

数据库设计三范式

- What?Why?How?



数据库设计概述

- 在设计和数据库有关的系统时，数据库表的设计至关重要，这些设计关系整个系统的架构，需要精心的仔细考虑。
- 数据库的设计主要包含了设计表结构和表之间的联系，在设计的过程中，有一些规则应该遵守



第一范式

- **第一范式(确保每列保持原子性)**
- 第一范式是最基本的范式。如果数据库表中的所有字段值都是不可分解的原子值,就说明该数据库表满足了第一范式。
- 第一范式的合理遵循需要根据系统的实际需求来定。比如某些数据库系统中需要用到“地址”这个属性,本来直接将“地址”属性设计成一个数据库表的字段就行。但是如果系统经常会访问“地址”属性中的“城市”部分,那么就非要将“地址”这个属性重新拆分为省份、城市、详细地址等多个部分进行存储,这样在对地址中某一部分操作的时候将非常方便。这样设计才算满足了数据库的第一范式,如下表所示。



第一范式

- 关系数据库的几种设计范式介绍
- 1、第一范式 (1NF)
- 在任何一个关系数据库中，第一范式 (1NF) 是对关系模式的基本要求，不满足第一范式 (1NF) 的数据库就不是关系数据库。
- 所谓第一范式 (1NF) 是指数据库表的每一列都是不可分割的基本数据项，同一列中不能有多值，即实体中的某个属性不能有多值或者不能有重复的属性。如果出现重复的属性，就可能需要定义一个新的实体，新的实体由重复的属性构成，新实体与原实体之间为一对多关系。在第一范式 (1NF) 中表的每一行只包含一个实例的信息。例如，对于图3-2中的员工信息表，**不能将员工信息都放在一列中显示**，也不能将其中的两列或多列在一列中显示；员工信息表的每一行只表示一个员工的信息，一个员工的信息在表中只出现一次。简而言之，第一范式就是无重复的列。



第一范式

用户信息表							
编号	姓名	性别	年龄	联系电话	省份	城市	详细地址
1	张红欣	男	26	0378-23459876	河南	开封	朝阳区新华路23号
2	李四平	女	32	0751-65432584	广州	广东	白云区天明路148号
3	刘志国	男	21	0371-87659852	河南	郑州	二七区大学路198号
4	郭小明	女	27	0371-62556789	河南	郑州	新郑市薛店北街218号

上表所示的用户信息遵循了第一范式的要求，这样在对用户使用城市进行分类的时候就非常方便，也提高了数据库的性能。



第一范式

- 所谓第一范式（1NF）是指数据库表的每一列都是不可分割的基本数据项，同一列中不能有多值，即实体中的某个属性不能有多值或者不能有重复的属性。如果出现重复的属性，就可能需要定义一个新的实体，新的实体由重复的属性构成，新实体与原实体之间为一对多关系。在第一范式（1NF）中表的每一行只包含一个实例的信息。
- 在任何一个关系数据库中，第一范式（1NF）是对关系模式的基本要求，不满足第一范式（1NF）的数据库就不是关系数据库。
- **列不可分。**



第二范式

- 第二范式(确保表中的每列都和主键相关)
- 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。



第二范式 (2NF)

- 第二范式 (2NF) 是在第一范式 (1NF) 的基础上建立起来的, 即满足第二范式 (2NF) 必须先满足第一范式 (1NF)。第二范式 (2NF) 要求数据库表中的每个实例或行必须可以被唯一地区分。为实现区分通常需要为表加上一个列, 以存储各个实例的唯一标识。如图3-2 员工信息表中加上了员工编号 (emp_id) 列, 因为每个员工的员工编号是唯一的, 因此每个员工可以被唯一区分。这个唯一属性列被称为主关键字或主键、主码。
- 第二范式 (2NF) 要求实体的属性完全依赖于主关键字。所谓完全依赖是指不能存在仅依赖主关键字一部分的属性, 如果存在, 那么这个属性和主关键字的这一部分应该分离出来形成一个新的实体, 新实体与原实体之间是一对多的关系。为实现区分通常需要为表加上一个列, 以存储各个实例的唯一标识。简而言之, 第二范式就是非主属性非部分依赖于主关键字



第二范式

学号	学生姓名	教师编号	教师姓名
023145	张三	988010	张老师
023146	李四	988010	张老师
023147	王五	988011	李老师
023145	张三	988011	李老师
023258	赵六	988010	张老师

- 存在多对多关系时只有一个字段作为主键是不够的。
- 这张表中的主键是学号和教师编号的组合
- 部分依赖，会产生冗余数据，需要分解表



第二范式

学号(主键)	学生姓名
023145	张三
023146	李四
023147	王五
023258	赵六

教师编号(主键)	教师姓名
988010	张老师
988011	李老师

数据不存在部分依赖的情况，我们就称之为符合第二范式

学号(主键)	教师编号(主键)
023145	988010
023146	988010
023147	988011
023145	988011
023258	988010



第二范式

- 第二范式 (2NF) 是在第一范式 (1NF) 的基础上建立起来的, 即满足第二范式 (2NF) 必须先满足第一范式 (1NF)。第二范式 (2NF) 要求数据库表中的每个实例或行必须可以被唯一的区分。为实现区分通常需要为表加上一个列, 以存储各个实例的唯一标识。要求实体的属性完全依赖于主关键字。
- **不能部分依赖**。即: 一张表存在组合主键时, 其他非主键字段不能部分依赖。



第三范式

- 满足第三范式 (3NF) 必须先满足第二范式 (2NF)。简而言之，**第三范式 (3NF) 要求一个数据库表中不包含已在其它表中已包含的非主关键字信息**。例如，存在一个部门信息表，其中每个部门有部门编号 (dept_id)、部门名称、部门简介等信息。那么员工信息表中列出部门编号后就不能再将部门名称、部门简介等与部门有关的信息再加入员工信息表中。如果不存在部门信息表，则根据第三范式 (3NF) 也应该构建它，否则就会有大量的数据冗余。简而言之，第三范式就是属性不依赖于其它非主属性



第三范式

- 第三范式需要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。
- 比如在设计一个订单数据表的时候，可以将客户编号作为一个外键和订单表建立相应的关系。而不可在订单表中添加关于客户其它信息（比如姓名、所属公司等）的字段。如下面这两个表所示的设计就是一个满足第三范式的数据库表。



例子

订单信息表

订单编号	订单项目	负责人	业务员	订单数量	客户编号
001	挖掘机	刘明	李东明	1台	1
002	冲击钻	李刚	霍新峰	8个	2
003	铲车	郭新一	艾美丽	2辆	1

客户信息表

客户编号	客户名称	所属公司	联系方式
1	李聪	五一建设	13253661015
2	刘新明	个体经营	13285746958



第三范式

学号(主键)	学生姓名	班级编号	班级名称	班级信息
023145	张三	987654	3班	特招班
023146	李四	987231	3班	特招班
023147	王五	987654	4班	普通班
...				
023258	赵六	987654	3班	特招班

班级名称和班级信息这一组值没有必要的重复出现，这些值并不直接依赖于学号（主键），

而是依赖于班级号，也就是说表中出现了**传递依赖**

解决这个问题的办法是把这些和学号（主键）没有直接联系的信息记录的新的表里面，即
班级表



第三范式

学号(主键)	学生姓名	班级编号
023145	张三	987654
023146	李四	987231
023147	王五	987654
...		
023258	赵六	987654

班级编号(主键)	班级名称	班级信息
987654	3班	特招班
987655	4班	普通班
...		



第三范式 (3NF)

- 满足第三范式 (3NF) 必须先满足第二范式 (2NF)。简而言之，**第三范式 (3NF) 要求一个数据库表中不包含已在其它表中已包含的非主关键字信息。**
- 在第二范式的基础上，数据表中如果不存在非关键字段对任一候选关键字段的传递函数依赖则符合第三范式。
- **不能存在传递依赖。即：除主键外，其他字段必须依赖主键。**



总结

- 第一范式
 - 列不可分
- 第二范式
 - 列必须直接依赖主键
- 第三范式
 - 传递依赖
 - 表里面的列不能出现其它表的非主键字段



表依赖关系

- 一对一
- 一对多
- 多对多
- -----
- 用户表
- 角色表

