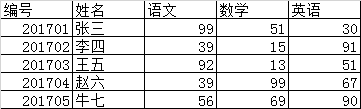
# 数据库

## 关系型数据库

当前在成熟应用且服务与各种系统的主力数据库还是关系型数据库。



代表：Oracle、SQL Server、MySQL

## 非关系型数据库

随着时代的进步与发展的需要，非关系型数据库应运而生。

代表：Redis、Mongodb

NoSQL数据库在存储速度与灵活性方面有优势，也常用于缓存。

## 数据库规范化

经过一系列的步骤，我们现在终于将客户的需求转换为数据表并确立这些表之间的关系，那么是否我们现在就可以在开发中使用呢？答案否定的，为什么呢！同一个项目，很多人参与了需求的分析，数据库的设计，不同的人具有不同的想法，不同的部门具有不同的业务需求，我们以此设计的数据库将不可避免的包含大量相同的数据，在结构上也有可能产生冲突，在开发中造成不便。

## 范式

要设计规范化的数据库，就要求我们根据数据库设计范式――也就是数据库设计的规范原则来做。范式可以指导我们更好地设计数据库的表结构，减少冗余的数据，借此可以提高数据库的存储效率，数据完整性和可扩展性。

设计关系数据库时，遵从不同的规范要求，设计出合理的关系型数据库，这些不同的规范要求被称为不同的范式，各种范式呈递次规范，越高的范式数据库冗余越小。目前关系数据库有六种范式：第一范式（1NF）、第二范式（2NF）、第三范式（3NF）、巴德斯科范式（BCNF）、第四范式(4NF）和第五范式（5NF，又称完美范式）。满足最低要求的范式是第一范式（1NF）。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），其余范式以次类推。一般说来，数据库只需满足第三范式(3NF）就行了。

## 三大范式

### 第一范式（1NF）

所谓第一范式（1NF）是指在关系模型中，对列添加的一个规范要求，所有的列都应该是原子性的，即数据库表的每一列都是不可分割的原子数据项，而不能是集合，数组，记录等非原子数据项。即实体中的某个属性有多个值时，必须拆分为不同的属性。在符合第一范式（1NF）表中的每个域值只能是实体的一个属性或一个属性的一部分。简而言之，第一范式就是无重复的域。

例如:表1-1中，其中”工程地址”列还可以细分为省份，城市等。在国外，更多的程序把”姓名”列也分成2列,即”姓”和“名”。

虽然第一范式要求各列要保存原子性，不能再分，但是这种要求和我们的需求是相关联的，如上表中我们对”工程地址”没有省份，城市这样方面的查询和应用需求，则不需拆分，”姓名”列也是同样如此。

表1-1   原始表



### 第二范式（2NF）

在1NF的基础上，非Key属性必须完全依赖于主键。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记录必须可以被唯一地区分。选取一个能区分每个实体的属性或属性组，作为实体的唯一标识。

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，第二范式就是在第一范式的基础上属性完全依赖于主键。

例如:表1-1中，一个表描述了工程信息，员工信息等。这样就造成了大量数据的重复。按照第二范式，我们可以将表1-1拆分成表1-2和表1-3：

l  工程信息表：（工程编号，工程名称，工程地址）：

表1-2   工程信息表



l  员工信息表（员工编号，员工名称，职务，薪资水平）：

表1-3   员工信息表



这样，表1-1就变成了两张表，每个表只描述一件事，清晰明了。

### 第三范式（3NF）

第三范式是在第二范式基础上，更进一层，第三范式的目标就是确保表中各列与主键列直接相关，而不是间接相关。即各列与主键列都是一种直接依赖关系，则满足第三范式。

第三范式要求各列与主键列直接相关，我们可以这样理解，假设张三是李四的兵，王五则是张三的兵，这时王五是不是李四的兵呢?从这个关系中我们可以看出，王五也是李四的兵，因为王五依赖于张三，而张三是李四的兵，所以王五也是。这中间就存在一种间接依赖的关系而非我们第三范式中强调的直接依赖。

现在我们来看看在第二范式的讲解中，我们将表1-1拆分成了两张表。这两个表是否符合第三范式呢。在员工信息表中包含：”员工编号”、”员工名称”、”职务”、”薪资水平”，而我们知道，薪资水平是有职务决定，这里”薪资水平”通过”职务”与员工相关，则不符合第三范式。我们需要将员工信息表进一步拆分，如下:

l  员工信息表：员工编号，员工名称，职务

l  职务表：职务编号，职务名称，薪资水平

现在我们已经了解了数据库规范化设计的三大范式，下面我们再来看看对表1-1优化后的数据表：

员工信息表(Employee)



工程信息表(ProjectInfo)



职务表(Duty)



工程参与人员记录表（Project\_ Employee\_info）

