

数据库范式

数据库范式简介



数据库范式是关系型数据库中表的一种设计规则

什么是范式 (NF = NormalForm)

- 范式是符合某一种设计要求的总结

在数据库中表的设计，必须保证其合理性

- 数据库表的设计关系整个系统的架构，关系到后续的开发效率和运行效率

如何设计合理的数据表

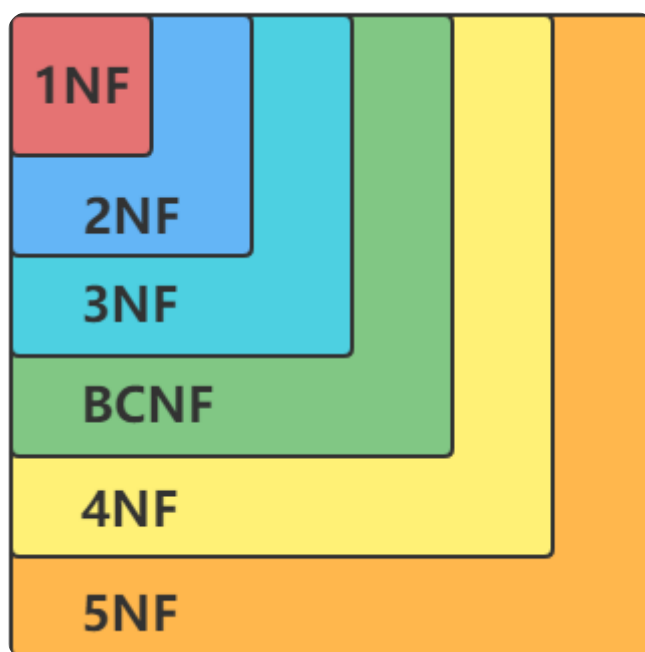
- 结构合理
- 冗余数据少
- 尽量避免插入删除修改异常
- 遵循一定的规则，在关系型数据库中这种规则就称为范式

关系型数据库有六种常见范式：

- 第一范式 (1NF)
- 第二范式 (2NF)
- 第三范式 (3NF)
- 巴斯-科德范式 (BCNF)
- 第四范式 (4NF)
- 第五范式 (5NF)

各个范式是依次嵌套包含的：

在第一范式的基础上进一步满足更多规范要求的称为第二范式 (2NF) ， 其余范式以此类推。



范式越高，设计质量越高，在现实设计中也越难实现

一般数据库设计，达到**第三范式**就足够了

数据库范式中的概念

元组：可以理解为一张表中的每条记录，也就是每一行数据。

属性：可以看作是“表的一列”。

- 主属性：
在一个关系中，如一个属性是构成某一个候选关键字的属性集中的
的一个属性，则称它为主属性。

例如：在关系——学生（学号，姓名，年龄，性别，班级）中，主属性是“学号”，那么其他的“姓名”、“年龄”、“性别”、“班级”都可以称为非主属性。

- 非主属性：

不包含在候选码中的属性称为非主属性，相对于主属性来定义的。

实时效果反馈

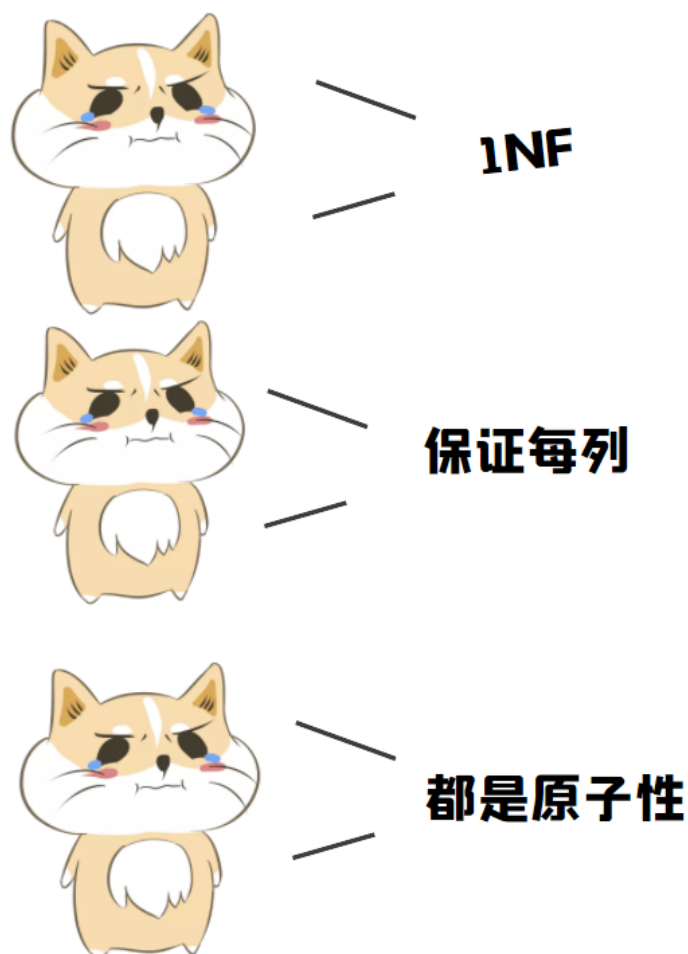
1. 什么是数据库范式？

- A** 是关系型数据库表的设计规则。
- B** 是关系型数据库中对象的使用规则
- C** 是关系型数据库中内存分配规则
- D** 是数据文件定义的规则

答案

1=>A

第一范式 (1NF)



- 最基本的范式，是其他范式的基础
- 数据库表每列都是不可分割的基本数据项，同一列中不能有多值，
确保每列保持原子性

学号	姓名	院系信息		课程名称	成绩
		系名	系主任		
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80

根据第一范式设计表

学号	姓名	系名	系主任	课程名称	成绩
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80

第一范式存在的问题

① 数据冗余

学号	姓名	系名	系主任	课程名称	成绩
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80

② 插入数据异常

学号	姓名	系名	系主任	课程名称	成绩
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80
		法律系	李莉	✗	

③ 修改数据复杂

学号	姓名	系名	系主任	课程名称	成绩
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80

4 删除异常

学号	姓名	系名	系主任	课程名称	成绩
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80

实时效果反馈

1. 第一范式要求确保每列保持

- A 结构性
- B 松散性
- C 原子性
- D 状态性

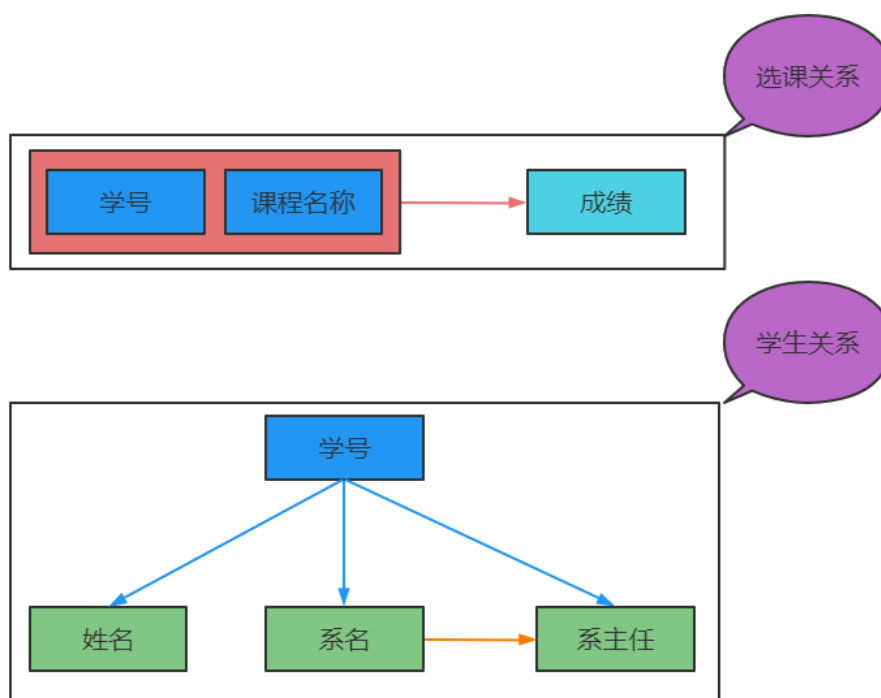
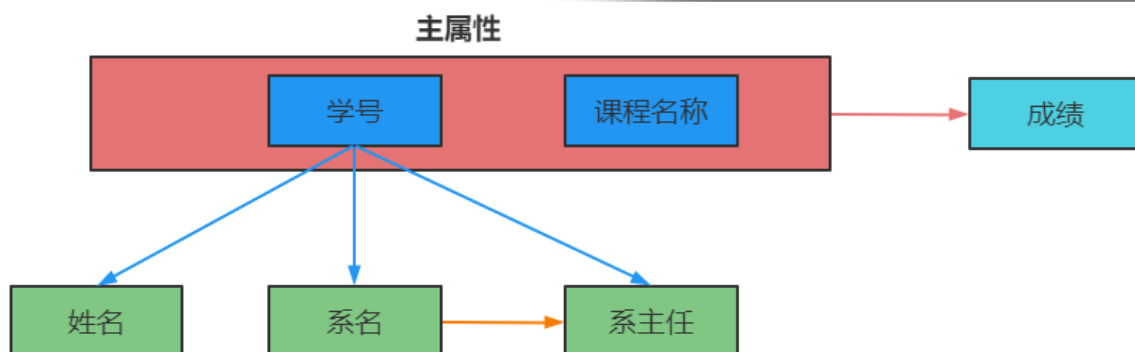
答案

第二范式 (2NF)



第二范式在第一范式的基础之上更进一层。第二范式需要确保数据库表中的每一列都和主键相关，而不能只与主键的某一部分相关（主要针对联合主键而言）。也就是说在一个数据库表中，一个表中只能保存一种数据，不可以把多种数据保存在同一张数据库表中。

学号	姓名	系名	系主任	课程名称	成绩
1221	海洋	信息工程系	王刚	高等数学	95
1221	海洋	信息工程系	王刚	大学英语	80
1221	海洋	信息工程系	王刚	计算机基础	88
1230	王强	信息工程系	王刚	高等数学	92
1230	王强	信息工程系	王刚	大学英语	87
1230	王强	信息工程系	王刚	计算机基础	76
1300	丽丽	经济系	张小明	高等数学	72
1300	丽丽	经济系	张小明	大学英语	98
1300	丽丽	经济系	张小明	普通化学	80



根据第二范式设计表

选课关系数据表

学号	课程名称	成绩
1221	高等数学	95
1221	大学英语	80
1221	计算机基础	88
1230	高等数学	92
1230	大学英语	87
1230	计算机基础	76
1300	高等数学	72
1300	大学英语	98
1300	普通化学	80

学生关系数据表

学号	姓名	系名	系主任
1221	海洋	信息工程系	王刚
1230	王强	信息工程系	王刚
1300	丽丽	经济系	张小明

第二范式存在问题

① 插入异常

学号	姓名	系名	系主任
1221	海洋	信息工程系	王刚
1230	王强	信息工程系	王刚
1300	丽丽	经济系	张小明
		法律系	李莉

② 删除异常

学号	姓名	系名	系主任
1300	丽丽	经济系	张小明

实时效果反馈

1.第二范式需要确保数据库表中的每一列都和__相关:

- A 主属性
- B 非主属性

答案

$1 \Rightarrow A$

第三范式 (3NF)

3NF

每一列数据都和主键直接相关



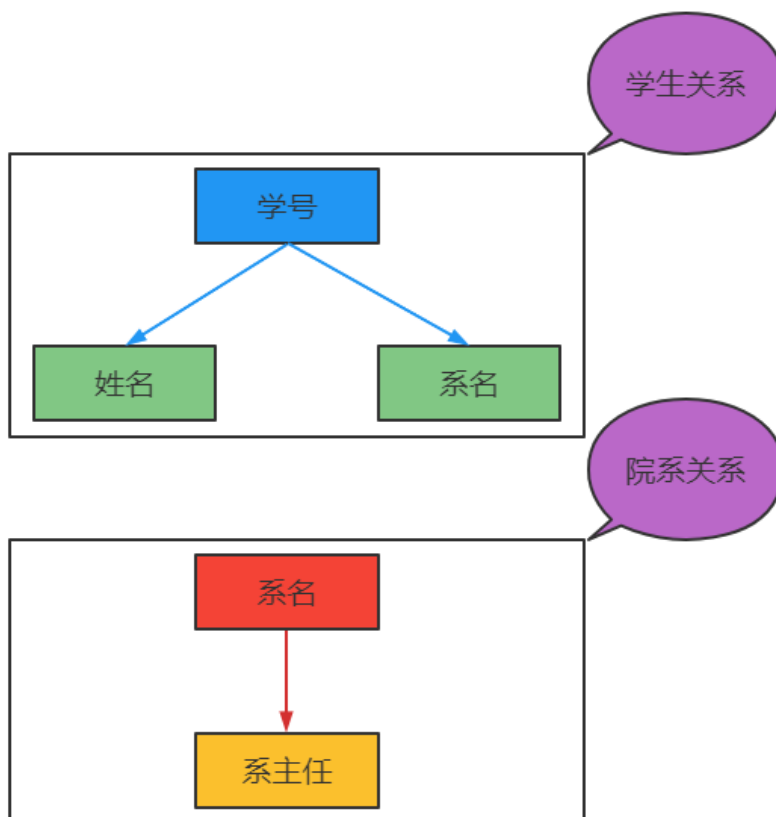
- 必须满足第二范式
- 确保数据表中的每一列数据都和**主键直接相关**，而不能间接相关

选课关系数据表

学号	课程名称	成绩
1221	高等数学	95
1221	大学英语	80
1221	计算机基础	88
1230	高等数学	92
1230	大学英语	87
1230	计算机基础	76
1300	高等数学	72
1300	大学英语	98
1300	普通化学	80

学生关系数据表

学号	姓名	系名	系主任
1221	海洋	信息工程系	王刚
1230	王强	信息工程系	王刚
1300	丽丽	经济系	张小明



根据第三范式设计表

选课关系数据表

学号	课程名称	成绩
1221	高等数学	95
1221	大学英语	80
1221	计算机基础	88
1230	高等数学	92
1230	大学英语	87
1230	计算机基础	76
1300	高等数学	72
1300	大学英语	98
1300	普通化学	80

学生关系数据表

学号	姓名	系名
1221	海洋	信息工程系
1230	王强	信息工程系
1300	丽丽	经济系

院系关系数据表

系名	系主任
信息工程系	王刚
经济系	张小明

符合3NF要求的数据库设计，基本上解决了数据冗余过大，插入异常，修改异常，删除异常的问题。¹²

实时效果反馈

1.第三范式需要确保数据库表中的每一列都和主属性__

- ☐ A 间接相关
- ☐ B 直接相关
- ☐ C 不相关
- ☐ D 部分相关

答案

1=>B

数据库设计范式总结

可以适当增加冗余 用于提高性能



数据库设计范式**优点**:

- 结构合理
- 冗余较小
- 尽量避免插入删除修改异常

数据库设计范式**缺点**:

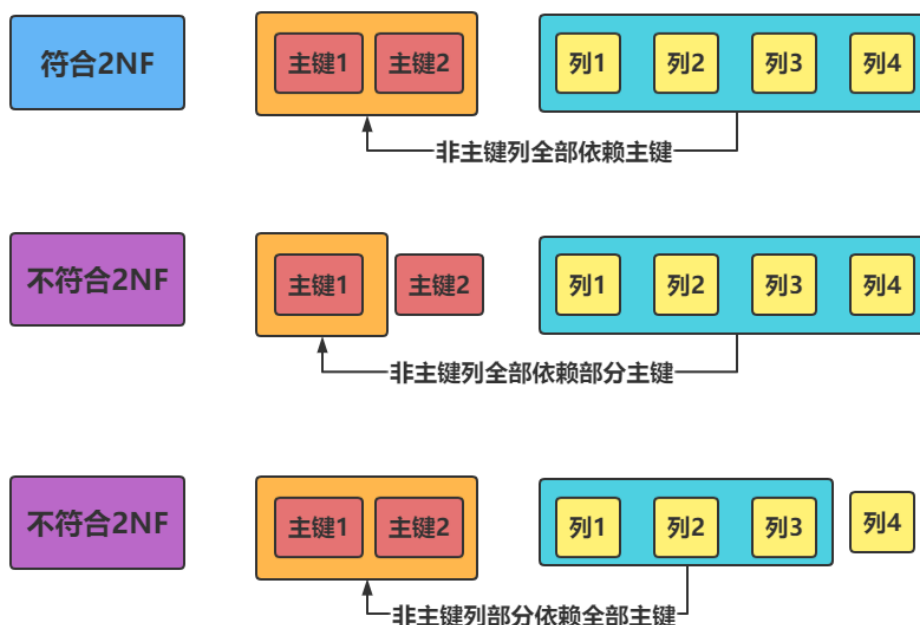
- 性能降低，多表查询比单表查询速度慢

在实际设计中，要整体遵循范式理论

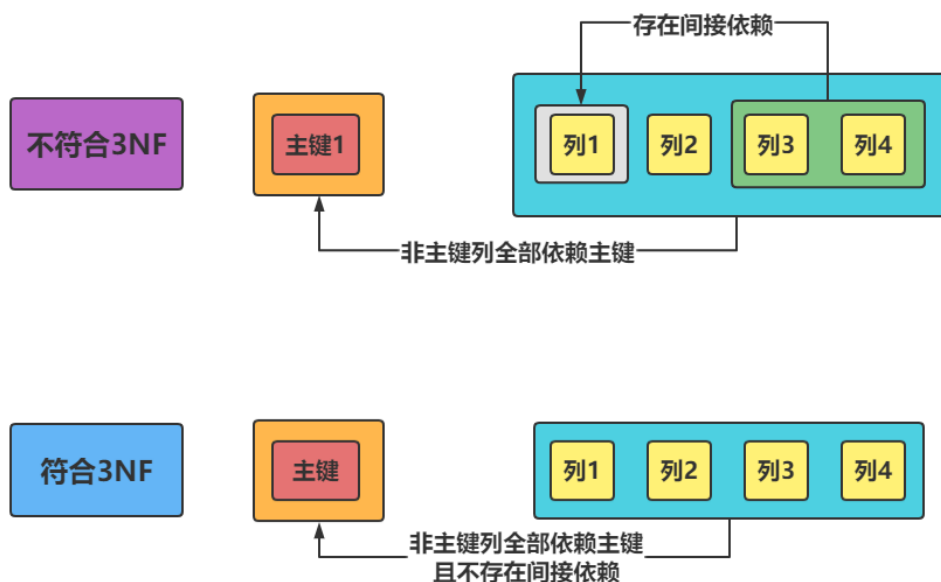
如果在某些特定的情况下还死死遵循范式也是不可取的，

因为可能降低数据库的效率，此时可以适当增加冗余而提高性能

- 第一范式（1NF）：字段不能再分
- 第二范式（2NF）：不存在部分依赖



- 第三范式 (3NF) : 不存在间接依赖



使用范式可以减少冗余，但是会降低性能

特定表的的设计可以违反第三范式，增加冗余提高性能



数据库里存放什么样的数据？

业务中的名词部分。

表关系简介

设计关系数据库的一个重要部分是将数据元素划分为相关的表，我们可以根据数据本身的关联性，将不同表之间的数据聚合在一起。

注意：

无论在表与表之间建立了什么样的关系，决定数据之间是否有关系的不是表，而是数据本身。

表与表的关系类型

表与表之间存在三种关系（relation），即一对一，一对多，多对多关系。

- **一对一**（one-to-one）：一种对象与另一种对象是一一对应关系，比如一个学生只能在一个班级里。
- **一对多**（one-to-many）：一种对象可以属于另一种对象的多个实例，比如一个班级里有很多学生。
- **多对多**（many-to-many）：两种对象彼此都是“一对多”关系，比如一个歌单里包含多首歌，同时一首歌可以属于多个歌单。

实时效果反馈

1. 决定数据之间是否有关系的是

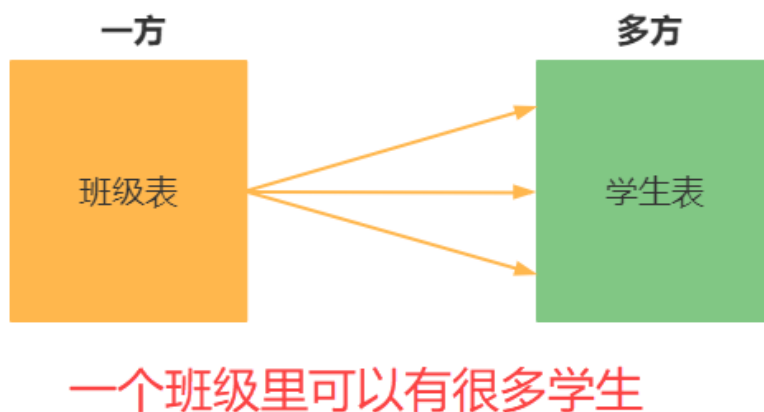
- ☒ A 表关系
- ☐ B 数据本身
- ☐ C SQL语句
- ☐ D 数据文件

答案

1=>B

一对多关系

一种对象可以属于另一种对象的多个实例。



一对多关系是建立在两张表之间的关系。一个表中的一条数据可以对应另一个表中的多条数据。

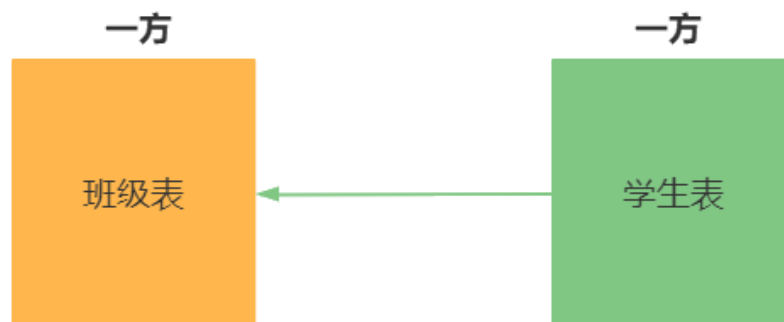
可以添加外键约束保证数据的参照完整性，外键永远在多方。外键允许重复，允许含有空值。

注意：

如果添加的外键约束，那么删除数据时，一定是先删除有外键一方的数据。

一对一关系

一种对象与另一种对象是一一对应关系。



一个学生只能在一个班级里

一对一关系是建立在一对多的基础之上，一个表中的一条数据只能对应另一个表中的一条数据。

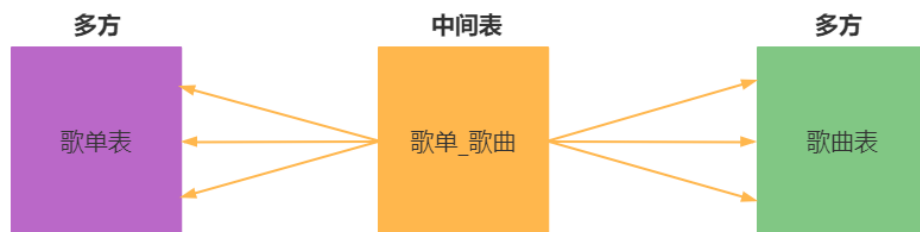
可以添加外键约束保证数据的参照完整性，外键可以在任何一方，需要让外键具备唯一约束。

注意：

如果添加的外键约束，那么删除数据时，一定是先删除有外键一方的数据。

多对多关系

两种对象彼此都是"一对多"关系



一个歌单包含多首歌，一首歌也可以属于多个歌单

需要建立一个中间表，中间表里建立两个列，然后需要用这两个列作为这个表的联合主键，然后每个列在作为外键参照各自的表的主键。

注意：

存数据：先两侧，再中间表。

删数据：先中间表（即有外键的表），再两侧。