均分  
from T_StuScore where examtype='正常'  
group by plancode
```

在分组的基础上查询平均成绩小于70分的课程

```
select plancode as 排课编号,AVG(score) as 考试平均分  
from T_StuScore where examtype='正常'  
group by plancode  
having AVG(score) <70
```

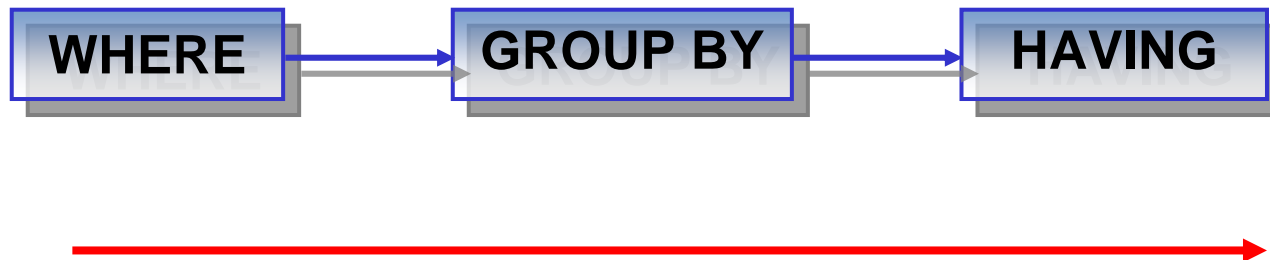


**HAVING** 中的条件列必须出现在**Group by** 或者聚合列中

	排课编号	考试平均分
1	1	69.333333

# 分组查询一对比

- ❑ **WHERE**子句从数据源中去掉不符合其搜索条件的数据
- ❑ **GROUP BY**子句在**WHERE**筛选出的数据中，根据分组列进行分组，并使用统计函数为各个分組组计算统计值
- ❑ **HAVING**子句去掉统计数据不符合要求的记录



执行顺序

# 总结

- DISTINCT
- 聚合函数：
- AVG
- COUNT
- MAX
- MIN
- SUM
- GROUP BY
- Join on
- 视图