

数据仓库设计

1 业务描述

为了完成总线型设计，本次将挑选作业 1 中关联性较强的的几个业务进行详细说明。

1.1 教务处

教务处主要的分析应用业务包括：

- 学生成绩分析：分析学生的考试成绩，统计不同专业、不同班级、不同年级的平均分、最高分和最低分，以及各科目成绩分布情况。
- 教学质量评估：对教师的教学效果进行评估，包括对学生的评价、教师的授课质量、教学方法等方面的统计分析。
- 课程安排分析：分析学生的选课情况，统计不同学院、不同专业、不同年级的选课情况，以及各课程选课人数分布情况。
- 教学计划制定：分析在校学生的成绩情况、毕业学生的就业情况，统计不同学院、不同专业、不同年级的情况。

1.2 学生处

学生处主要的分析应用业务包括：

- 学生成绩分析：分析学生的考试成绩，统计不同专业、不同班级、不同年级的平均分、最高分和最低分，以及各科目成绩分布情况。

1.3 财务处

财务处主要的分析应用业务包括：

- 财务报表分析：对学校的财务报表进行分析，评估学校的财务状况，并预测未来财务走势。
- 资金管理分析：对学校的资金使用情况进行分析，评估学校的资金管理效率，并对未来的资金需求进行预测。
- 财务风险评估：对学校的财务风险进行评估，并给出风险预警，以防止财务危机的发生。

1.4 人事处

- 员工数据分析：人事处需要了解员工数量、结构、流动等情况，以实现员工管理的有效性。
- 工资分析：人事处需要对员工工资数据进行分析，确定工资结构是否合理，以便于决策者制定出有效的薪酬政策。
- 教职工绩效分析：人事处需要分析员工的绩效数据，以实现员工绩效管理的有效性，并对员工进行激励和奖励。由于学校中员工可以细分为教师和职工，因此绩效分析也会有所不同。对教师的评估需要着重于教学质量、科研产出等，而普通职工则类似一般公司中的绩效考核。

2 数据仓库维度建模

2.1 总线结构设计

使用总线结构设计数据仓库，要求为各个部门设计一致性维度和事实。接下来详细分析公共维度的选择以及给出最终的总线矩阵。

2.1.1 业务过程

从第一部分的业务描述中，抽出以下这些关键业务过程：

- 教务处/学生处：成绩分析、教学质量评估、课程安排
 - 成绩分析综合分析每个学生的所有成绩。
 - 教学质量评估通过学生完成评教任务，结果记录在评教事实表当中。
 - 课程安排分析需要分析学生跨专业选课记录。
- 财务处：财务报表分析
- 人事处：工资分析

2.1.2 公共维度

根据抽取出的业务过程，我们可以确定公共维度及其属性如下：

- 学生维度：学号（PK）、姓名、联系方式、所属院系（FK）、年级
- 教师维度：职工编号（PK）、姓名、工龄、部门（FK）
- 职工维度：职工编号（PK）、姓名、职称、工龄、所属院系（FK）
- 课程维度：课程编号（PK）、课程名称、开课院系（FK）
- 日期维度：日期关键字（PK）、日期、年月
- 学期维度：学期关键字（PK）、开始日期、结束日期
- 院系维度：院系编号（PK）、名称、位置
- 部门维度：部门编号（PK）、名称、位置

注：一名教师同时拥有教师维度和职工维度，它们的使用场景不同。

2.1.3 总线矩阵

由此我们可以画出如下的总线矩阵：

<input type="checkbox"/>		A 业务/维度 ?	<input checked="" type="checkbox"/> 日期	<input checked="" type="checkbox"/> 学期	<input checked="" type="checkbox"/> 学生	<input checked="" type="checkbox"/> 教师	<input checked="" type="checkbox"/> 职工	<input checked="" type="checkbox"/> 课程	<input checked="" type="checkbox"/> 院系	<input checked="" type="checkbox"/> 部门
1		成绩分析	✓	✓	✓			✓	✓	
2		教学评估	✓	✓		✓		✓	✓	
3		选课分析		✓	✓			✓	✓	
4		财务报表分析	✓							✓
5		工资分析	✓				✓			✓

2.2 事实表构建

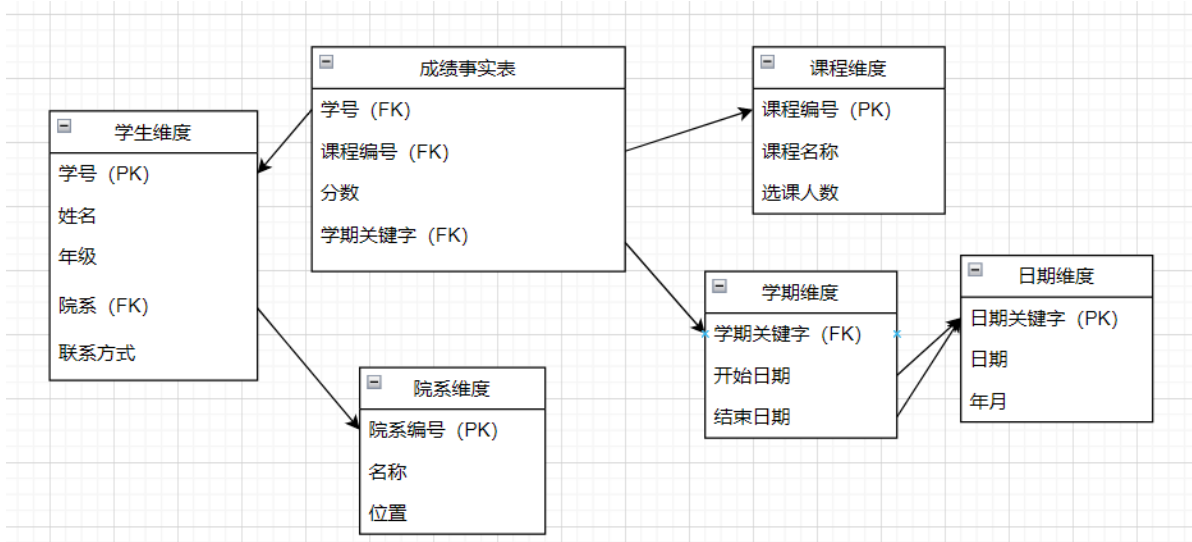
事实表采用雪花型结构与维度进行关联。如果考虑效率问题，那么是星型结构优先。但是这里我发现一些小问题。比如在考虑选课事实表时，可以考虑将院系维度直接加入到事实表中。可是教师的院系和课程的开课院系可能是不一样的。因此院系维度应该被教师维度和课程维度分别持有。

成绩分析事实表

粒度说明：成绩分析表中，一行是一条学生的成绩记录

对事实表中的维度、度量值进行说明：

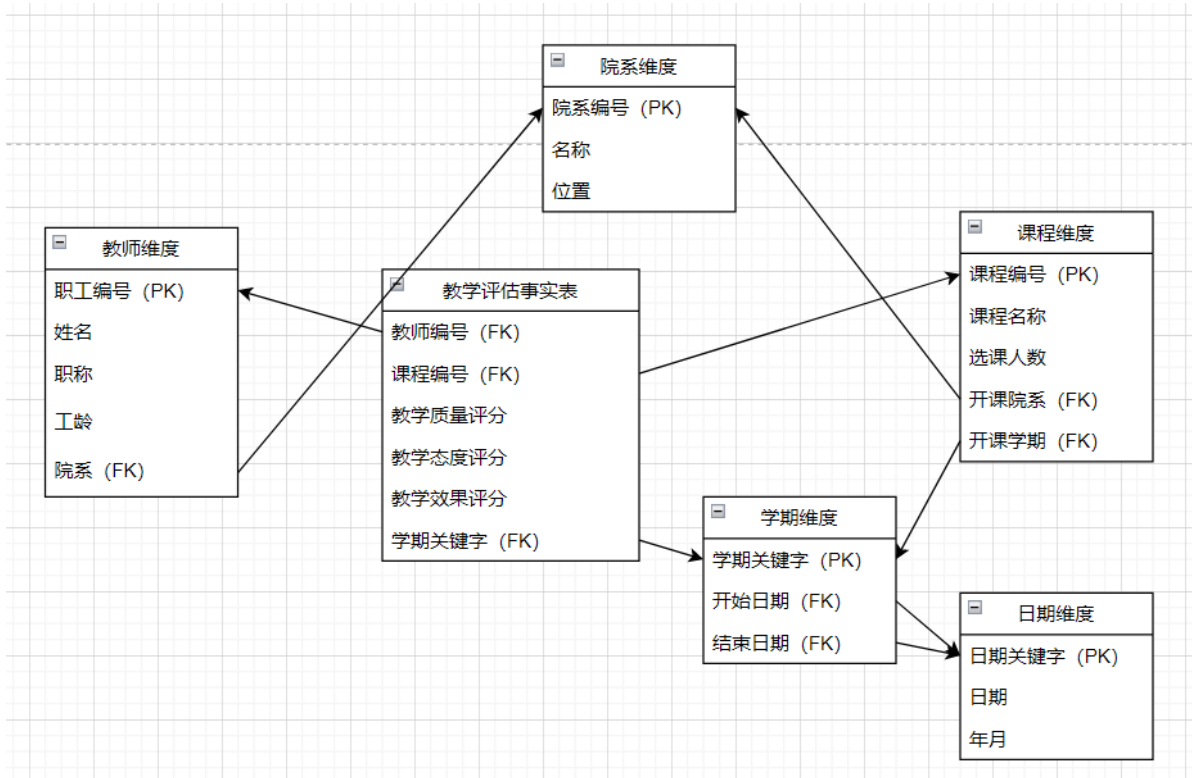
- 学号：学生编号，外键关联学生维度表
- 课程编号：学生的该门课程编号，外键关联课程维度表
- 分数：学生取得的分数，0-100 分
- 学期：该门课程所在学期，外键关联学期维度表



教学质量评估事实表

粒度说明：教学质量评估事实表中一行是一条评教记录
对事实表中的维度、度量值进行说明：

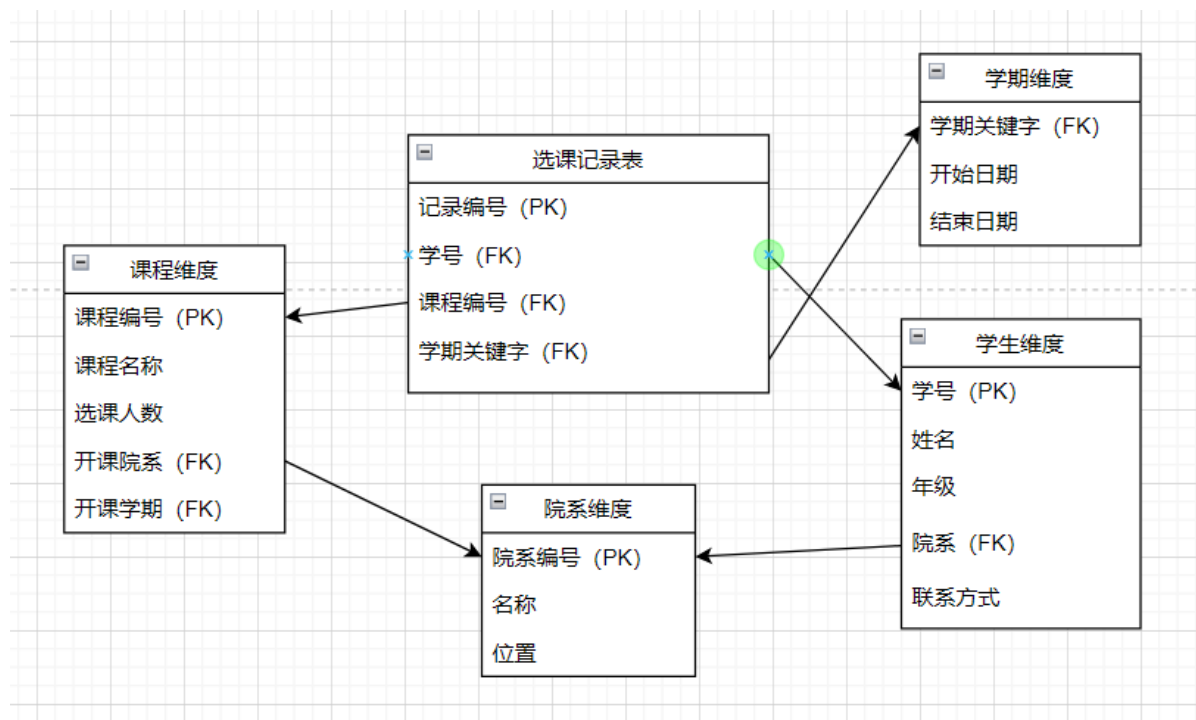
- 教师编号：评教的教师的编号，外键关联教师维度表
- 课程编号：评教的课程编号，外键关联课程维度表
- 评教学期：评教的学期，外键关联学期维度表
- 教学质量评分：对教师教学质量的评分，从 1-10 评分
- 教学态度评分：对教师教学态度的评分，从 1-10 评分
- 教学效果评分：对教师教学效果的评分，从 1-10 评分



选课情况事实表

粒度说明：选课情况事实表中，一行是一条某学生的选课记录。
对事实表中的维度、度量值进行说明：

- 学号：该选课学生的学号，外键关联学生维度表
- 教师编号：选课记录的编号，唯一确定一条选课记录
- 课程编号：所选课程编号，外键关联课程维度表
- 学期关键字：选择该课程的学期，外键关联学期维度表

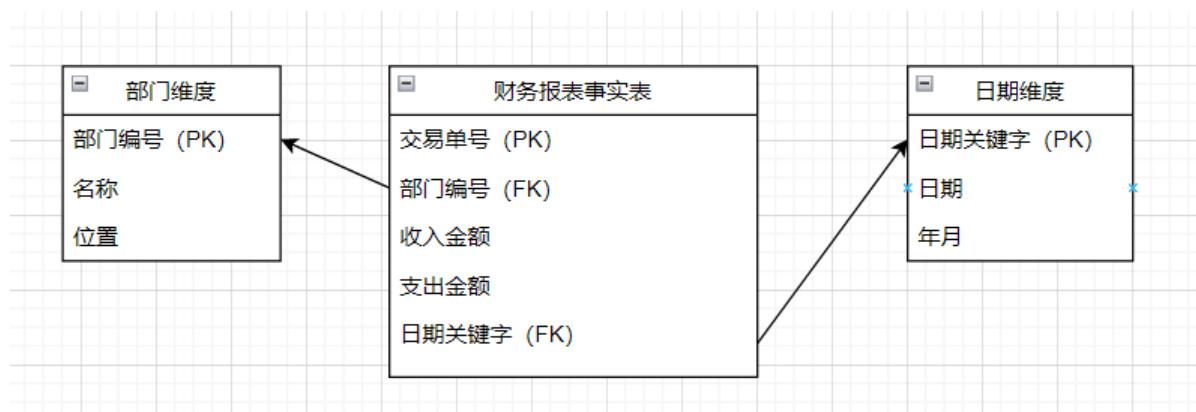


财务报表分析事实表

粒度说明：财务报表分析事实表中，一行是一条财务记录。

对事实表中的维度、度量值进行说明：

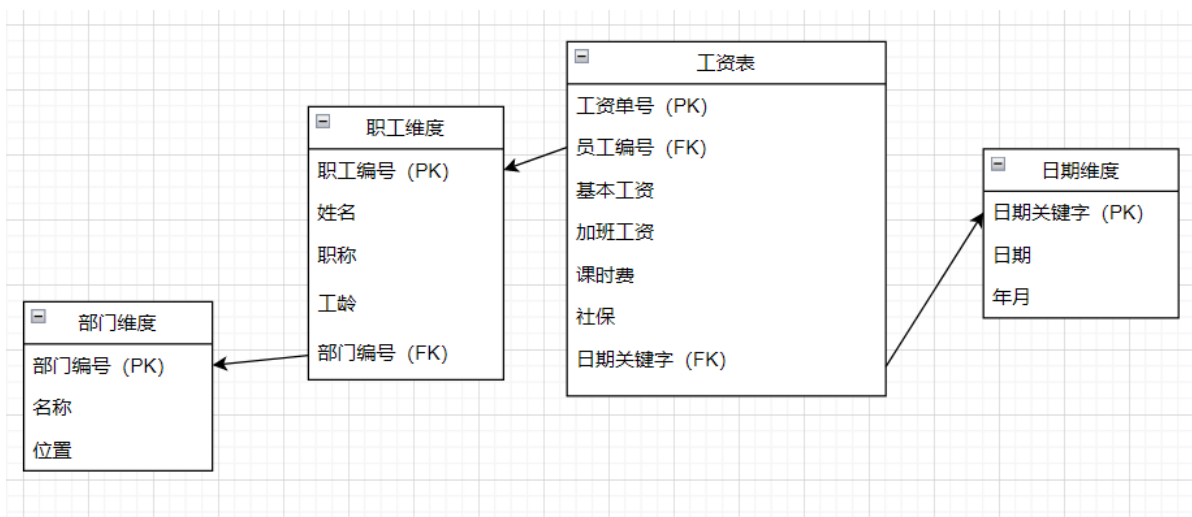
- 交易单号：标识每次交易的唯一编号，用于索引
- 部门编号：相关部门的编号，外键关联部门维度表
- 日期关键字：交易日期，外键关联日期维度表



工资分析事实表

粒度说明：工资分析事实表中，一行是一条工资单记录。

- 工资单号：标识工资单的唯一编号，用于索引
- 日期关键字：交易日期，外键关联日期维度表



3 数据映射关系

3.1 成绩分析

学生每门课程得到分数后，由教师或助教将该分数录入教务系统，生成一条成绩记录。此时得到的是操作性数据。教务处需要进行成绩分析的时候，可以通过一名学生所有的成绩记录计算出该学生每学期的成绩浮动变化、整体趋势等分析型数据；也可以横向地计算一门课程不同班级的分数差异等等，这些都是从操作型数据得到的。

3.2 教学质量评估

学生在每个学期末被要求完成评教任务，对他们该学期上的所有课程进行评分。得到的各项分数被加入到一条评教记录当中，这是操作型数据。当教务处需要对教师的教学质量进行评估的时候，选取该老师所有开设的课程的相关记录，计算出学生所反应的该教师的教学水平。通过这些计算得到的便是分析型数据。

3.3 选课情况

学生在学期初在教务系统中进行选课，每次选课操作均产生一条选课记录加入到表中，得到操作型数据。教务处在之后进行课程时间安排时可以分析这些操作型数据，通过计算找到某一专业学生倾向于选择的其他专业的课程，合理安排课程时间，尽可能地减少冲突的可能性。这时得到的是操作型数据。

3.4 财务报表

学校每个部门在进行采购等财务事务时，均会由财务处生成财务报表记录，基本的财务报表记录是操作型数据。财务部可以根据财务报表记录以及各部门相关的其他收益方面记录，计算资金管理水平、财务风险等。经过这些计算得出的分析型数据可以对财务处决策起到至关重要的帮助作用。

3.5 工资分析

学校里的教职工按月获得工资，每次工资单都会被财务处负责人员记录在系统当中，得到操作型数据。财务处可以使用这些记录，计算某一位职工/教师的绩效与收入的相关程度，获得分析型数据。