

# 《交通大数据技术》



2024/6/14

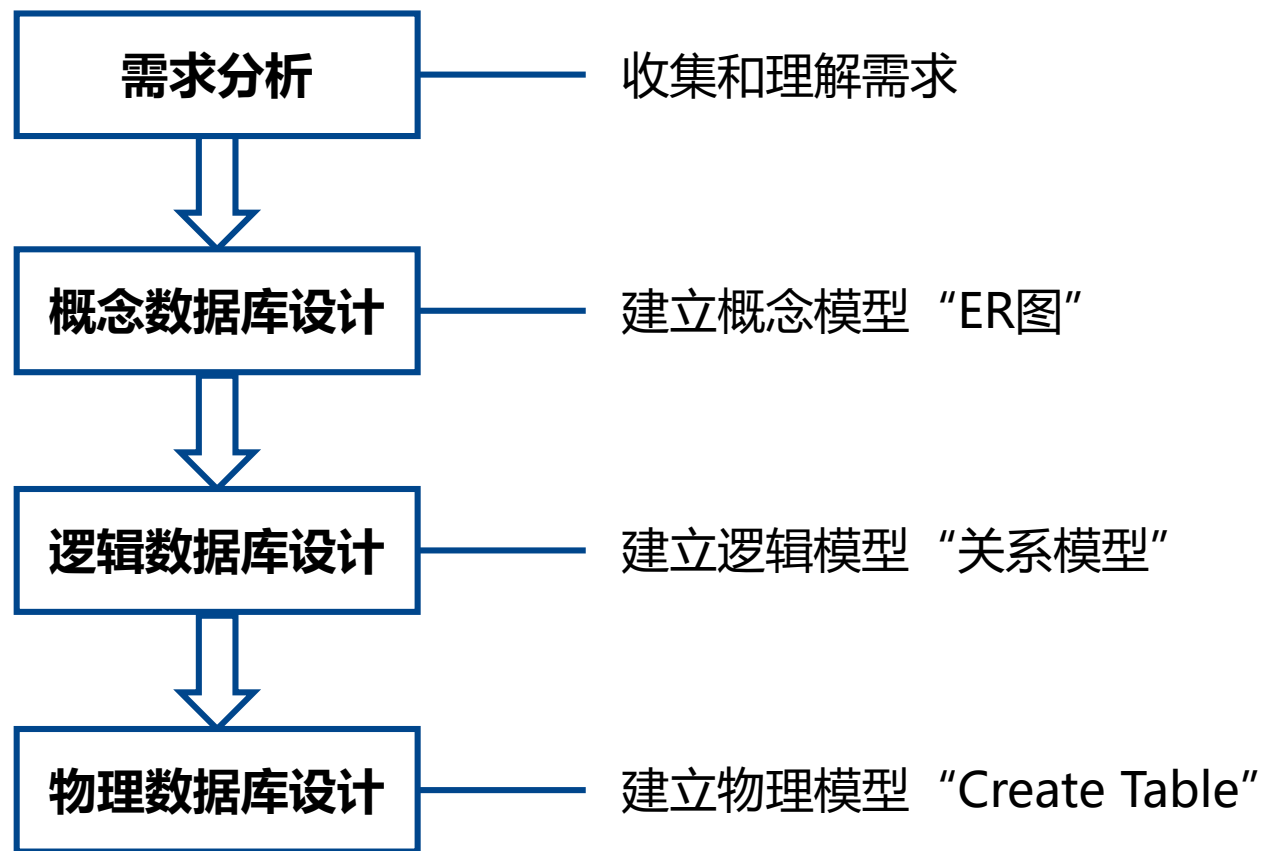
交通大数据技术

**马晓磊**  
**交通科学与工程学院**  
**2024年春季学期**

# E/R图 (Entity/Relationship)

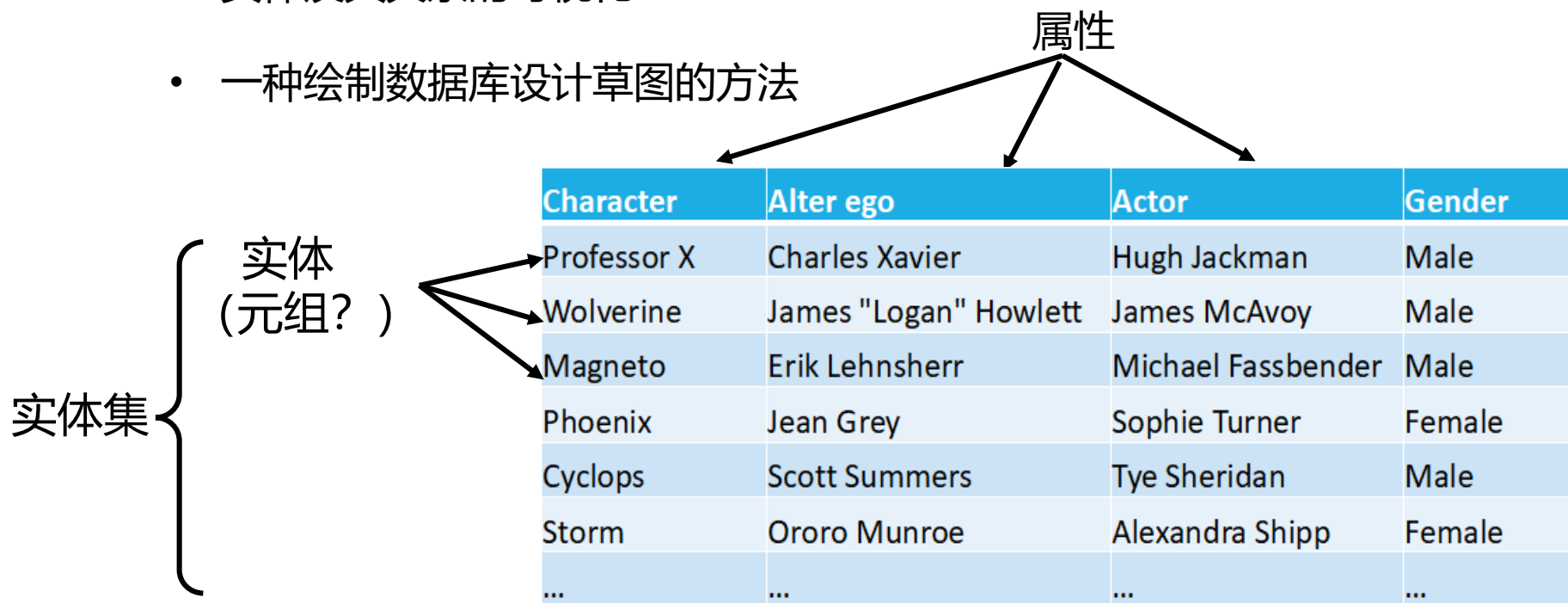


# 数据库设计概述



## 什么是E/R 图?

- 一种数据建模技术
- 实体及其关系的可视化
- 一种绘制数据库设计草图的方法



- 实体集：用矩形表示



- 属性：用椭圆表示

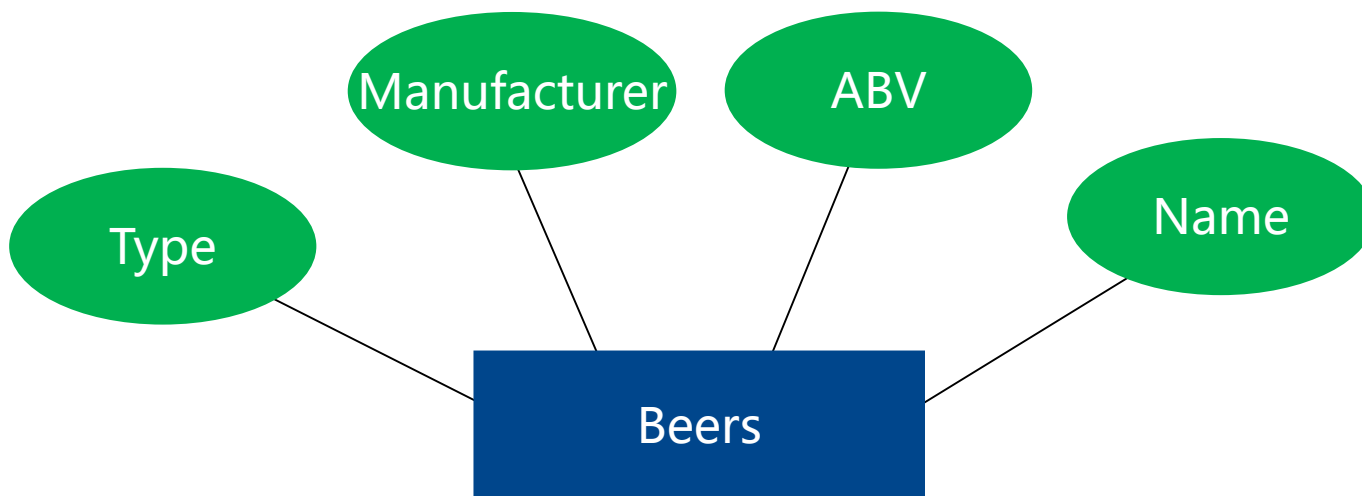


- 关系：用菱形表示



# E/R图中的属性

属性通过一条线连接到关联的实体集



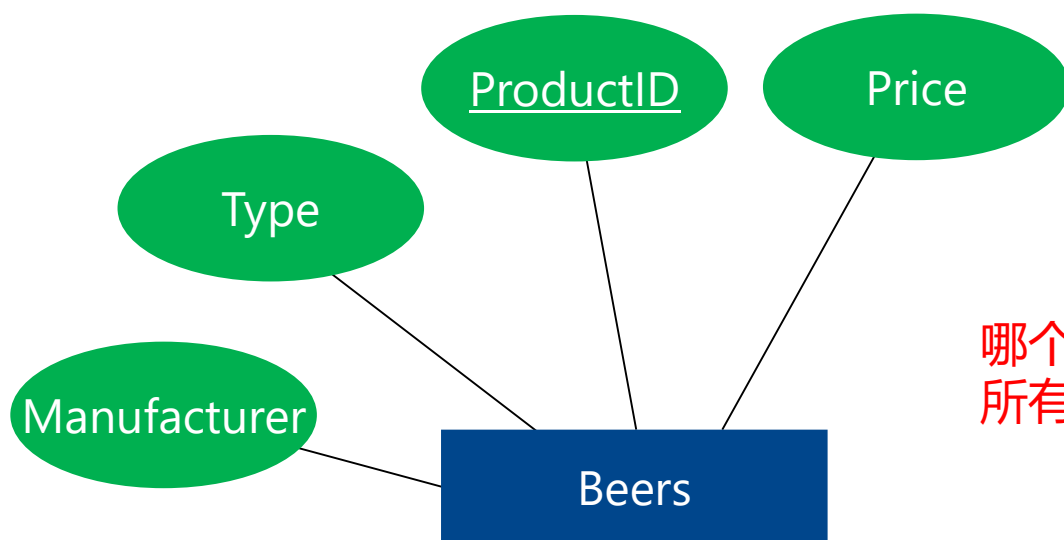
每个实体的每个属性都有一个值，例如：

Name	Manufacturer	Type	ABV
Old Rasputin	North Coast	Imperial Stout	9.0

# E/R图中的键

每个实体集都必须有一个键。

实体集合E的键是一个或多个属性的集合K，这样，给定E中任意两个不同的实体e1和e2，e1与e2不能具有键K中每个属性的相同值。



哪个属性集可以作为班级中所有学生的键？

# 关系

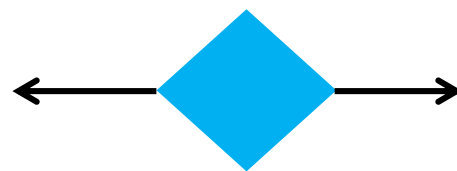
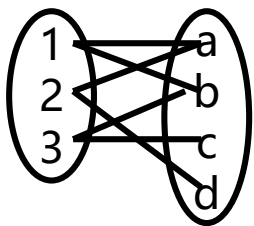
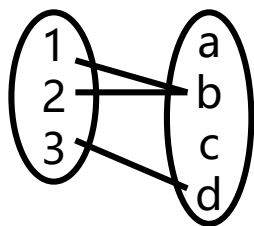
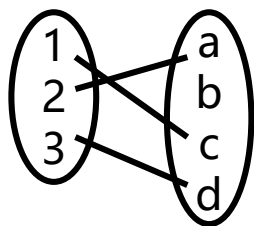
关系是两个或多个实体集之间的连接。

- 一对一

特殊情况  
↓

- 多对一

- 多对多

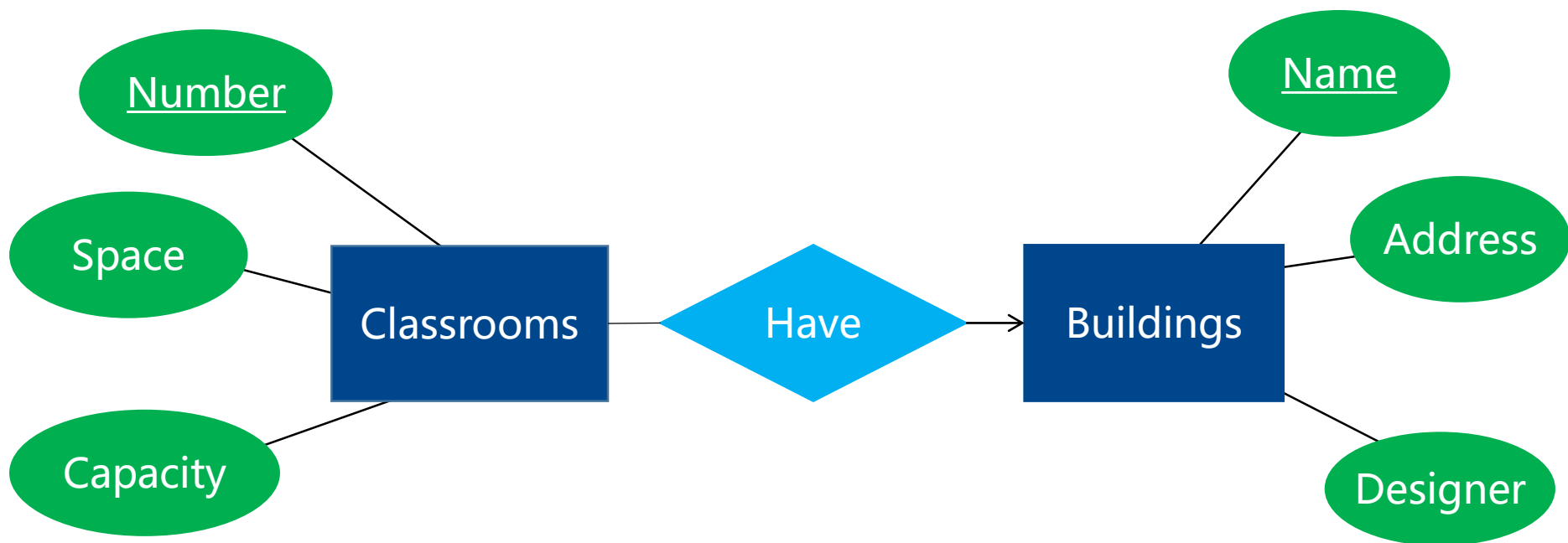




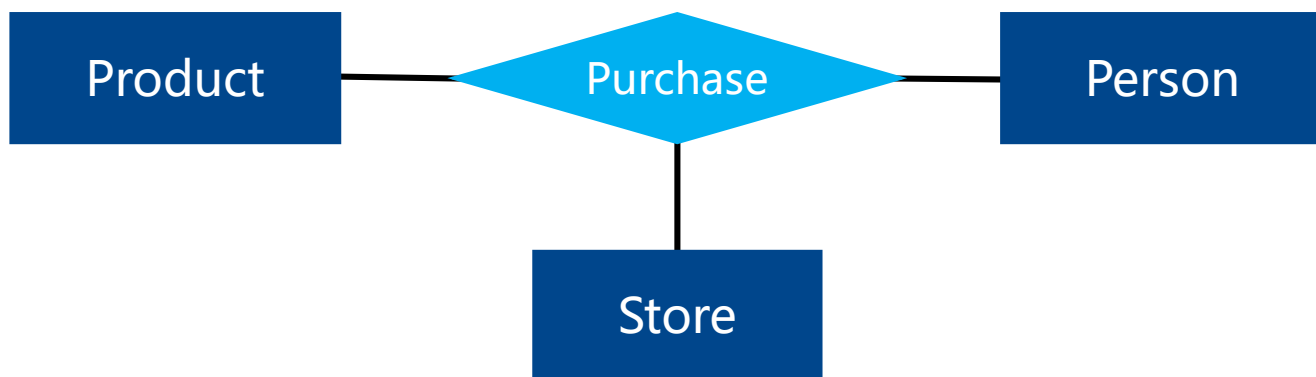
在ER图中，表示实体间多对多关系的正确方式是什么？

- ☐ A 直接将两个实体方框连接
- ☐ B 使用带有两端箭头的直线连接实体
- ☒ C 使用一个菱形连接两个实体，并标明关系名称
- ☐ D 在两个实体间画一个双向箭头

提交



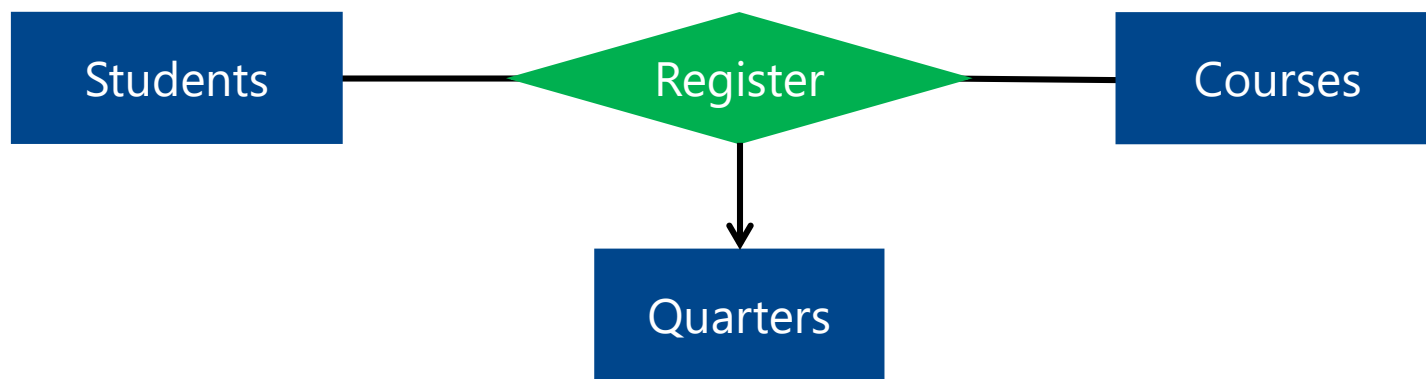
建立buyers, products和stores之间的购买关系模型：



三种多对多关系：

需要所有三个实体的信息才能形成独特的组合

箭头代表什么意思？

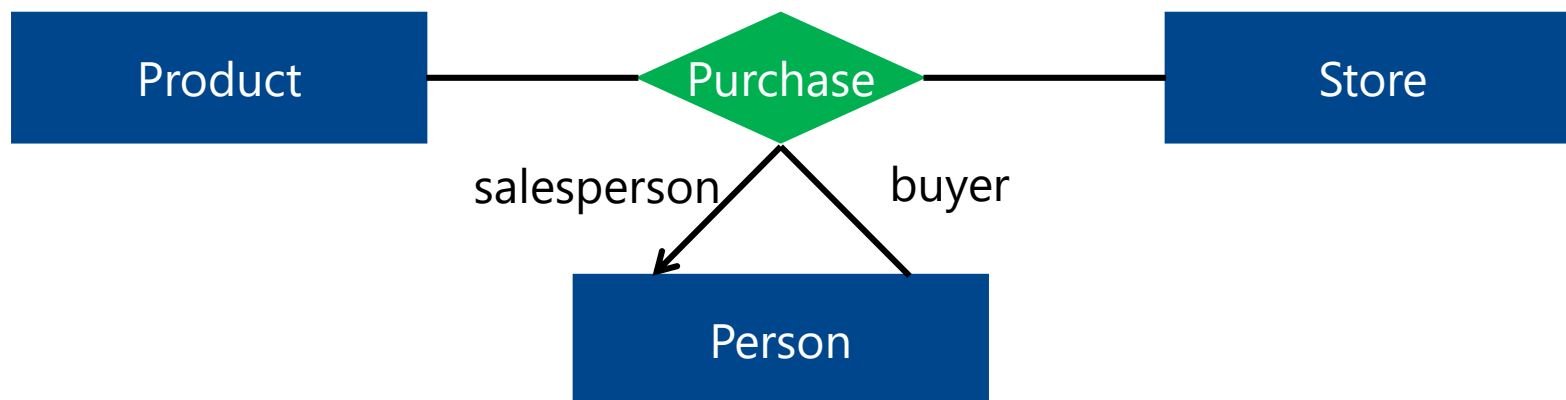


三重多对一关系：

**Student** 和 **Course** 的任何唯一组合将仅与唯一的 **Quarter** 相关联（假设不允许重修）。

# 关系中的角色

如果我们需要一个实体在一个关系中设置两次呢？



递归关系：当实体集出现在关系中时，我们从实体集中绘制尽可能多的线

存储信息，例如：

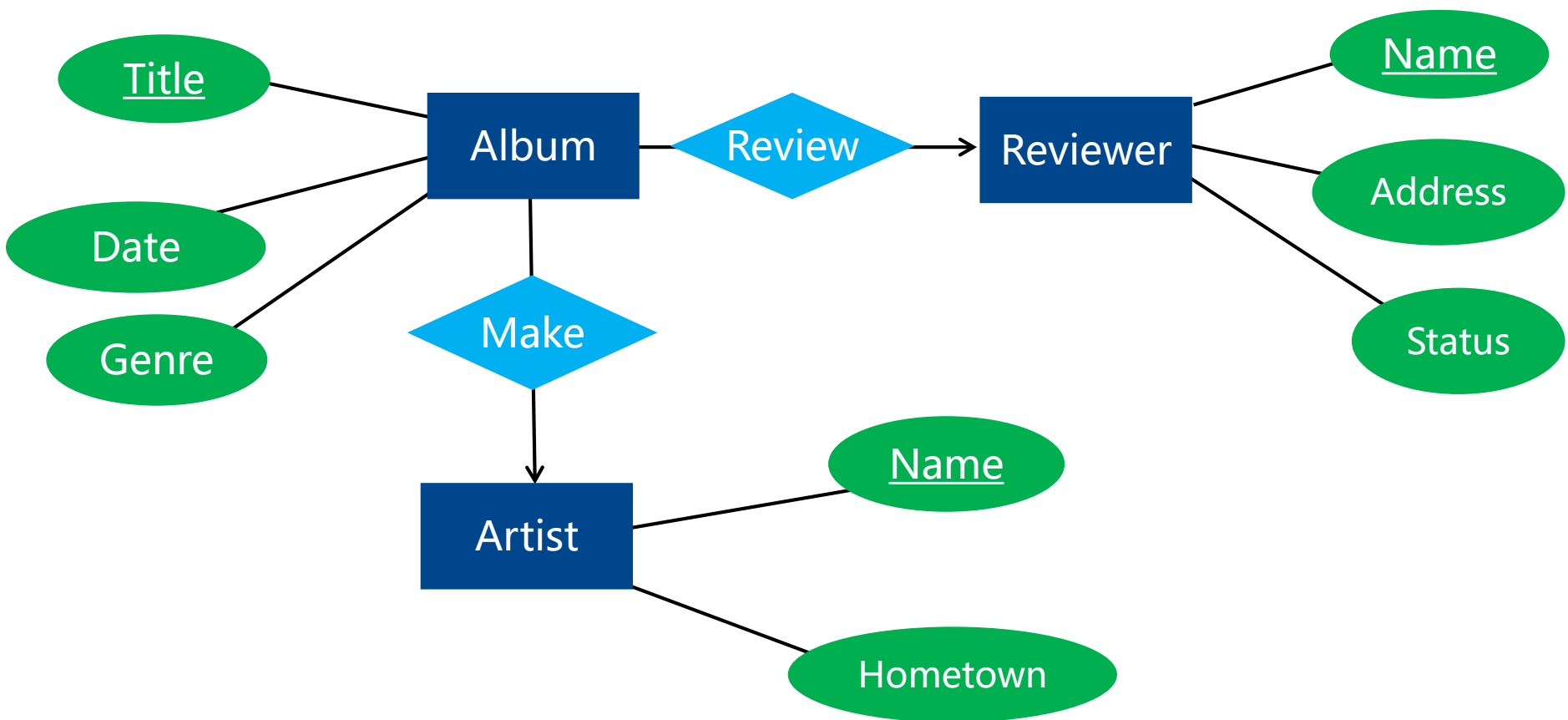
- Albums: name, date, genre, etc.
- Reviewers: name, status, address, etc.
- Artists: name, type, hometown, etc.

执行规则，包括：

- 每一张专辑都是由一个艺术家制作的，但是一个艺术家可以制作很多张专辑。
- 每一张专辑都可以由一个评审员评审

作答

# E/R图示例

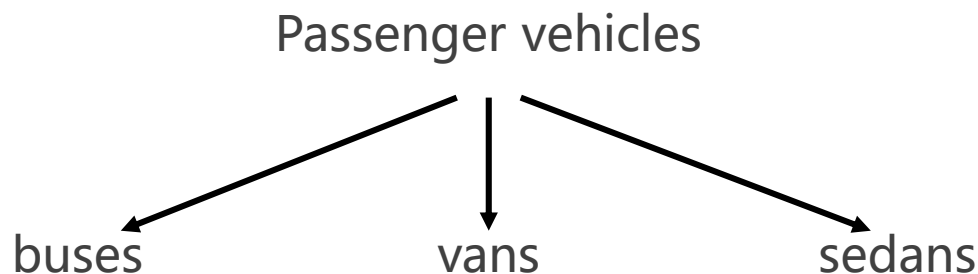


# E/R模型中的子类

世界本来就是等级森严的。有些实体是其他实体的特例。因此，在这种情况下，我们需要一个**子类**的概念来提高存储效率。

例如：Passenger vehicles

- 所有类型共享的许多公共属性（如Model, Make, Year）。
- 仅与某些类型相关的附加字段（例如，bus的 license class.）





# E/R模型中的子类

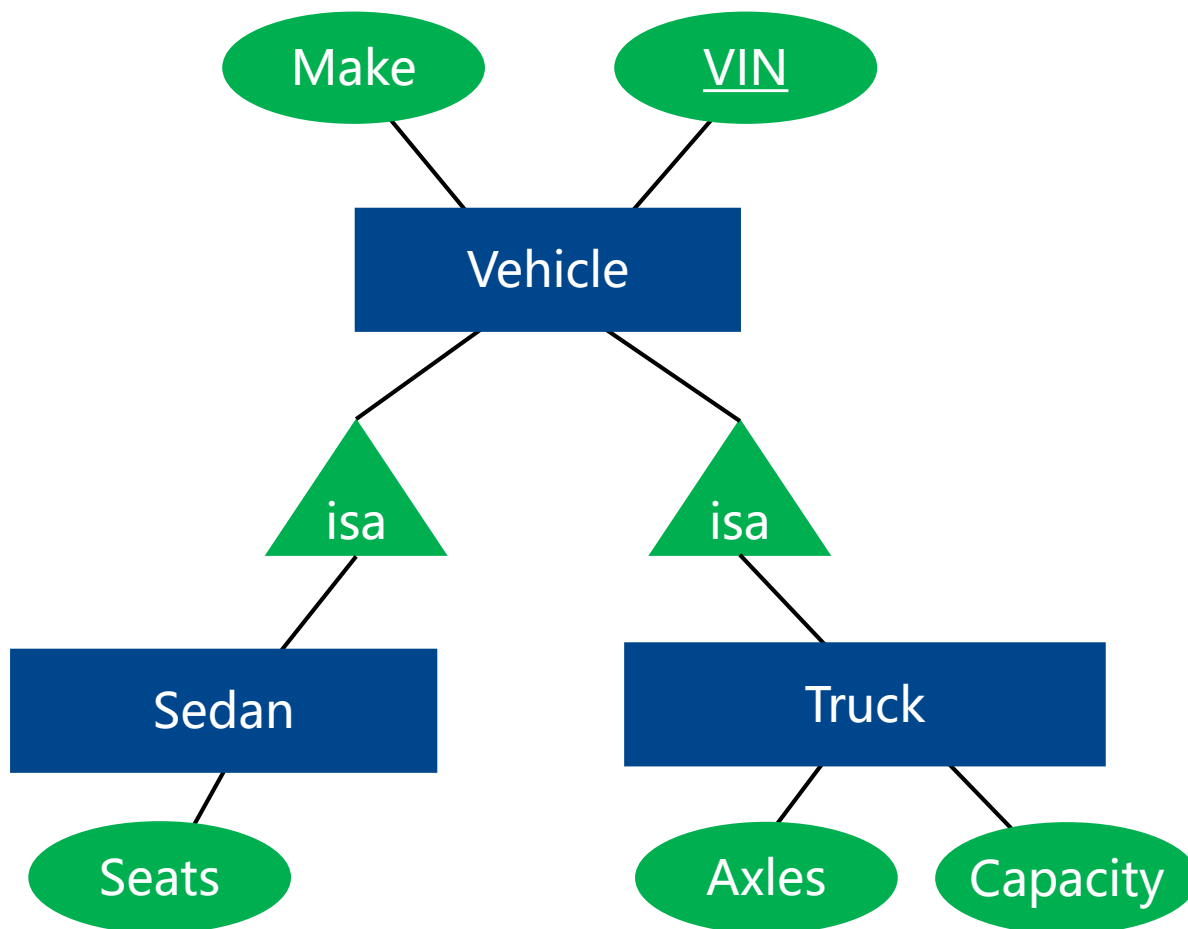
子类由  表示

连接到顶点的实体是父实体。

连接到底部的实体是子实体。

子实体继承其父属性，包括任何键，它也可能有自己的键。

# E/R模型中的子类



# E/R模型中的子类

## 从记录的角度考虑：

我们使用子类是因为某些类型的信息不会与所有子类相关

例如，这里我们要存储**Trucks**的**Axles**和**Capacity**，但这不适用于**Sedans**和其他**Vehicles**。

- Vehicle

Field 1	Field 2
VIN	Make

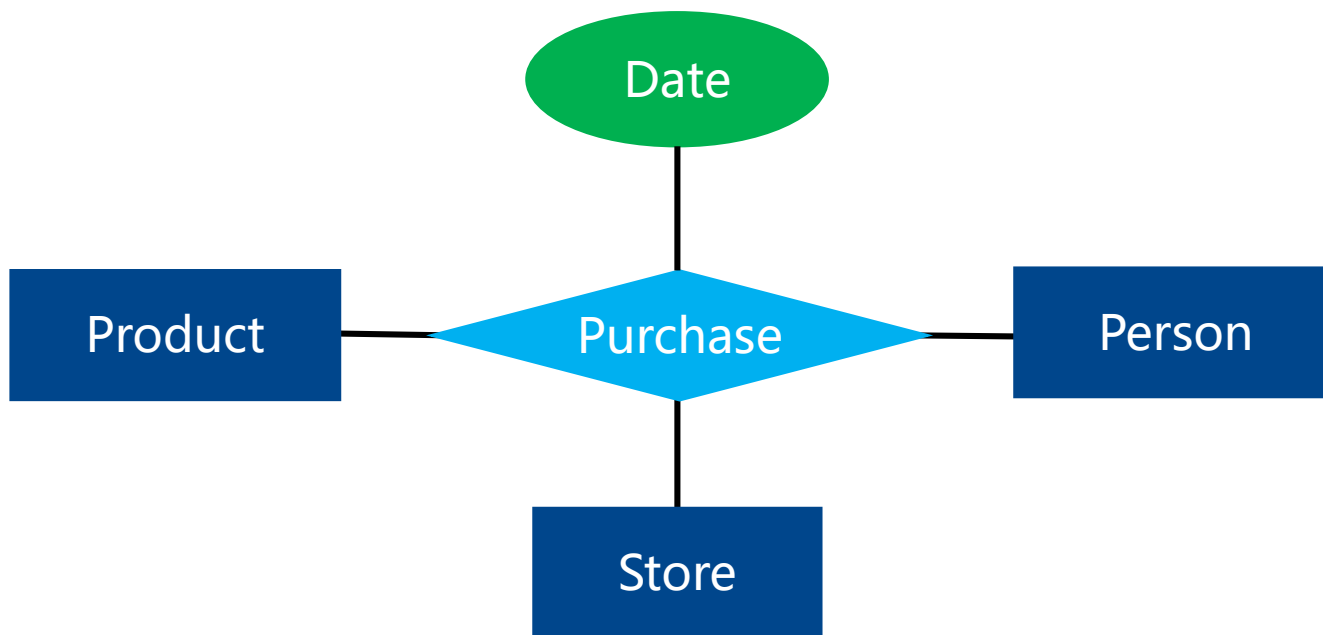
- Sedan

Field 1	Field 2	Field 3
VIN	Make	Seats

- Truck

Field 1	Field 2	Field 4	Field 5
VIN	Make	Axles	Capacity

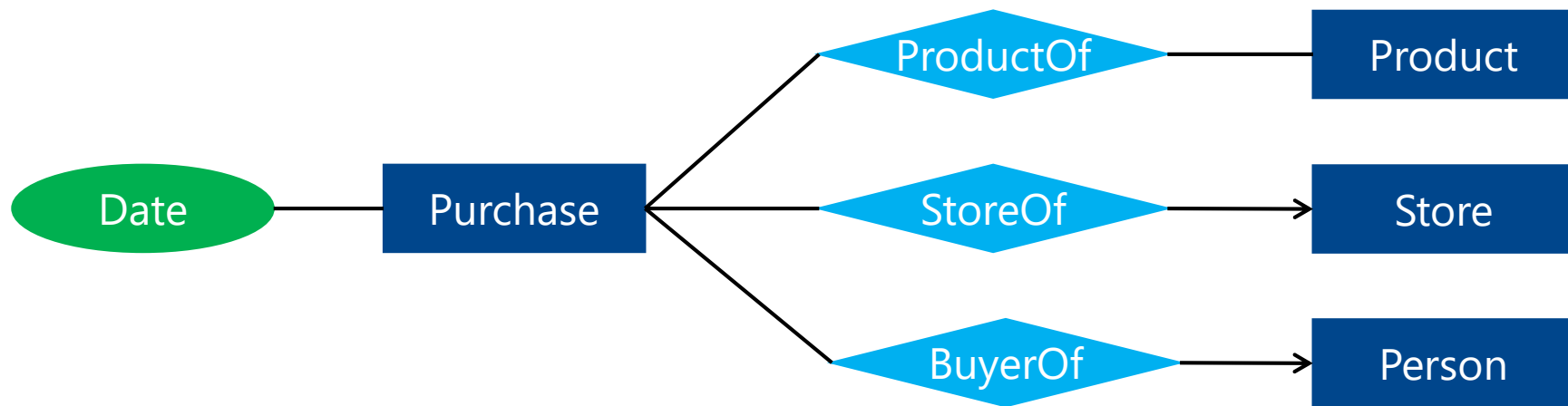
**Date**仅定义为采购事件的描述符。



这是个好办法吗？

- 关系上的属性很容易使图表混淆。
- 一些数据建模方法（如ODL）不支持多路关系。

# 将多关系转换为二项



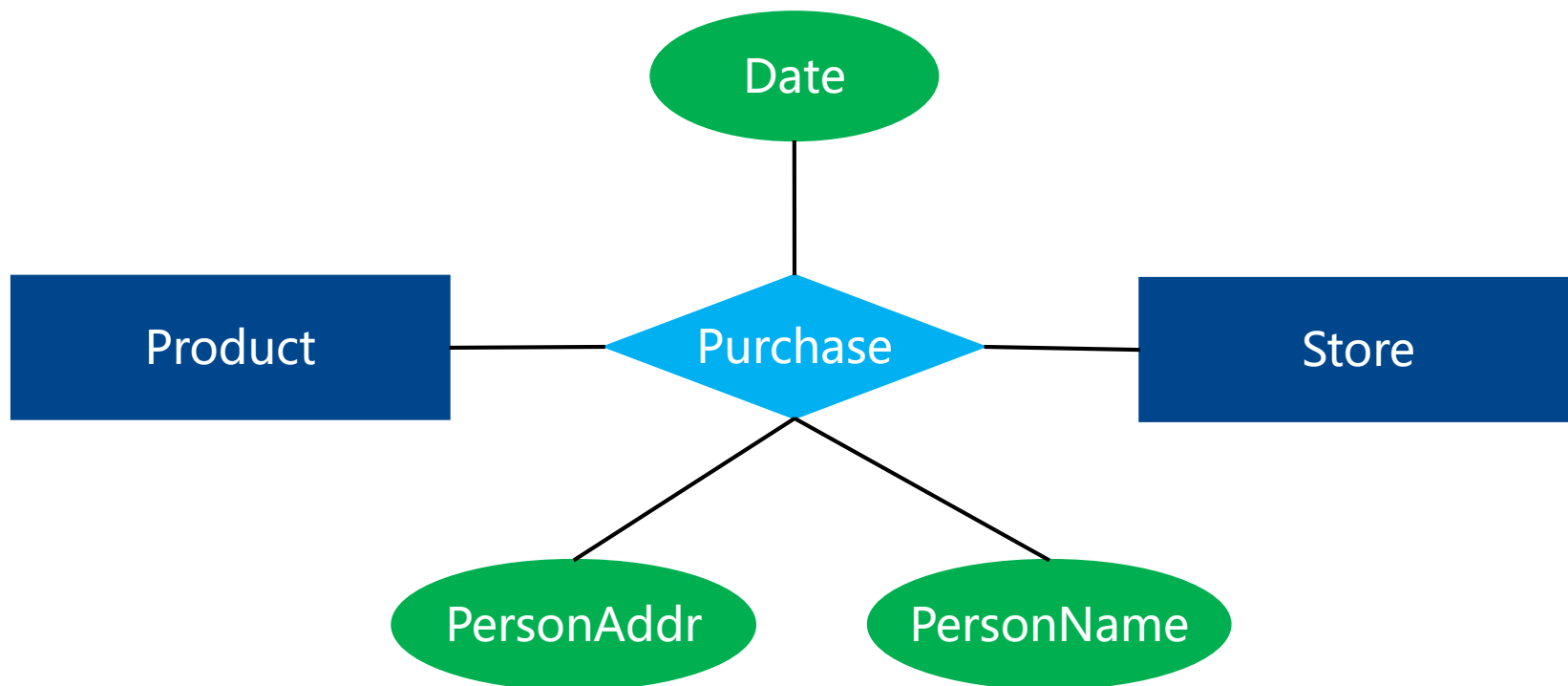
## 步骤

1. 创建一个新的实体集，命名类似于关系
2. 在新实体集和多路关系中存在的实体集之间创建关系

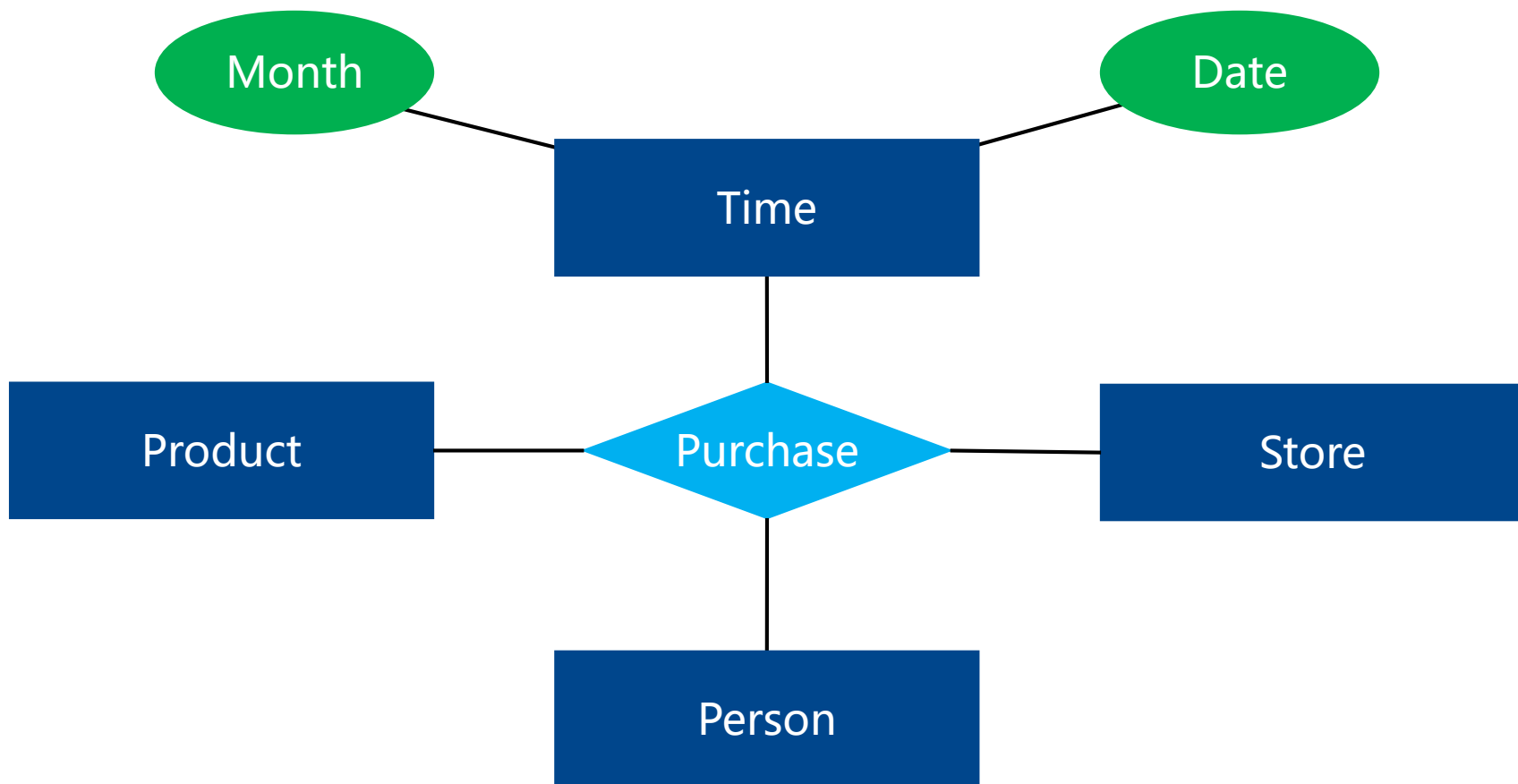
# 设计原则

1. 真实的！实体集及其属性应反映现实。
2. 避免冗余。我们应该小心，每件事只说一次。
3. 选择正确的关系。不必要的关系，尽管它们可能是真的，但会导致冗余和复杂性。
4. 选择合适的元素类型。它是实体集、关系还是属性？

# 设计原则：有错误吗？



# 设计原则：有错误吗？





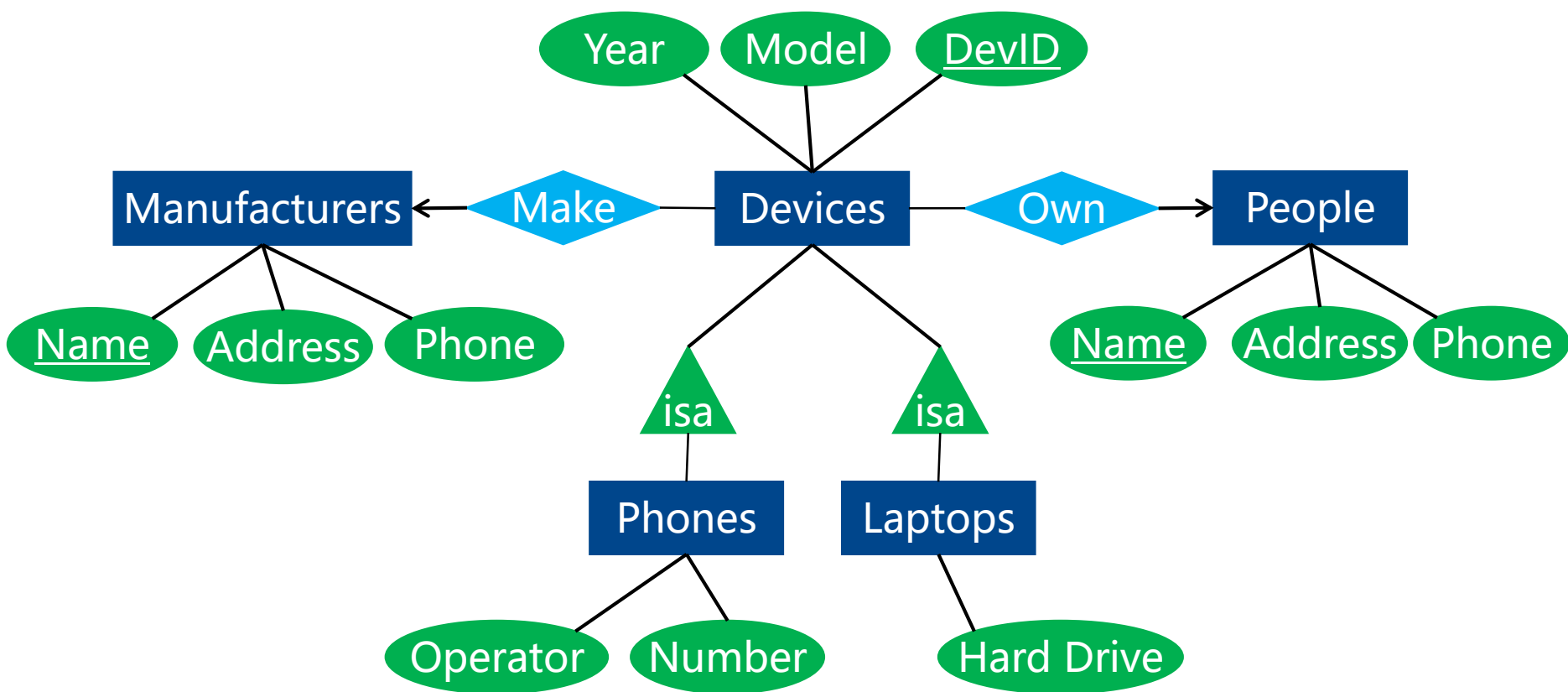
**假设我想建立一个数据库，记录这个房间里所有人的电子设备，包括：**

- 设备说明
- 制造商和所有者信息
- 其他信息？

**使用E/R图绘制数据库设计：**

- 包括实体、属性和关系。
- 每个实体都应该有一个键。
- 仅在必要时使用多关系和子类

# 示例



## 一些常用的约束

### 键约束:

- 属性或属性集在关系表中必须具有唯一且不为空的值。
- 示例：一个课程必须具有唯一的SLN。

### 单值约束:

- 实体集的每个属性都有一个值。
- 例如，一个人只能有一个生日。

# E/R图中的约束

## 实体的键是否受单个值约束？

- 一定地。键是单值约束的主要来源。

多对一关系意味着单个值约束。



## 单值约束和键约束有何不同？

- 非键单值属性可以采用null值（键不能为null）

# E/R图中的约束

## 域约束：

- 每个属性X的值必须是X域中的原子值。
- 示例：人名不能是数字。

## 一般限制条件：

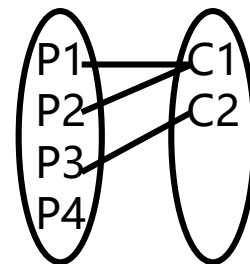
- 在数据库中应保留的其他任意约束，如检查约束和范围约束。
- 例如：州名，一个人的年龄。

## 引用完整性约束：

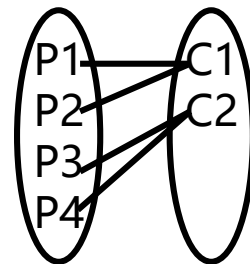
- 要求某个对象引用的值实际存在于数据库中。
- 例如，如果你在UW学习课程，你的课程必须存在于UW课程数据库中

# 引用完整性约束

在下面的E/R图中引用完整性约束是强制的吗？



不是。为什么？下面的图表呢？

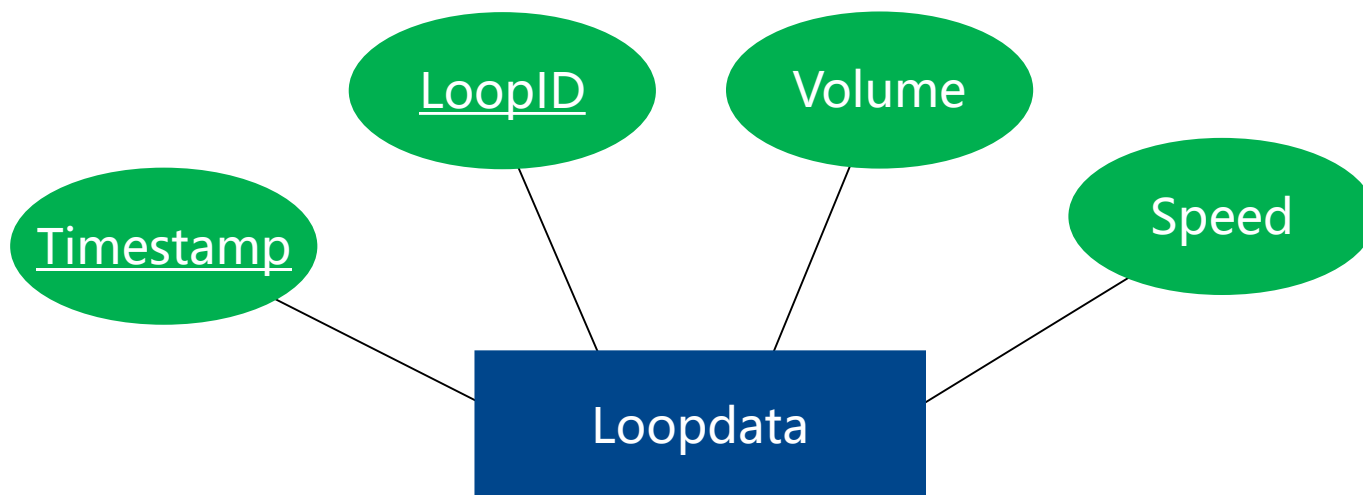


这一条确实是，作为一种产品的制造商，只有一家公司存在。

## 数据库中引用完整性约束的可能结果是什么？

- 禁止删除另一个表所需的实体。如果有学生已注册该课程，则无法删除该课程。
- 级联更新/删除：如果我从公司表中删除一个公司，那么与该公司关联的所有产品也将被删除

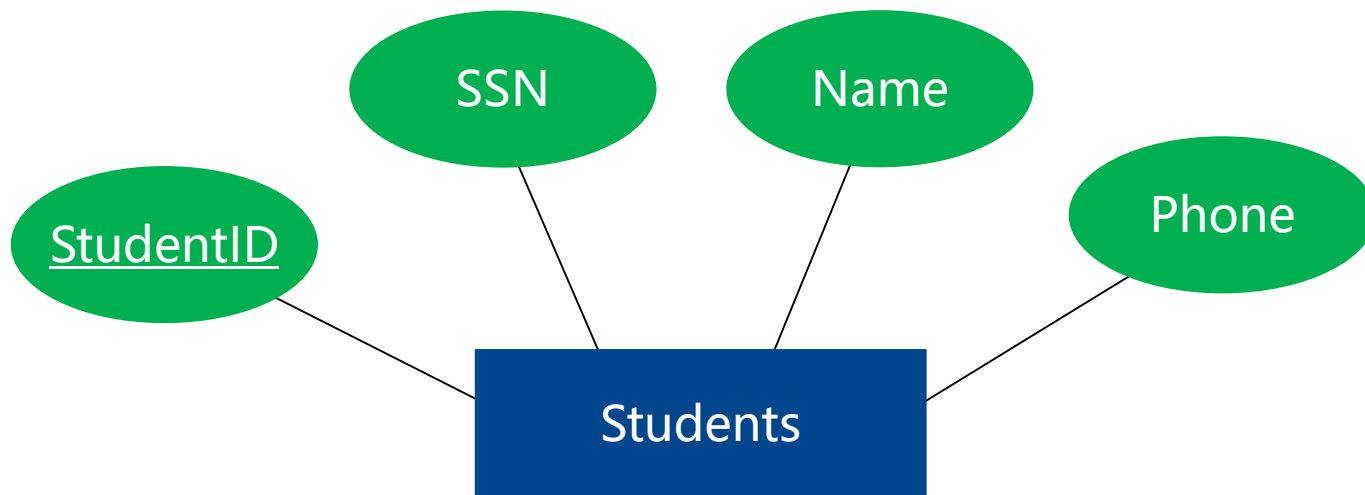
# E/R图中的键



**Timestamp** 和 **LoopID** 组合是唯一标识单个观测值所必需的  
如何表示：在构成键所需的所有属性下加下划线



# E/R图中的键



**StudentID** 和 **SSN** 都可以作为键.

如何在E/R图中指定多个键?

没有标准的方式! 我们得选一个。

当实体集的键来自与其相关的其他实体时，实体集是弱的。

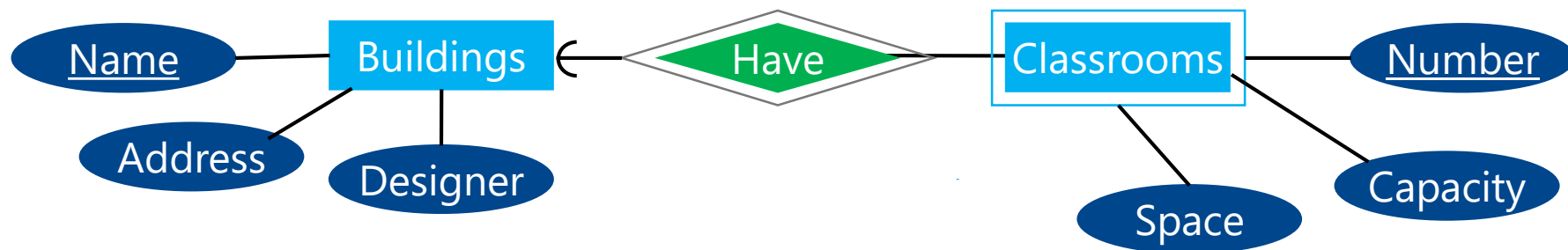
弱实体集的原因：

- 实体集根据与子类（如建筑物和教室）无关的分类划分为层次结构
- 将多路关系转换为二进制关系（例如，采购发票）

在ER图中：

- 弱实体集用**双边框矩形**显示。
- 弱实体的支持关系将显示为**双边界的菱形**。

# 弱实体集

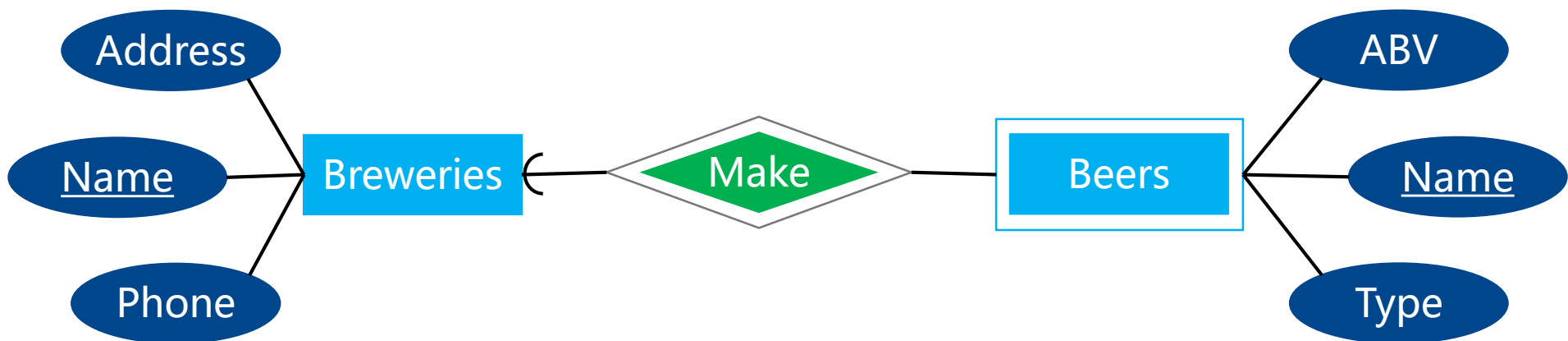


如果E是一个弱实体集，那么它的键包括:

- 零个或更多自身属性，以及
- 从E到其他实体集的某些多对一关系所达到的实体集的主属性。

弱实体集通常意味着多对一关系（包括作为特例的一对一关系）和引用完整性。

# 弱实体集



重要提示：不要过度使用弱实体集。

在多数情况下，引入一个属性作为键就足够了（例如，为**Beers**创建一些 ProductID）

在ER图中，如何正确表示一个强实体集和弱实体集之间的关系？

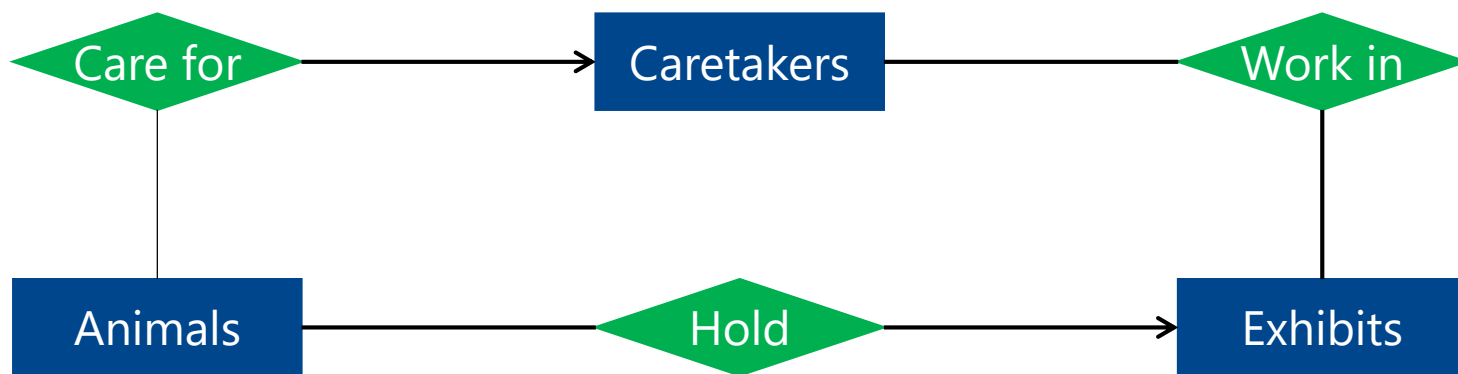
- ☐ A 使用单线的菱形连接它们
- ☒ B 弱实体集用双线的方框表示
- ☐ C 关系用带有“弱”标签的菱形表示
- ☒ D 使用一个带有双线边界的菱形连接它们，并且弱实体集显示部分依赖

提交

# E/R图回顾

- 每个实体集都应该有一个键。
- 如果有多个候选键，请选择一个。
- 不要添加不必要的关系

**哪种关系是不必要的？**

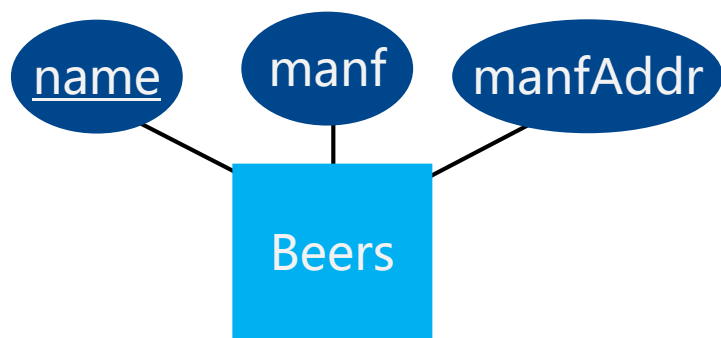


- 不要包含不必要的实体集。

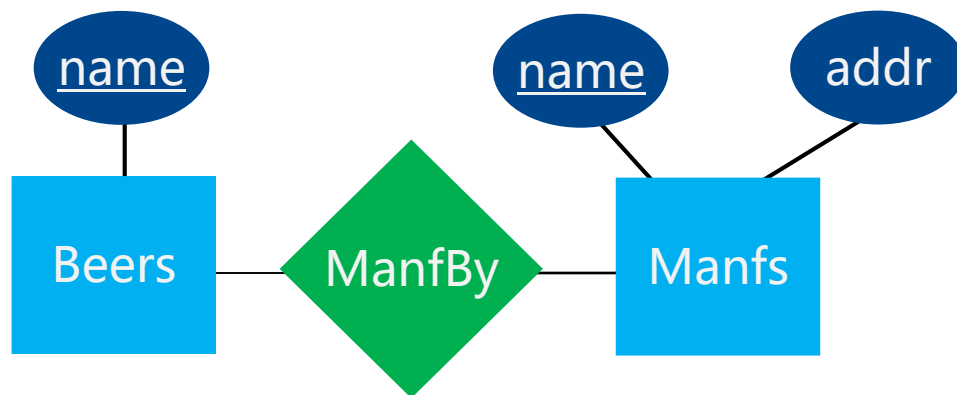


- 多对一关系中的箭头方向
- 双矩形和菱形表示弱实体集
- 正确分配属性，不要只在添加实体集时分配属性。
- 引用完整性是一个非常严格的约束----明智地使用
- 不要过于复杂！

# E/R图回顾



这种设计为每种啤酒重复一次制造商的地址；如果制造商暂时没有啤酒，地址就会丢失。



这种设计只给出每个制造商的地址一次。



## 慕课视频片段

视频名称：Video



温馨提示：此视频框在点击“上传手机课件”时会进行转换，用手机进行观看时则会变为可点击的视频。此视频框可被拖动移位和修改大小

**为表示北航的课程和注册系统的数据库绘制ER图。**

我们至少要代表以下内容：

- Students: StudentID, name, etc.
- Courses: SLN, name, section, etc.
- Departments: name, chair, etc.
- Buildings: name, address, etc.
- Classrooms: number, space, etc.
- Other information you are interested to save.

**你需要考虑的假设：**

- Classrooms 只能通过building name 和 classroom number唯一标识.
- 一门course 只能由一个特定的department提供
- 北航系统的其他假设

# E/R图回顾

