

商务智能——数据仓库设计

191850189王涛

1.关于某高校信息化建设环境的假设

2.主题域的划分与主题的设计

3.公共维度

4.总线矩阵

5.多维建模

5.1 生源质量分析

5.1.1 对于该分析应用业务的描述:

5.1.2 定义粒度

5.1.3 选取维度

5.1.4 选取度量值

5.1.5 多维建模设计

5.1.6 说明

5.2 图书馆借阅情况分析

5.2.1 对于该分析应用业务的描述:

5.2.2 定义粒度

5.2.3 选取维度

5.2.4 选取度量值

5.2.5 多维建模设计

5.2.6 说明

5.3 食堂销售情况分析

5.3.1 对于该分析应用业务的描述:

5.3.2 定义粒度

5.3.3 选取维度

5.3.4 选取度量值

5.3.5 多维建模设计

5.3.6 说明

5.4 受欢迎课程分析

5.4.1 对于该分析应用业务的描述:

5.4.2 定义粒度

5.4.3 选取维度

5.4.4 选取度量值

5.4.5 多维建模设计

5.4.6 说明

## 1.关于某高校信息化建设环境的假设

---

某高校（在下文的介绍中泛称“学校”，但全部代指该高校）为更好地为广大师生服务，决定对信息化建设进行升级，希望能够综合全校的相关数据，并以此为基础，在教务课程安排、考试安排、交换项目服务、财务缴费管理、图书馆管理、食堂就餐服务、宿舍管理、日常文娱活动管理、科研项目申报、招生等方面改善相关的业务与服务。

## 2.主题域的划分与主题的设计

---

为了让主题的设计更加全面合理，首先根据信息化建设的覆盖范围划分4个主题域：学生学业管理主题域、教职工服务主题域、财务管理主题域、后勤保障主题域，在不同的主题域中设计相应的主题。

### 一、主题设计中遵循的原则：

- 1.矩阵识别的是业务过程，而不是组织的业务部门
- 2.如果数据来源不同，功能不同，或者矩阵行代表的内容无法在单个迭代过程中合理完成，就应当创建独立的矩阵行
- 3.先考虑简单过程后考虑复杂过程，增量式构建

### 二、主题设计：

#### 1.学生学业管理主题域（共5个主题）：

招生、学生选课、学生交换项目、考试安排、学业反馈

#### 2.教职工服务主题域（共3个主题）：

教务课程安排、科研项目申报、职工招聘

#### 3.财务管理主题域（共4个主题）：

学业缴费、职工工资发放、教师工资发放、教师科研经费发放

#### 4.后勤保障主题域（共2个主题）：

学生图书借阅、食堂餐饮销售

## 3.公共维度

---

### 1.日期维度：

日期关键字（PK）、日期、日期完整描述、星期、日历月、日历年、财政年月、节假日指示符、周日指示符

### 2.院系维度：

院系关键字（PK）、院系代码、院系名称、院系所在书院、院系的科类（文/理/工/交叉）

### 3.学生维度：

学生关键字（PK）、入学年份、当前所在院系、原院系、学号、姓名、性别、身份证号、学制、联系电话、银行卡号

**说明：**学校允许学生进入大学学习后，根据兴趣转一次专业，即学生所在专业是可能变化的，但是变化的频率很低，所以学生维度是缓慢变化维，在这里选择使用类型3进行处理，原因在于通过增加属性，维护一个历史版本，在这里已经足够使用。

#### **4.课程维度：**

课程关键字（PK）、开课院系、任课教师、课程时间、考试时间、上课地点、学分

#### **5.生源维度**

生源学校关键字（PK）、生源学校所在省、生源学校所在市

#### **6.教师维度：**

教师关键字（PK）、所属院系、入职年份、教师编号、姓名、身份证号、职称、联系电话、银行卡号

#### **7.科研项目维度**

项目关键字（PK）、项目名称、项目级别

#### **8.职工维度：**

职工关键字（PK）、入职年份、入职部门、姓名、性别、身份证号、联系电话、银行卡号

#### **9.部门维度**

部门关键字（PK），部门名称

## **4.总线矩阵**

---

	日期 维度	院系 维度	学生 维度	课程 维度	生源 维度	教师 维度	科研项目 维度	职工 维度	部门 维度
招生	√	√			√			√	√
学生选课	√	√	√	√		√			
学生交换 项目	√	√	√	√					
考试安排	√	√	√	√		√			
学业反馈	√	√	√		√				
教务课程 安排	√	√	√	√		√			
科研项目 申报	√	√	√			√	√		
职工招聘	√	√						√	√
学业缴费	√		√						√
职工工资 发放	√		√					√	√
教师工资 发放	√	√				√			
教师科研 经费发放	√	√				√	√		
学生图书 借阅	√	√	√						
食堂餐饮 销售	√								

## 5.多维建模

## 5.1 生源质量分析

### 5.1.1 对于该分析应用业务的描述：

大一学年结束后，根据学生在校一学年的学业成绩、获奖情况等信息以及生源地信息对生源质量进行分析，找出各地学生的特点、差异等，据此有针对性地给予学生学业帮助，对生源地给出相应的反馈，以便更好地连接中学和大学教育。

### 5.1.2 定义粒度

每个学生（同届）大一学年的学业情况为一条记录

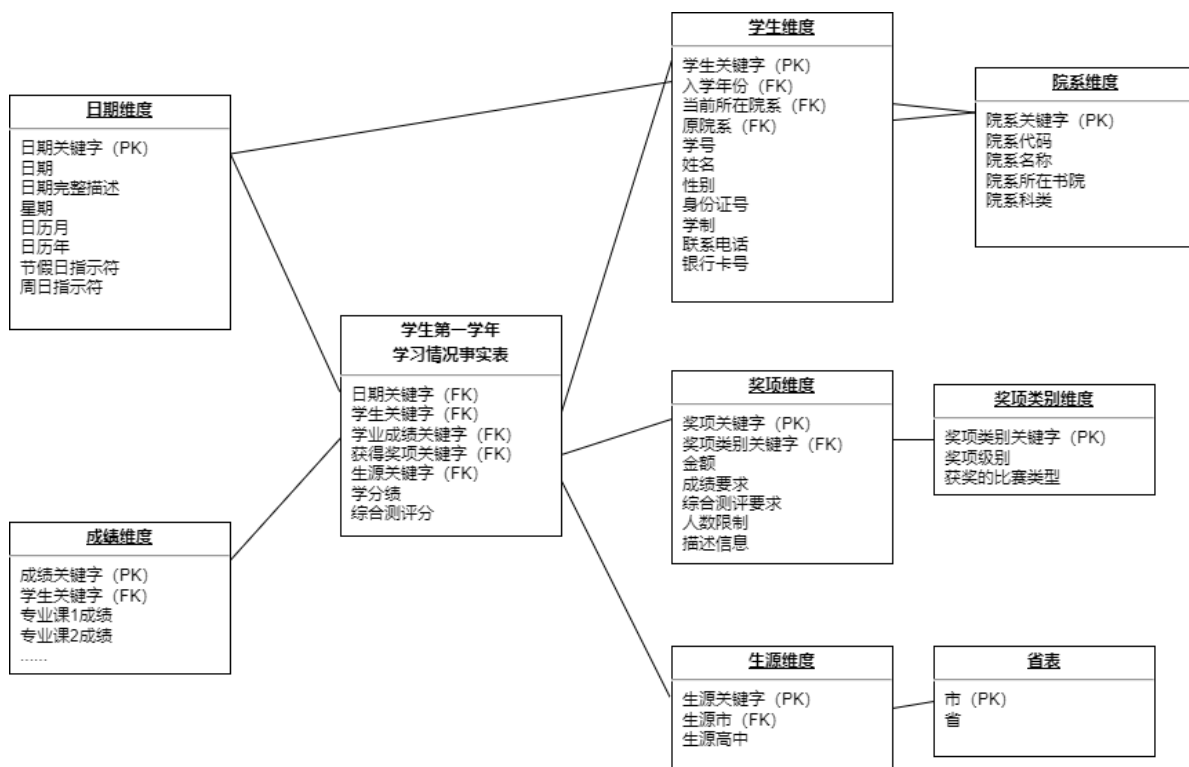
### 5.1.3 选取维度

日期、学生、学业成绩、获奖情况、生源

### 5.1.4 选取度量值

学分数、综合测评分

### 5.1.5 多维建模设计



### 5.1.6 说明

使用雪花模型，生源信息维度中省市以层级分开，奖项维度里，用外键连接奖项类别维度，使层次更加清晰，同时能够节省空间，减少数据冗余，使得数据仓库中的数据结构规范化。

## 5.2 图书馆借阅情况分析

### 5.2.1 对于该分析应用业务的描述：

通过对图书馆的借阅情况分析，分析比较受欢迎的图书类别以及各个院系同学的借阅情况。

### 5.2.2 定义粒度

学生的单条借阅记录

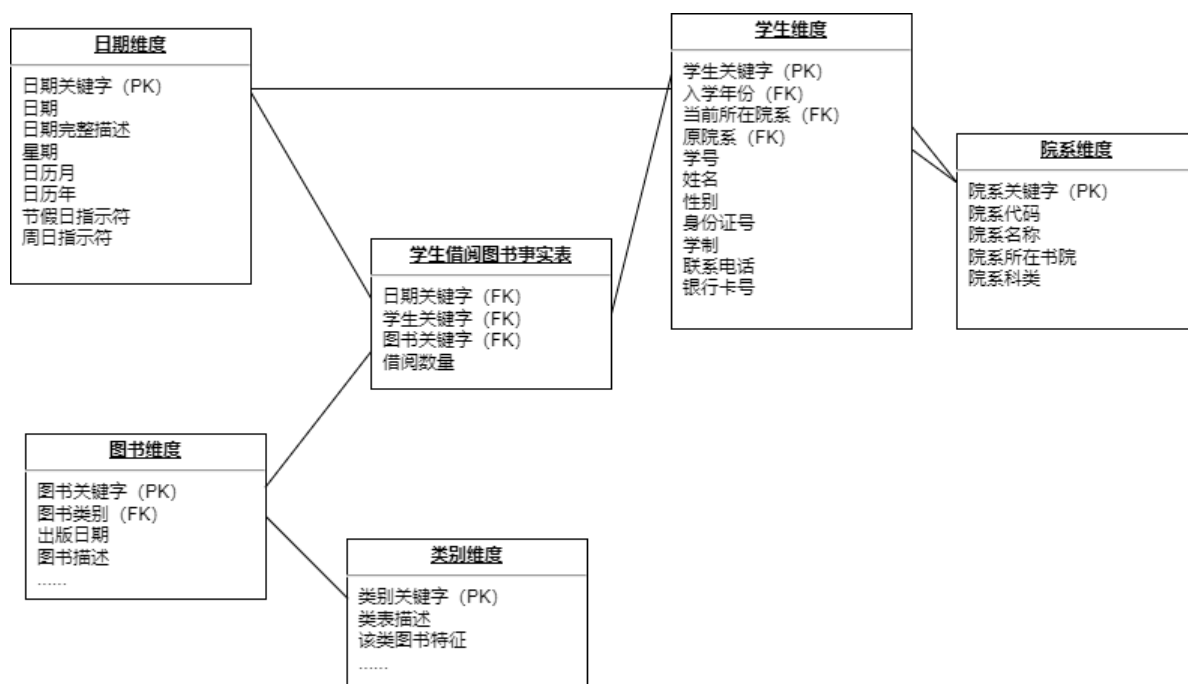
### 5.2.3 选取维度

日期、学生、图书

### 5.2.4 选取度量值

借阅数量（图书馆每本书都有多本）

### 5.2.5 多维建模设计



### 5.2.6 说明

采用雪花模型，图书维度中用外键连接类别维度，层次更加清晰，维表更加规范化；此外，由于图书馆通常藏书量非常大，这样做也能节省一定的空间。

## 5.3 食堂销售情况分析

### 5.3.1 对于该分析应用业务的描述：

对食堂的销售数据进行分析，分析比较受师生员工欢迎的菜，以及相关菜品的利润信息，在成本可控的情况下为师生提供更加可口的饭菜。

### 5.3.2 定义粒度

单条食堂售卖记录

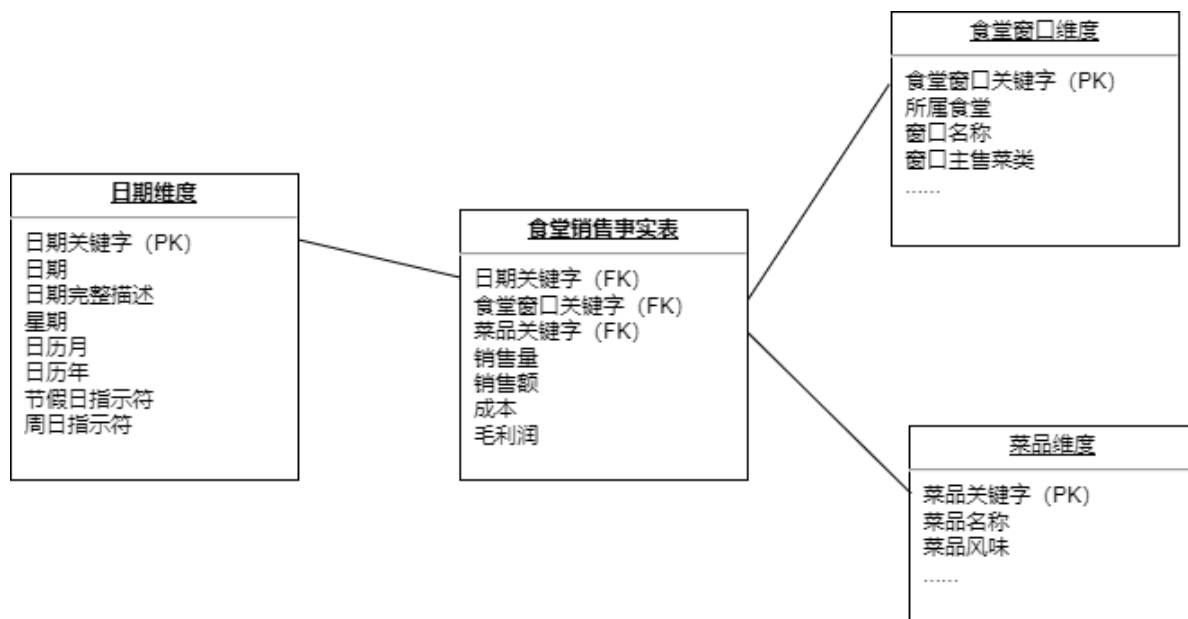
### 5.3.3 选取维度

日期、食堂窗口、菜品

### 5.3.4 选取度量值

销售量、销售额、成本、毛利润

### 5.3.5 多维建模设计



### 5.3.6 说明

使用星型模型，记录什么时间什么窗口售卖了多少什么菜，比较方便，也更直观明了。

## 5.4 受欢迎课程分析

### 5.4.1 对于该分析应用业务的描述：

根据学生的选课情况、期末得分分布以及对课程的评分等信息，分析受学生欢迎的优质课程有哪些。

### 5.4.2 定义粒度

学生的单条选课信息

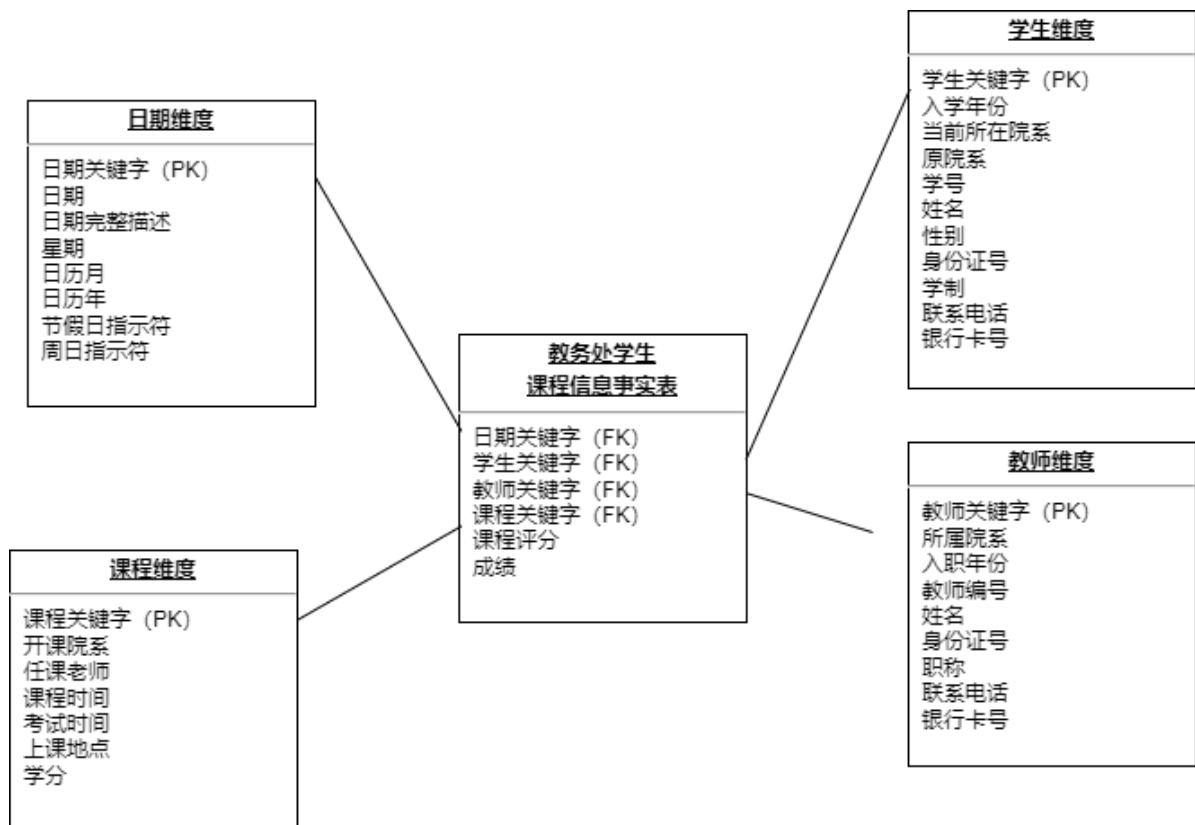
### 5.4.3 选取维度

日期、学生、教师、课程

## 5.4.4 选取度量值

课程评分、成绩

## 5.4.5 多维建模设计



## 5.4.6 说明

使用星型模型，课程分析设计日期、学生、教师、课程四个维度，用星型表比较直观清晰。