

作业一 组织构建与主题的明确

一、高校的业务背景，分析部门，分析需求

高校的业务背景

高校部门的金字塔模型

高校分级呈金字塔式样，共分五级：

第一级是党务群团部门，大致包括学校党政办公室、组织部、宣传部、统战部、校纪委、学生处、校工会等。

第二级是行政职能部门，大致包括教务处、科研处、人事处、财务处、后勤保障处等。

第三级是各大学院部门。

第四级是教辅及直属单位。

第五级是科研机构。

第一级：党务群团部门

党政办公室

协助党政处理日常事务。

组织部

负责党的基层组织建设、党员教育管理、党员发展、党费管理等。

附属业务：情况通知、党费记录、发展评定。

宣传部

是中共中央主管意识形态方面工作的综合职能部门。

附属业务：印刷出版宣传、电视出版物宣传。

统战部

是党委主管统一战线工作的职能部门，承担了解情况、掌握政策。

校纪委

是党内监督专责机关，在校党委和上级纪检组织的双重领导下开展工作。

学生处

是负责全校普通本科学生招生、就业、教育管理、资助、军训、国防教育等工作的职能部门。

附属部门：招生办、就业服务咨询中心、奖学金/助学金管理中心、部队通讯部门等。

附属业务：学生信息管理、奖助学金管理。

校工会

是工会会员代表大会的常设机构和工作机构。

第二级：行政职能部门

教务部门

教务部门有教学质量科、教材科、综合科、实习检验科，

附属业务：负责教学质量管理、教材管理、教师管理、学籍管理、考务管理、教学改革研究。

总务部门（后勤部门）

总务部门总理各项事务，对单位内的行政后勤工作的保障，

附属业务：负责水电暖处理、校园环境保护、校内医疗保健、设施维修采购、收费管理、获取反馈。

人事部门

人事部门是一个企业或者团体的人力资源和企业公共事务的管理部门。

附属业务：负责招聘、教师发展、人才交流。

宿管部门

宿管部门负责对全校学生宿舍作息、纪律和卫生的检查维护。

附属业务：负责学生宿舍资产管理、设备维修保养、入学离校工作、假期住宿工作。

政教部门

全校学生的思想政治教育工作。

附属业务：学生的思想、政治、品德管理。

科研部门

科研部门用于管理学校科学研究。

附属业务：产学研办公室、学术委员会办公室、学报编辑部。

财务部门

财务部门指组织领导和具体从事财务管理工作的职能部门。

附属业务：会计、预算、核算、债务部门等。

审计部门

审计是对资产作出证据搜集及分析，以评估企业财务状况，然后就资料及一般公认准则之间的相关程度作出结论及报告。

附属业务：运作、履行、财务报表。

发展规划部门

制定战略、工作重点和方向。

附属业务：高等教育研究所、社会合作办公室、教学质量评估部门。

国际合作交流部门

与其它学校、研究所、政府部门合作。

附属业务：外事办公室、产学研办公室。

信息化建设与管理部门

构建信息系统、进行数据存储、修改和安全的管理。

附属业务：数据中心、网络安全部门。

第三级：各级学院部门

学校的学院部门很多，这里选取他们共有的业务。

出版业务

印刷业务

宣传业务

合作业务

采购业务

租赁业务

第四级：教辅及直属单位

图书馆

搜集、整理、收藏图书资料以供人阅览、参考的机构。

附属业务：出版、印刷、通行记录管理。

档案馆

收集、保管档案的机构。负责接收、征集、管理档案和开展档案利用等。

附属业务：档案保存管理、档案数字化记录。

校医院

为学生以及教职工提供医疗服务。

附属业务：医疗采购、医疗信息处理、医疗废物处理。

学校出版部

学校内进行图书、图画、杂志、报纸和电子物品等有版权物品的出版活动的组织。

附属业务：出版、印刷。

附属学校

学校在其它地区开设的附属中学、附属小学等。

第五级：科研机构

同第三级：各级学院部门。

分析部门

党务群团部门——组织部、学生处

行政职能部门——教务部门、总务部门、人事部门、宿管部门、财务部门、审计部门

各级学院部门

教辅及直属单位——图书馆、档案馆

科研机构

分析需求

组织部

组织部负责党的基层组织建设、党员教育管理、党员发展、党费管理等。

学校需要对党员、团员的党费、团费进行数字化记录，并和学校的数据中心进行连接；

学校需要记录学生的发展评定情况，便于学生的评奖评优、留学等活动的相关信息填写。

学生处

学生处是负责全校普通本科学生招生、就业、教育管理、资助、军训、国防教育等工作的职能部门。

学校需要对招生系统的数据进行探查，例如，在“南星梦想计划”中，学校招生办会将有意向的学生信息录入系统。

学校需要对学生曾经获得的奖助学金进行记录。这样可以保证流程的公正透明。

教务部门

教务部门负责教学质量、教材管理、教师管理、学籍管理、考务管理和教学改革研究。

学校需要对教师的信息和学生的学籍信息进行管理，以便于杜绝无关人员。

学校需要管理教材数量和教材采购信息。

总务部门

总务部门负责水电暖处理、校园环境保护、校内医疗保健、设施维修采购、收费管理、获取反馈。

学校的总务部门应该将反馈收集起来，便于长期的反馈流程。

学校应该保存设施采购、维修的详细报表，便于财务查账。

学校应该掌控总务部门的收费记录。

人事部门

人事部门是一个企业或者团体的人力资源和企业公共事务的管理部门。

人事部门应该保有一部分的人力资源储备信息，并且为这些预备人员进行分级。

宿管部门

宿管部门负责对全校学生宿舍作息、纪律和卫生的检查维护。

宿管部门同总务部门，应该有设备维修保养的数字化报表。

宿管部门应该掌握学生的入学离校信息和假期住宿信息。

财务部门&审计部门

财务部门指组织领导和具体从事财务管理工作的职能部门。审计是对资产作出证据搜集及分析，以评估企业财务状况，然后就资料及一般公认准则之间的相关程度作出结论及报告。

学校应当掌握各部门的财务报表和审计状况，并且赋予其较高的密级。

各级学院部门&科研机构

无论是何种类型的学院，都担负着诸如合作、租赁和采购的任务。

学校的教职工同样也是学院的教职工，需要细化职工信息。

学院部门应该保证信息的实时性，同样要和学校的数据中心进行连接。

图书馆

图书馆是负责搜集、整理、收藏图书资料以供人阅览、参考的机构。

图书馆应该保有图书的电子借阅记录并且和学校的信用系统进行连结。

图书馆应该保有电子化的图书查询系统，最终实现不需要人力寻找图书资源的理想化状态。

图书馆应该有学生的图书馆通行记录。

档案馆

档案馆是负责收集、保管档案的机构。负责接收、征集、管理档案和开展档案利用等。

档案馆应该负责对档案的保存管理和数字化记录。

二、完成数据仓库的系统规划

注：作业一未要求进行概念模型设计，但以下文字已经有了概念模型的雏形。

划分主题

组织部

组织部财务主题

组织部财务主题需要对党员、团员的团费进行数字化记录；

- 1.财务主题应该包含日期信息、党团员信息、费用缴纳信息；
- 2.财务主题的每一个项目应该是一条包含以上三条消息的记录；
- 3.对党团员信息进行细化，例如编号的形式，以更明细地区分每一个人；
- 4.日期信息包含时间具体到分钟且便于筛选；
- 5.费用缴纳信息包括应缴纳款项、实际缴纳款项、缴纳状态；

主码键是团费缴纳编号。

组织部发展教育主题

组织部发展教育主题需要对学生的的发展评定情况进行记录；

1.发展教育主题应该包括党团员编号、姓名，发展评定进度，材料提交情况，评奖评优情况等；

2.材料提交，评奖评优情况是泛数组类型，其中包含具体的日期和备注；

主码键是党团员编号。

学生处

学生处招生主题

学生处招生主题会对有意向加入本高校的学生信息进行录入；

1.招生主题应该包括意向学生姓名及其编号、意向学生所在学校、目前的成绩情况、联系方式；

主码键是意向学生编号。

学生处就业主题

学生处就业主题与招生主题相似；

1.就业主题包括学生姓名和学号，是否就业，以及就业方向；

主码键是学生学号。

学生处资助主题

学生处资助主题含有在校学生的奖助学金获取情况；

1.资助主题包括学生姓名和学号，奖学金获得情况和助学金获得情况；

2.奖学金/助学金获得情况中的每一条都应该含有获得时间和获得款项，以及该奖金的备注；

主码键是学生学号。

教务部门

教务部门学生信息/教师信息主题

教务部门的学生信息/教师信息主题是整个数据仓库中最为重要的主题之一，因为许多主题都能够通过学号来获得学生的主要信息和来自于其它主题的信息。

1.学生信息主题包括学生的姓名、学号、入学时间、所在专业、银行账户、成绩情况、评奖评优与不良记录情况等。

2.主码键是学生学号。

3.教师信息主题包括教师姓名、工号、入职时间、职务、工资账户、评奖评优情况、联系方式（电子邮箱）等。

4.主码键是教师工号。

5.该主题的两个子主题是：

学生信息主题

教师信息主题

教务部门教材采购发放主题

教材的采购发放主题是典型的采购类型主题，有以下子主题：

供应商主题：

包括含有的教材名/ID，供应商编号

采购主题：

包括待采购的教材名/ID，数量

发放主题：

包括待采购的教材名/ID，数量

仓库主题：

包括含有的教材名/ID，数量

主码键除供应商主题都是ID，供应商主题的主码键是供应商编号。

总务部门

反馈主题

反馈主题用来收集反馈信息，并与该部门的其它主题联动；

- 1.反馈主题应该包括反馈条目接收时间，反馈内容，反馈指向主题，反馈条目编号；
- 2.反馈条目接收时间应该含有一个触发器，便于在反馈保存一段时间后加以提醒；
- 3.反馈指向主题是总务部门的另一个主题；

主码键是反馈条目编号。

设施维修采购主题

同教务部门的教材采购发放主题，但是多了一个维修主题：

该主题包括被维修的项目编号，被维修的项目日期，维修编号。

主码键为维修编号。

收费记录主题

收费记录主题用于记录每次收费。

- 1.该主题包括收费单据编号，收费数额，收费原因，收费备注，收费面向群体。

该主题的主码键为收费单据编号。

人事部门

人力资源储备主题

人力资源储备主题与学生处的招生主题相似，

都有人员编号，姓名，地区，联系方式。

但是多了以下项目：

- 1.分级：分级表明该人员拥有在何种工作上的潜力；

2.年龄：人力资源的年龄应该随着时间推移递增。

主码键为人力资源编号。

宿管部门

反馈主题

同总务部门的反馈主题。

设备维修保养主题

同总务部门的设施维修。

学生出入主题

宿管部门应该掌握学生的入学离校信息和假期住宿信息。

1.该条目应该包括学生的姓名、学号、是否在校、是否在宿舍、上一次通过感应门时间，假期是否住宿。

主码键为学生的学号。

宿舍检查主题

宿管部门应该定期对宿舍进行检查并对结果进行保存。

1.该条目应该包括宿舍号、宿舍长学号、宿舍成员学号、获得的反馈条目、宿舍检查时间、宿舍检查情况备注。

主码键为宿舍号和宿舍检查时间。

财务部门/审计部门

财务预算编制主题

财务部门需要划清学校在一段时间内的财务预算。

1.条目包括预算ID、预算接收方、预算时间、预算备注内容。

主码键为预算ID。

财务预算监察主题

针对财务预算编制主题，提出调整建议，防止超支。

1.条目包括预算监察ID、对应的编制主题ID、建议等。

主码键为预算监察ID。

财务政策监察主题

针对以上两者，执行国家的财务会计政策、税收政策和法规。

1.条目包括以上二者的ID、政策监察ID、法规建议。

主码键为政策监察ID。

财务流转主题

负责在编制预算后，对分配不合理的部门之间进行流转。

1.条目包括两个预算编制主题，流入/流出，款项，日期，公证编号等。

主码键为公证编号。

各级学院部门&科研机构

每个部门和机构有略微差异，在这里只列出梗概。

采购主题

仓库主题

租赁主题

合作主题

学术研究主题

图书馆

图书借阅记录主题

1.包括编号、对应图书、借阅人、借阅时间、借阅有效期。

主码键为编号。

图书归还记录主题

1.包括编号、对应图书、归还人、归还时间。

主码键为编号。

图书馆通行记录主题

1.包括编号、人员明细、I/O

主码键为编号。

图书查询仓库主题

1.包括书编号、书借阅情况、剩余本数、所在位置。

主码键为书编号。

档案馆

数字化保存主题

1.将档案馆的资料数字化保存在网络中，让更多人可以在线上参观档案。

2.包括档案编号、分类、对应的数字化媒体。

主码键为档案编号、分类

档案接收与转移主题

1.包括档案编号、分类、第三方、转入/转出标记、时间、ID。

主码键为ID。

按照部门划分主题域

划分主题域的方法有很多：

1.按照业务系统划分

2.按照需求划分

3.按照功能划分

4.按照部门划分

很显然，按照部门划分主题域是最为简单的。在后续的作业中，将使用按照部门划分主题域的方法。

作业二 数据仓库构建

在作业二中，我们选取一个部门来构建模型——图书馆。

多维数据模型的构建过程分为四个步骤：

选择业务过程或主题、选择粒度、确定维度、确定度量。

PPT/作业中的构建过程相似：

选取要建模的业务处理过程、选择粒度、确定维度、选择度量值。

想要完成图书馆的正常运行，以下事务是必不可少的：

书籍采购、书籍摆放、通行管理、借阅与归还。

一、业务描述

图书馆的运行主要分为以下四个部分：

书籍采购

图书馆需要书籍来保证核心资源。

书籍来源的方式有采购、社会捐赠等，这里我们统一用采购来称呼。

书籍摆放

有了书籍之后，就需要对书籍的摆放进行处理，无论是书籍添加到某个位置，还是从某个位置转移或移除，统一定义为摆放。

通行管理

有了书籍，另外一个必不可少的要点就是人了。人员的出入，是图书馆管理的两大核心之一。

借阅与归还

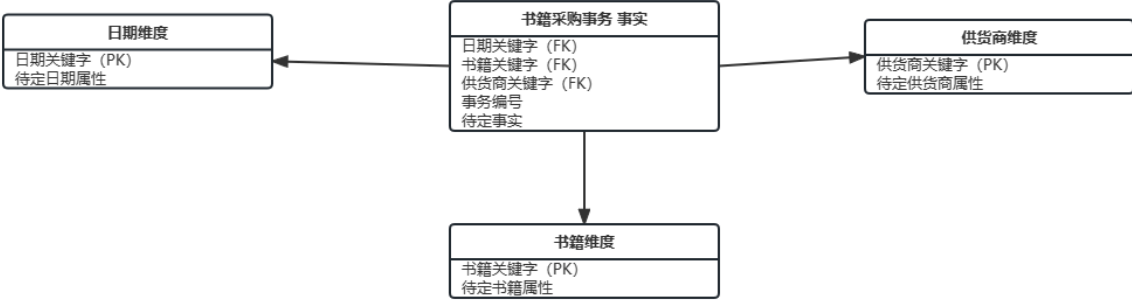
第二大核心是借阅与归还。这一事务不仅涉及了前面的三个事务，是最后也是最重要的一环。

二、书籍采购

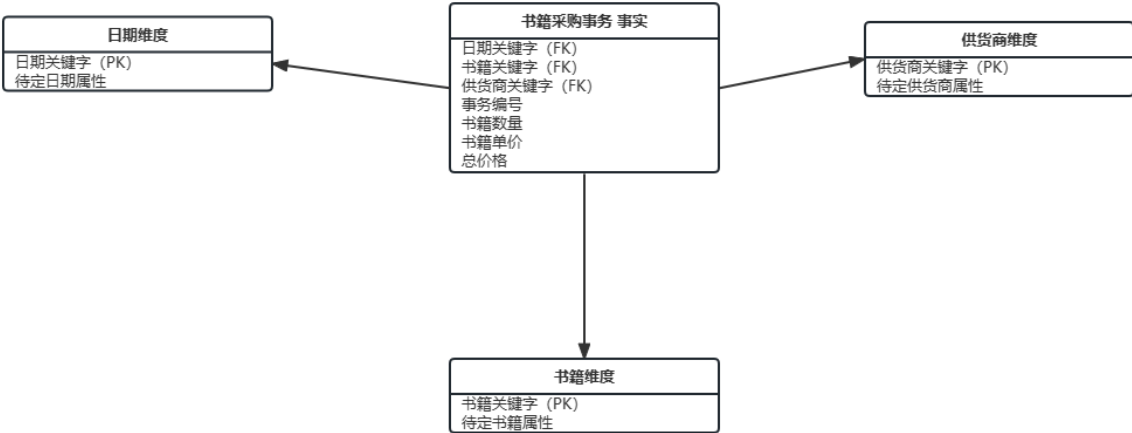
选择粒度

粒度为特定供货商、特定书籍、特定批次采购。

确定维度



确定事实



事实说明

包括书籍数量、单价、总价；
其中总价由数量*单价乘积得到。

日期维度设计

| 日期维度 |
|------------|
| 日期关键字 (PK) |
| 年份 |
| 月份 |
| 日期 |
| 年份月份组合 |
| 月份日期组合 |
| 年月日组合 |
| 星期数 |
| 星期具体日期 |

日期维度表

| 日期关键字 | 年份 | 月份 | 日期 | 年份月份组合 | 月份日期组合 | 年月日组合 | 星期数 | 星期具体日期 |
|---------------|------|----|----|---------|--------|-----------|-----|--------|
| 20230210-rand | 2023 | 2 | 10 | 2023年2月 | 2月10日 | 2023/2/10 | 7 | Fri |

由年份、月份、日期生成除星期和星期具体日期外的其它内容；

由工业化软件生成当前星期数和星期具体日期。

供货商维度设计

| 供货商维度 |
|--|
| 供货商关键字 (PK) 是否为供应商 是否为捐赠方 供应商等级 供应商姓名 供应商联系方式 供应商地址 供货商书籍书单 |

供货商维度表

| 供货商编号 | 是否为供应商 | 是否为捐赠方 | 供应商等级 | 供应商姓名 | 供应商联系方式 | 供应商地址 | 供货商书籍书单 |
|-------|--------|--------|-------|-------|-------------|---------|---------|
| 1 | 1 | 0 | 1 | Ok | 13333333333 | 翻斗花园25号 | 《这就是中国》 |
| 2 | 0 | 1 | | darel | 15555555555 | | |

两种类型：供应商/捐赠方

捐赠方只含有姓名和联系方式。

！供货商维度中的书单关键字实际上也是外码

书籍维度设计

| 书籍维度 |
|---|
| 书籍关键字 (PK) 所属类型 所属类型编号 纸张类型 书籍尺寸 书籍厚度 书籍页数 出版社 作者 |

书籍维度表

| 书籍关键字 | 所属类型 | 所属类型编号 | 纸张类型 | 书籍尺寸 | 书籍厚度 | 书籍页数 | 出版社 | 作者 |
|-------|------|--------|------|-------|--------|------|-------|----|
| 书的名字 | 早教 | 7 | B5 | 17*12 | 1.14cm | 514页 | 零氮出版社 | Ok |

通过书名对不同书进行区分。

现行总线矩阵

| | 日期 | 供货商 | 书籍 |
|------|----|-----|----|
| 书籍采购 | × | × | × |

三、书籍摆放

当前值链

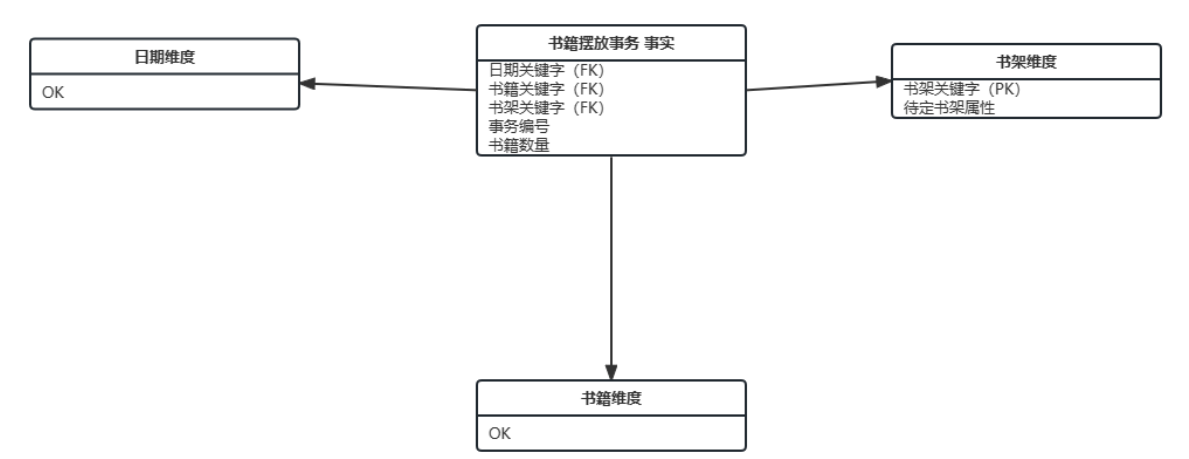
书籍采购——书籍摆放

选择粒度

对于书籍摆放来说，周期快照和累计快照意义不大，选择普通的事务即可。

拟定每本书摆在每个书架上为一个粒度。

确定维度&确定事实



事实说明

书籍的摆放事务包括书籍数量、事务编号、日期关键字、书架关键字和书籍关键字。

书架维度设计

| 书架维度 |
|------------|
| 书架关键字 (PK) |
| 所在楼层 |
| 所在楼层的位置 |
| 书架编号 |
| 书架层数 |
| 书架所属类型 |
| 书架是否已满 |
| 包含书单及其数量 |

书架维度表

| 书架关键字 | 所在楼层 | 所在楼层位置 | 书架编号 | 书架层数 | 书架所属类型 | 书架是否已满 | 包含书单数量 |
|-----------|------|--------|--------|------|--------|--------|--------|
| G302-0001 | 3 | G2 | G302-1 | 5 | 早教 | 是 | 略 |
| F108-0003 | 1 | F8 | F108-3 | 3 | 科幻 | 否 | 略 |

书架的关键字涵盖了楼层、位置和书架编号；

书架所属类型与书籍类型一致；

！ 包含书单与书籍属性一致。

类似的事务——书籍拿取

书籍的拿取看似需要再设计一个事务，但是只要把书籍数量改为负数即可。

现行总线矩阵

| | 日期 | 供货商 | 书籍 | 书架 |
|------|----|-----|----|----|
| 书籍采购 | × | × | × | |
| 书籍摆放 | × | | × | × |

四、通行管理

当前值链

书籍采购——书籍摆放

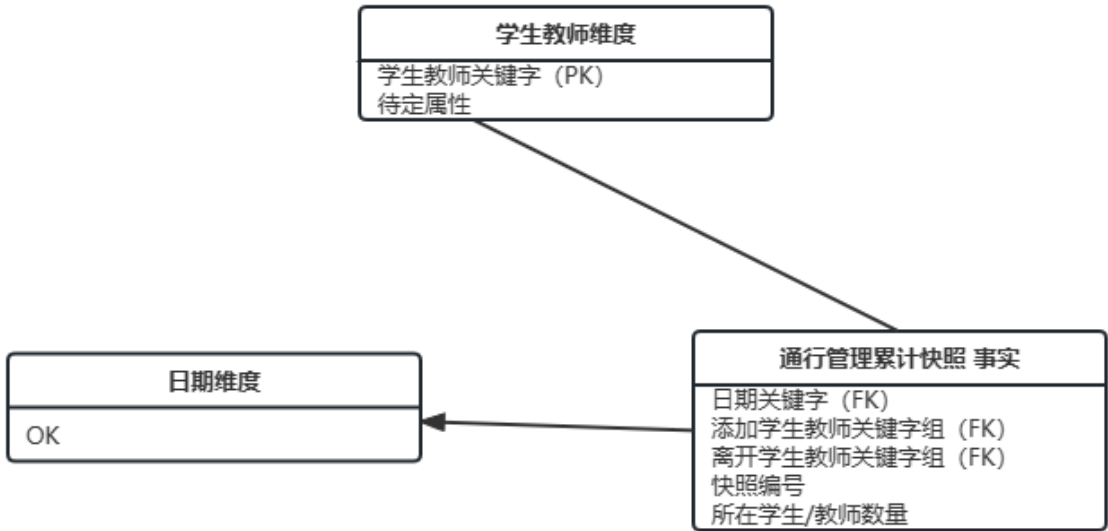
通行管理

选择粒度

相对于书籍的采购与摆放，图书馆人流量不仅对于事务敏感，而且对于累计快照和周期粒度都很敏感。

所以这里选择累积快照。

确定维度&确定事实



事实说明

通行管理累计快照用于记录当前在图书馆的学生人数；

定时更新的快照包含了关键字组；

通过两个关键字组，以及上一次的所在学生/教师数量，可以生成当前的学生/教师数量。

学生教师维度设计

| 学生教师维度表 |
|--|
| 关键字 (PK) 学工号 来过几次图书馆 借阅记录 (FK) 归还记录 (FK) |

学生教师维度表

| 关键字 | 学工号 | 来过几次图书馆 | 借阅记录 | 归还记录 |
|-----------|-----------|---------|-------|-------|
| 201250123 | 201250123 | 100 | array | array |

现行总线矩阵

| | 日期 | 供货商 | 书籍 | 书架 | 学生教师 |
|------|----|-----|----|----|------|
| 书籍采购 | × | × | × | | |
| 书籍摆放 | × | × | | × | |
| 通行管理 | × | | | | × |

五、借阅与归还

当前值链

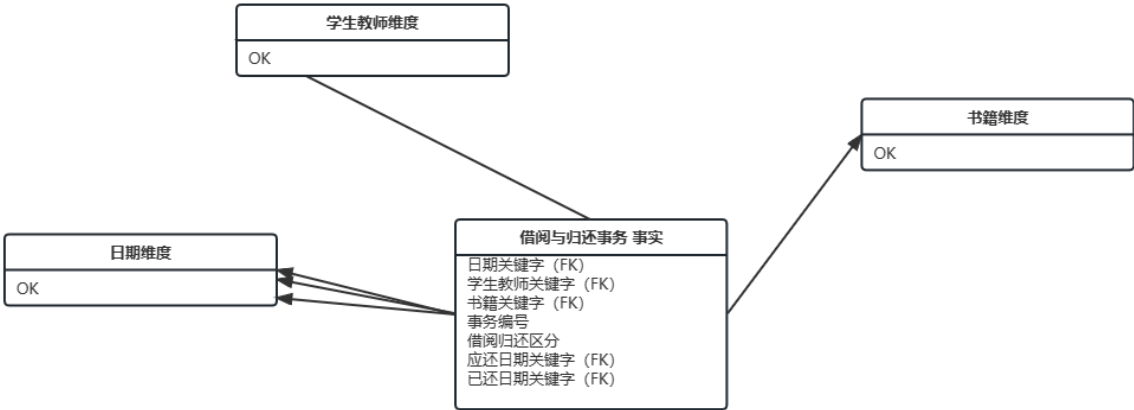
书籍采购——书籍摆放——借阅与归还

通行管理——借阅与归还

选择粒度

以每个学生借一本书为单位

确定维度&确定事实



事实说明

借阅与归还需要同时用到日期、学生教师、书籍三个维度；

借阅与归还事务的事实需要对借阅与归还进行区分，

并设计冗余实现对返还日期的储存。

现行总线矩阵

| | 日期 | 供货商 | 书籍 | 书架 | 学生教师 |
|-------|----|-----|----|----|------|
| 书籍采购 | × | × | × | | |
| 书籍摆放 | × | × | | × | |
| 通行管理 | × | | | | × |
| 借阅与归还 | × | | × | | × |

六、映射关系说明（操作型数据和分析型数据）

1.作业一中的各部门的主题模型实际上是ER模型，因为它们并没有涉及到对具体细节的处理，处理的范围还局限在“该干什么”，而不是“该怎么干”；作业二中大量采用星型模型，真正做到了面向主题；

2.分析型数据是具有时间跨度的一系列操作的汇总，比如通行管理中的快照；而操作型数据做不到实时通行人数管理；

3.分析型数据没有复杂的ER模型，采用总线模型的方式减轻了对性能的要求。

七、公共维度、总线矩阵的作用

公共维度和总线矩阵贯穿了多个多维建模的过程，每次在进行新的建模时，都会查看总线矩阵发现已经在前序业务中写好的维度，极大地减少了复杂度。

八、维度建模思想

1.事务 周期快照 累计快照融合

根据不同的需求采用不同的方式，如书籍摆放使用事务快照，而通行管理采用累计快照。

2.值链

每次在新的建模时采用值链能够使建模流程更加清晰。