SIM Global Education

UOW AUSTRALIA

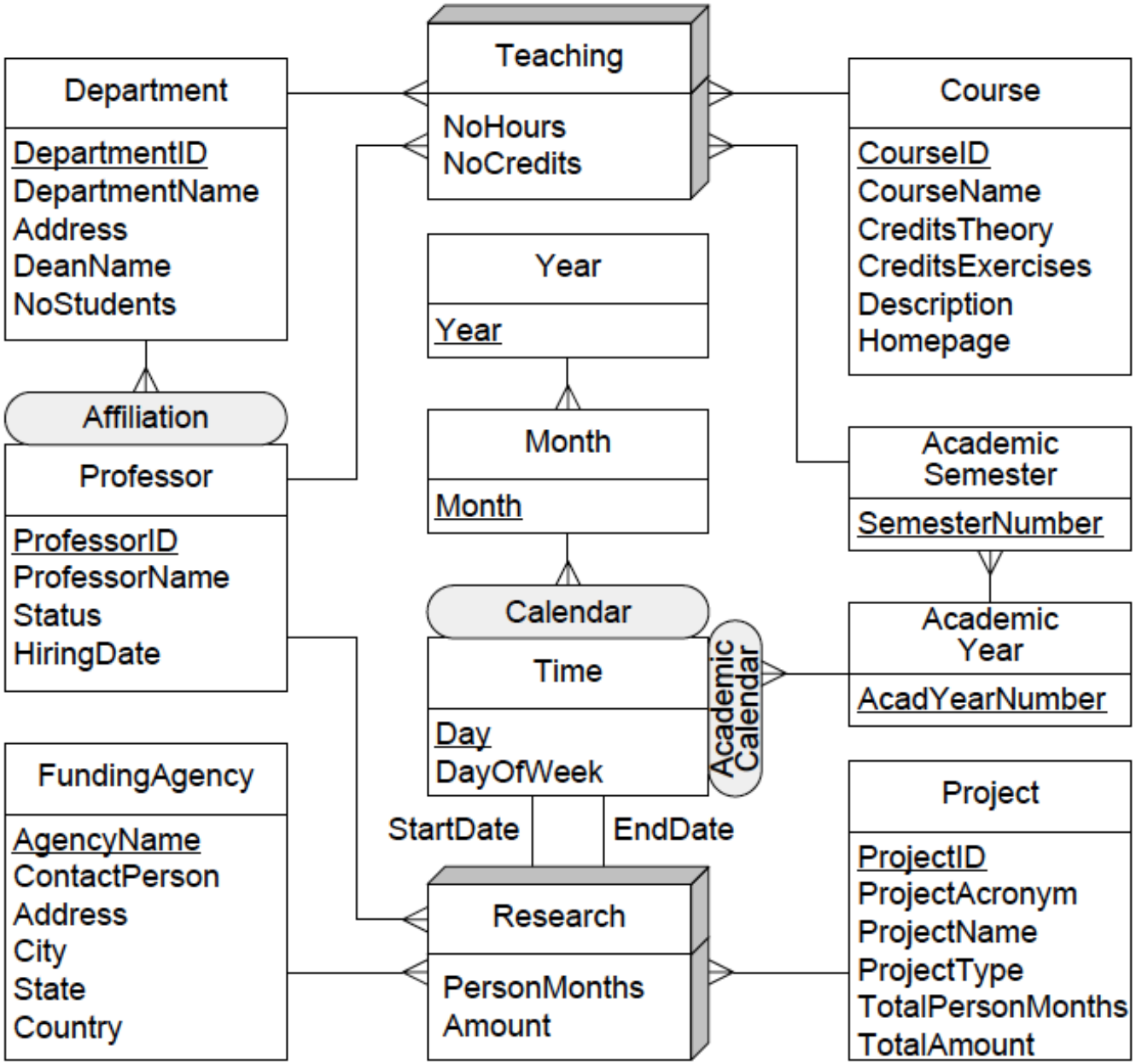
ISIT312 – Big Data Management

Conceptual Warehouse Design

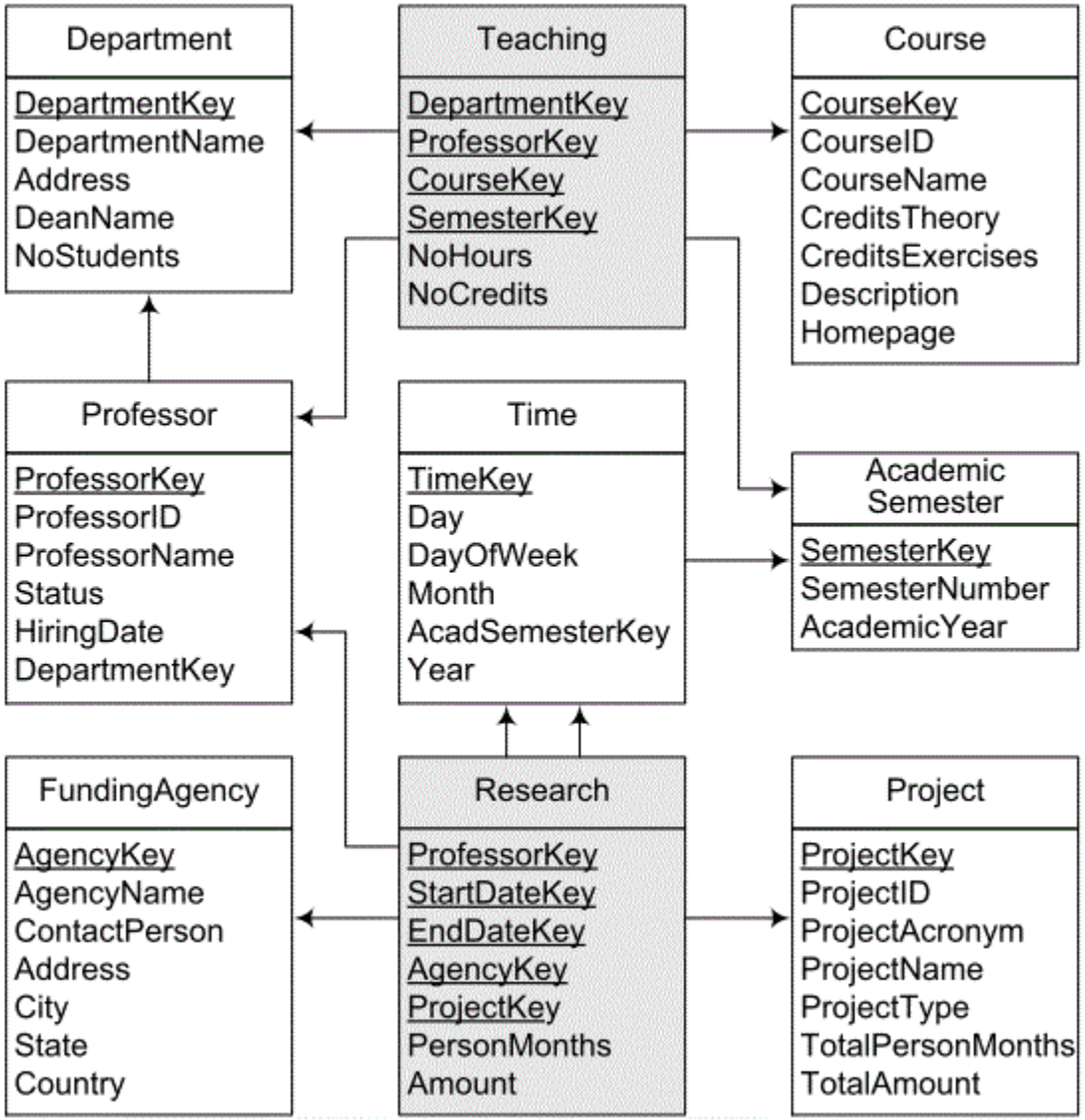
Sionggo Japit

sjapit@uow.edu.au

2 October 2025


Teaching
NoHours
NoCredits
Department
DepartmentID
DepartmentName
Address
DeanName
NoStudents
Course
CourseID
CourseName
CreditsTheory
CreditsExercises
Description
Homepage
Year
Year
Affiliation
Professor
ProfessorID
ProfessorName
Status
HiringDate
Month
Month
Academic Semester
SemesterNumber
Calendar
Time
Day
DayOfWeek
Academic Calendar
Academic Year
AcadYearNumber
FundingAgency
AgencyName
ContactPerson
Address
City
State
Country
StartDate
EndDate
Research
PersonMonths
Amount
Project
ProjectID
ProjectAcronym
ProjectName
ProjectType
TotalPersonMonths
TotalAmount

Constellation Schema



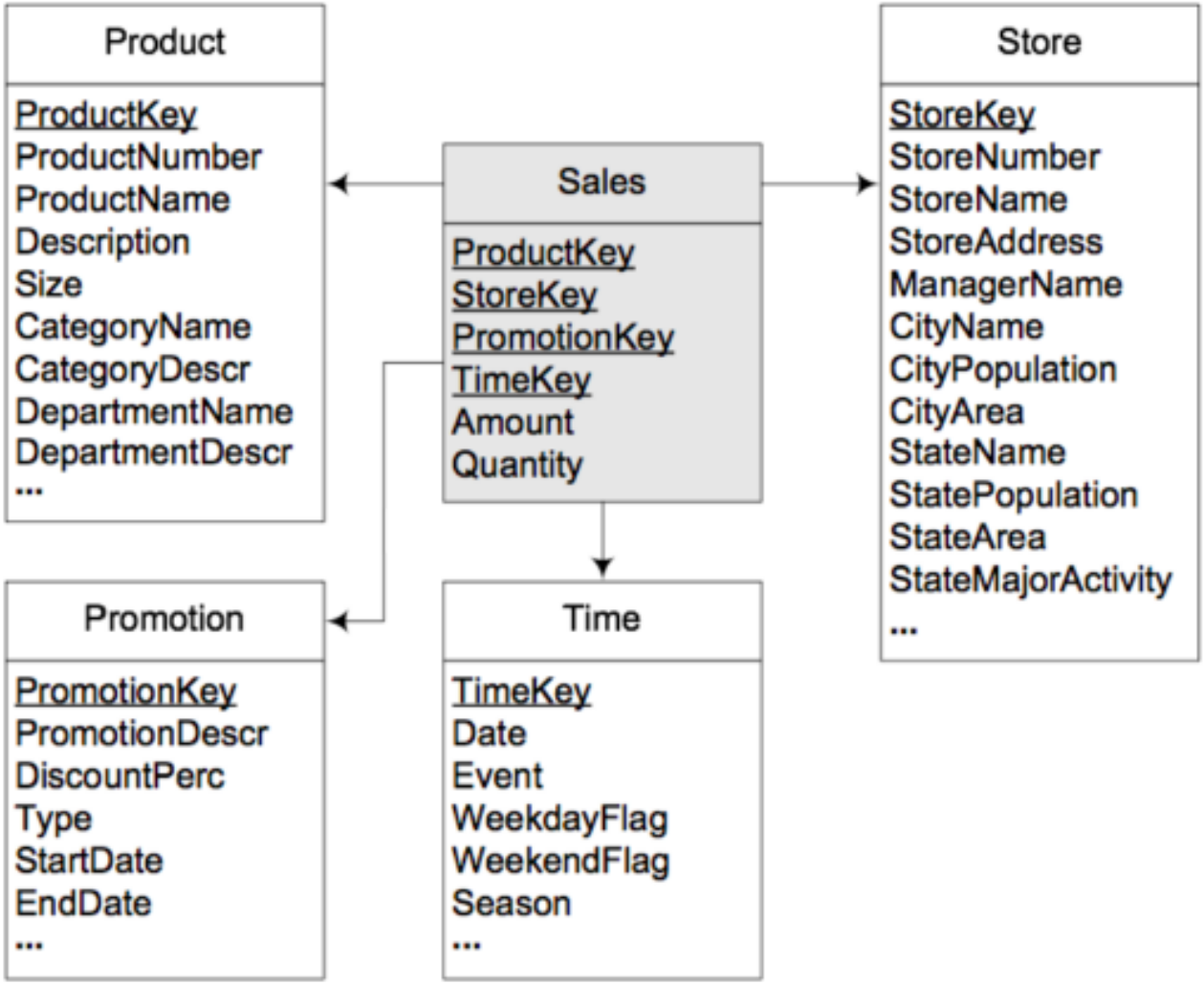
Constellation Schema

Relational Data Warehouse Design – Star Schema

ROLAP systems can be organized using one of the following logical design:

i. Star Schema – Star Schema consists of one fact table and a set of dimension tables.
   - Referential integrity constraints between fact table and dimension tables exist
   - Dimension tables may contain redundancy in the presence of hierarchies
   - Dimension tables are denormalized, and fact tables are normalized.

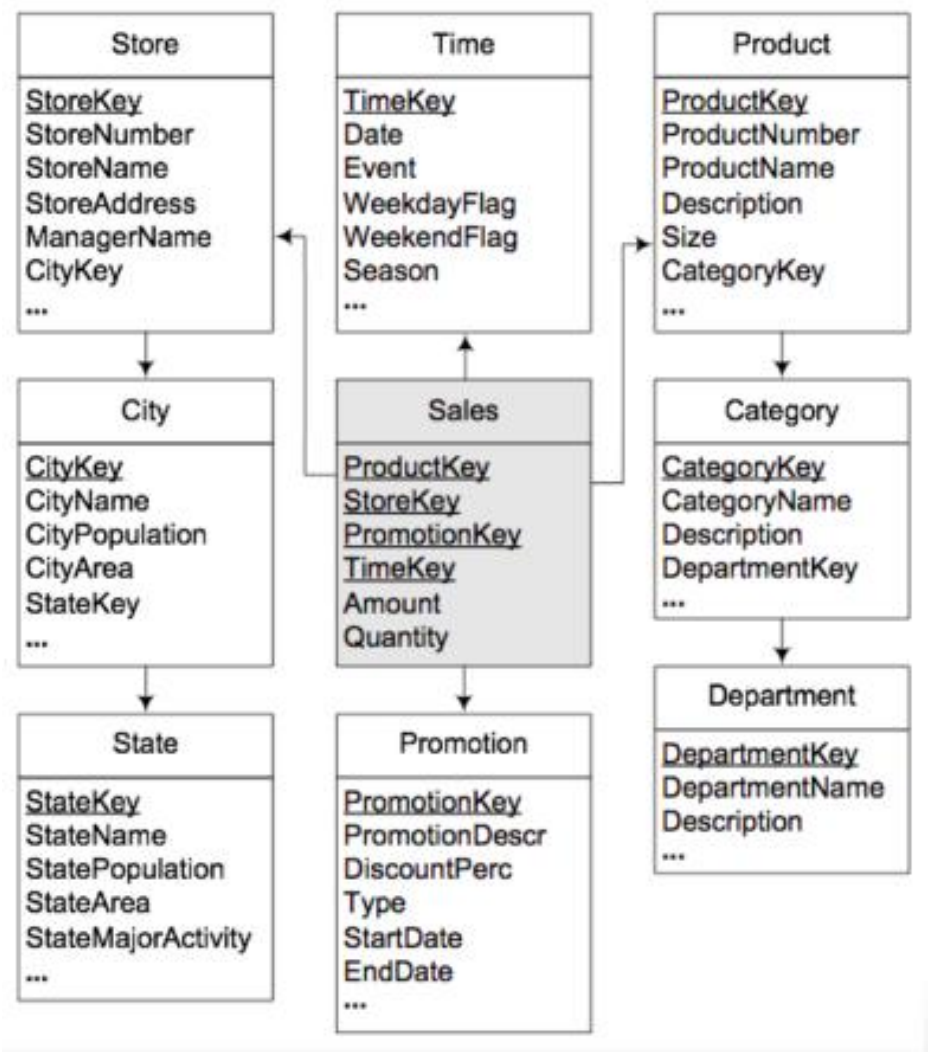
An example of star schema



Relational Data Warehouse Design – Snowflake Schema

ii. Snowflake schema – Snowflake schema is similar to star schema except that the dimension tables are normalized to avoid redundancy. The fact that the dimension tables are normalized, the performance may be degraded because of cartesian product.

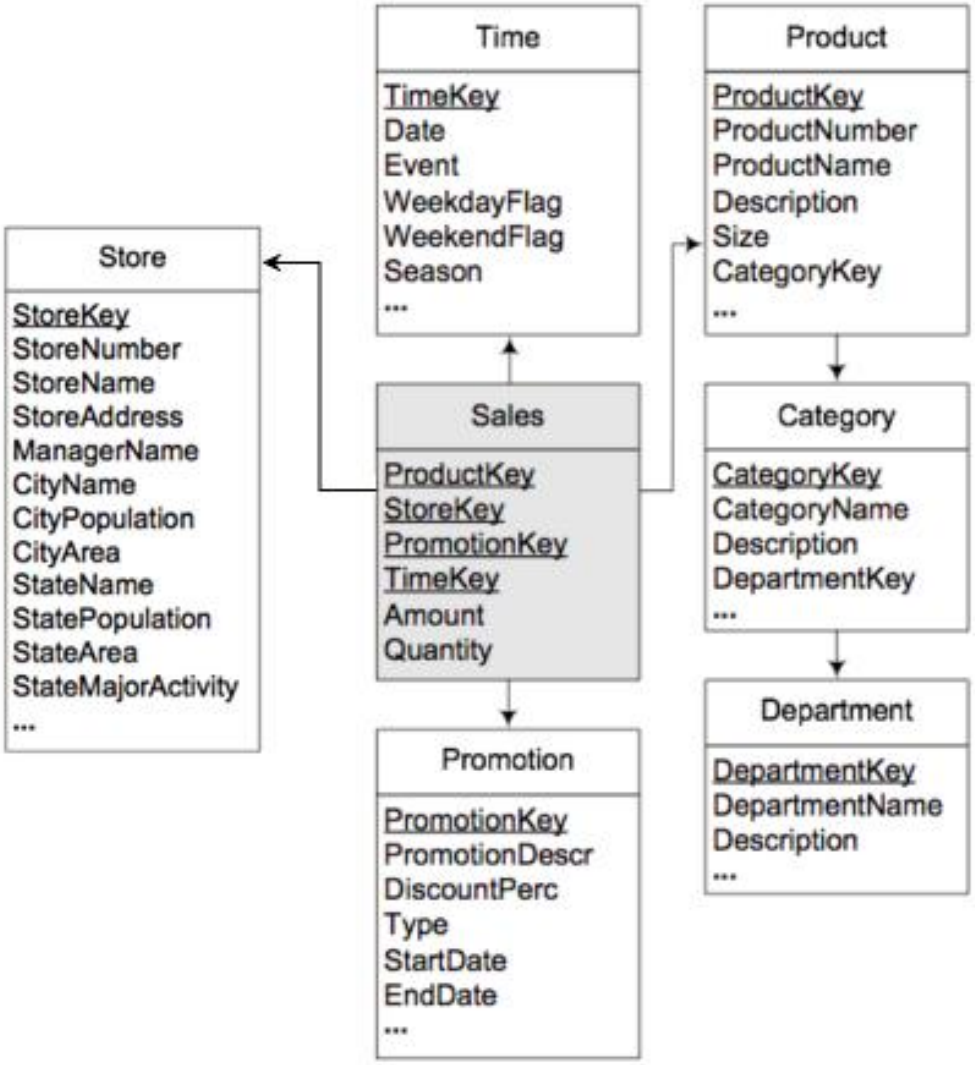
An example of a snowflake schema



Relational Data Warehouse Design – Starflake Schema

iii. Starflake schema – A starflake schema is a combination of star schema and a snowflake schema, that is, some of the dimension tables are normalized while some are not.

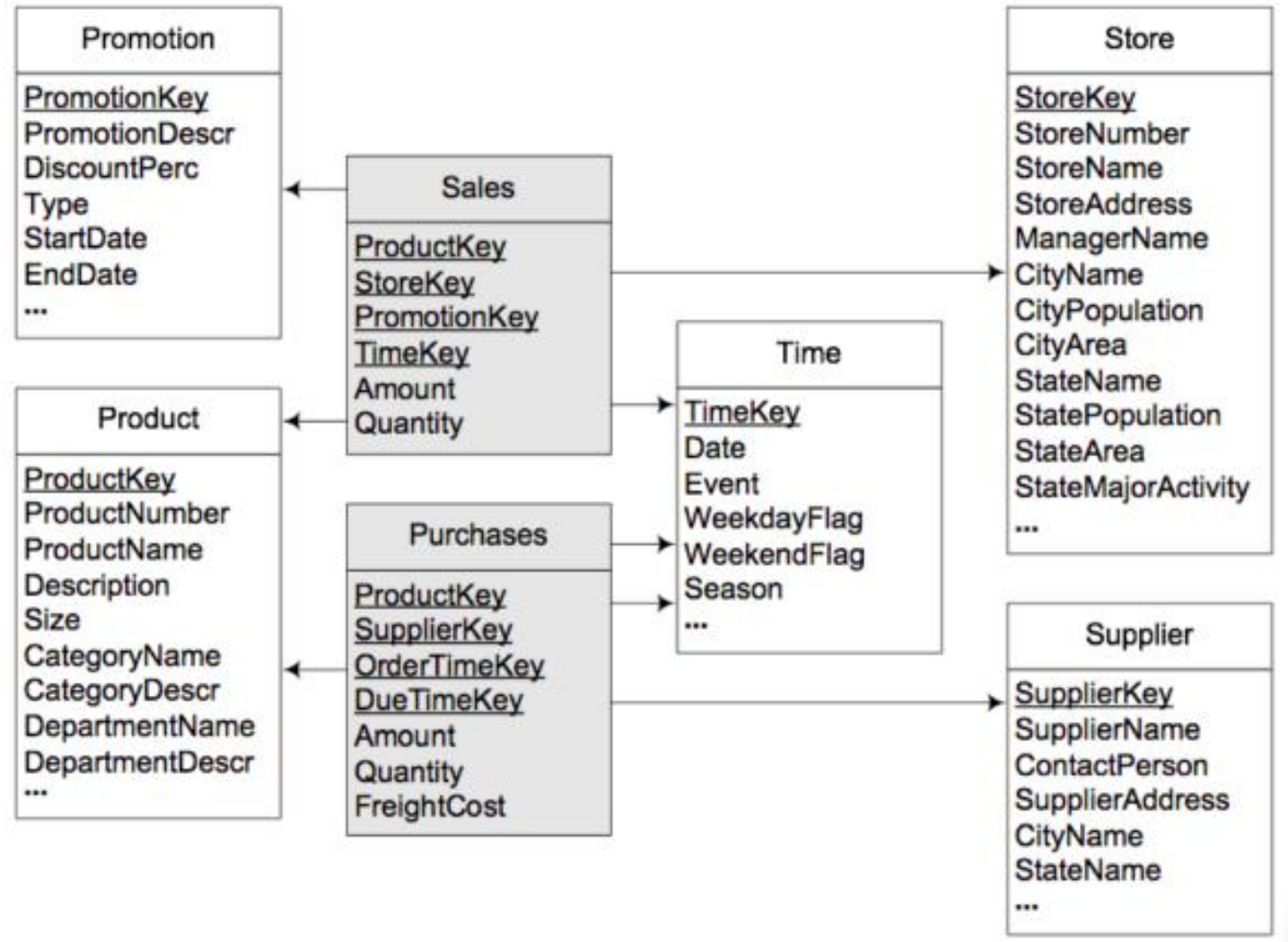
An example of a starflake schema



Relational Data Warehouse Design – Constellation Schema

iv. Constellation schema – Constellation schema consists of multiple fact tables that share dimension tables.

An example of a constellation schema



Conceptual Warehouse Design

Example

Conceptual Warehouse Design

A University consists of various department. Each department is described by a department identification, department name, address of the department, the dean of the department and total number of students registered to the department. Professors are affiliated to departments. The professors may involve in teaching as well as research. Professor is described by professor identification, name, status and the hiring date. The researches the professors done may be in the form of projects which are funded by research funding agency. The research agencies are identified by an agency number and described by the contact person, address of the agency that include city, state and country.

Conceptual Warehouse Design

The project is described by a project identification, project acronym, project name, project type, total person months and total amount.

Each department offers many courses. The course is described by a course identification, course name, credit for theory, credit for exercises, description of the courses, and an URL to the homepage of the course.

Conceptual Warehouse Design

The University want to maintain a data warehouse that contains information about teaching and research activities. On the one hand, the information about teaching activities is related to dimensions department, professor, course, and time, the latter at a granularity of academic semester. Measures for teaching activities are number of hours and number of credits. On the other hand, the information about research activities is related to dimensions professor, funding agency, project, and time, the latter twice for the start date and the end date, both at a granularity of day. In this case, professors are related to the department to which they are affiliated. Measures for research activities are the number of person months and amount.

Create a conceptual schema of a data warehouse domain described above. Use a notation explained to you during the lecture classes, presentation 04 Conceptual Data Warehouse Design.

Identify the entity

Department
DepartmentID
DepartmentName
Address
DeanName
NoStudents

Professor
ProfessorID
ProfessorName
Status
HiringDate

Project
ProjectID
ProjectAcronym
ProjectName
ProjectType
TotalPersonMonths
TotalAmount

FundingAgency
AgencyName
ContactPerson
Address
City
State
Country

Course
CourseID
CourseName
CreditTheory
CreditExercise
Description
Homepage

Identify the hierarchy

Calendar:

- Time
  - Year
  - Month
  - Day

Academic Calendar:

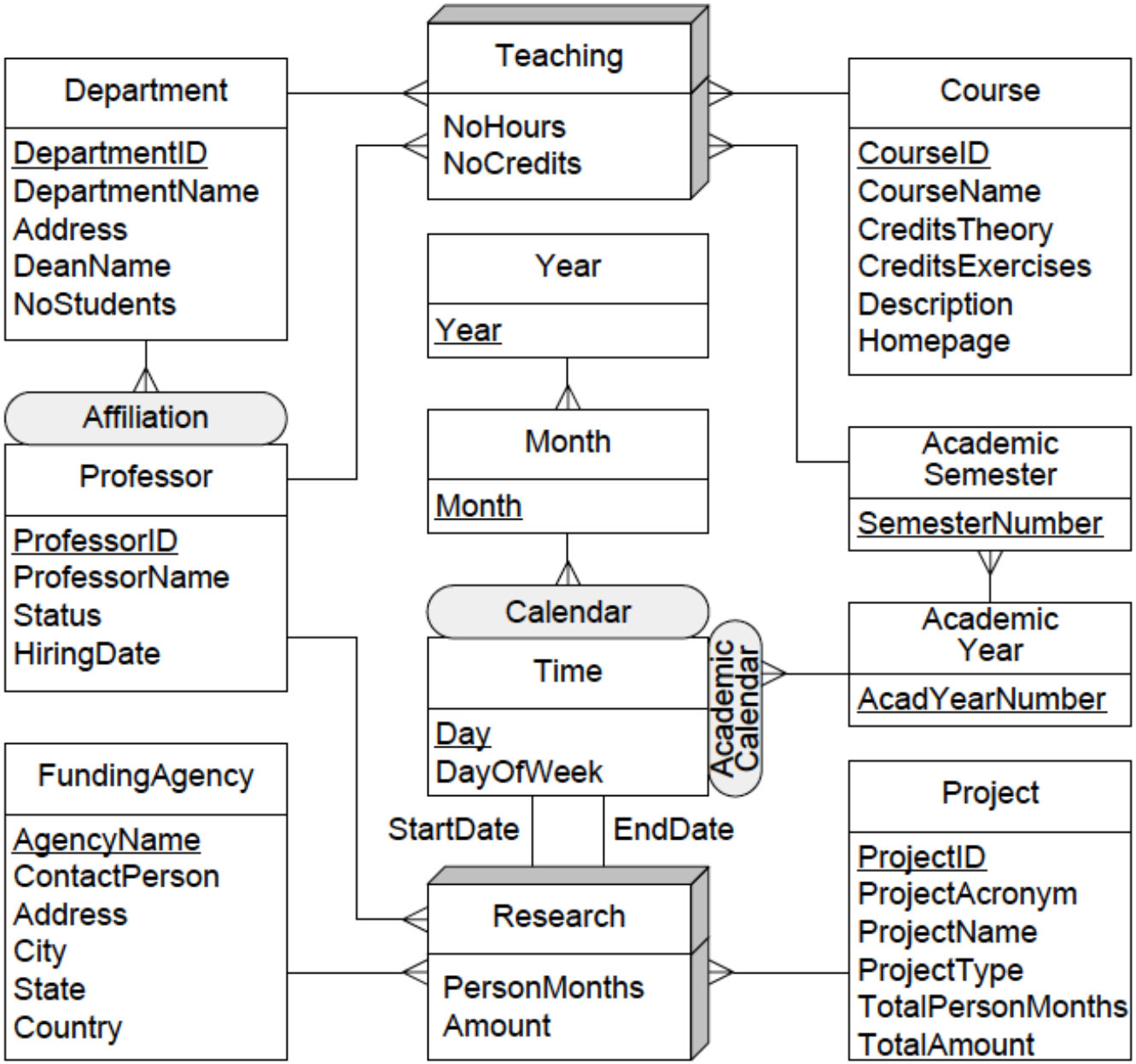
- Time
  - Academic Year
  - Academic Semester

Address:

- Country
- States
- City

Identify the fact table and measures

- Teaching
  - Number of hours
  - Number of credits
- Research
  - Person months
  - Amount

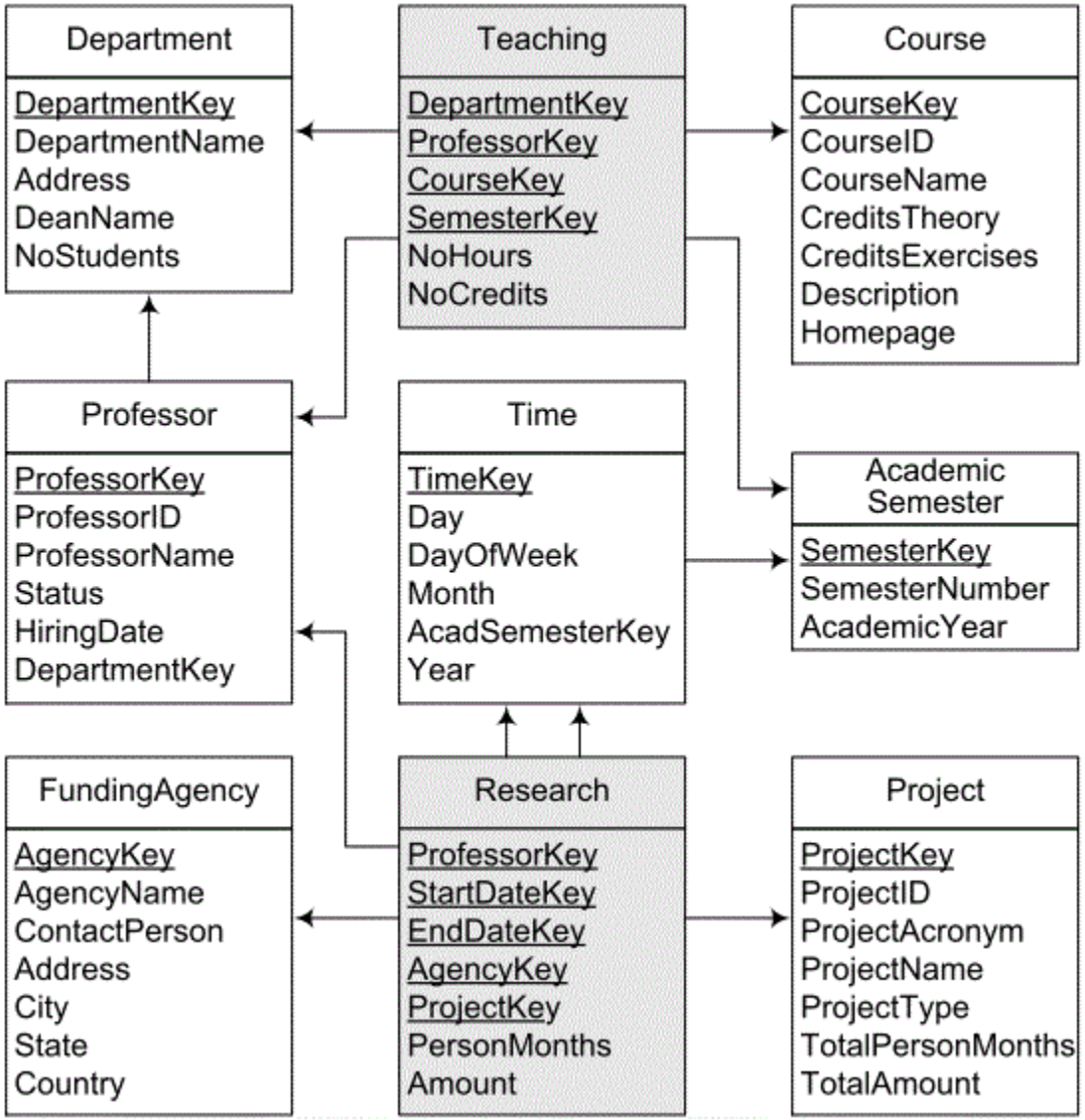

Teaching
NoHours
NoCredits
Department
DepartmentID
DepartmentName
Address
DeanName
NoStudents
Course
CourseID
CourseName
CreditsTheory
CreditsExercises
Description
Homepage
Year
Year
Affiliation
Professor
ProfessorID
ProfessorName
Status
HiringDate
Month
Month
Academic Semester
SemesterNumber
Calendar
Time
Day
DayOfWeek
Academic Calendar
Academic Year
AcadYearNumber
FundingAgency
AgencyName
ContactPerson
Address
City
State
Country
StartDate
EndDate
Research
PersonMonths
Amount
Project
ProjectID
ProjectAcronym
ProjectName
ProjectType
TotalPersonMonths
TotalAmount

Constellation Schema

Constellation Schema

Transform the conceptual schema diagram obtained in the previous activity to a Constellation schema.

Constellation Schema



Constellation Schema

Query

Queries

For the constellation schema shown in the previous slide, write in SQL the queries given below:

a. By department, total number of teaching hours during the academic year 2012 – 2013.

```
SELECT    DepartmentName, SUM(NoHours)
FROM      Teaching T, AcademicSemester S,
          Department D
WHERE     T.SemesterKey = S.SemesterKey
AND       T.DepartmentKey = D.DepartmentKey
AND       AcademicYear = '2012-1023'
GROUP BY  DepartmentName;
```

Queries

- By department, total amount of research projects during the calendar year 2012.

```
SELECT      DepartmentName, SUM(Amount)
FROM        Research R,
            Professor P,
            Department D,
            Time T
WHERE       R.ProfessorKey = D.ProfessorKey
AND         P.DepartmentKey = D.DeprtmentKey
AND         R.StartDateKey = T.TimeKey
AND         Year = '2012'
GROUP BY    DepartmentName;
```

Queries

- By department, total number of professors involved in research projects during the calendar year 2012.

```
SELECT      DepartmentName, Count(ProfessorID)
FROM        Research R,
            Professor P,
            Department D,
            Time T
WHERE       R.ProfessorKey = P.ProfessorKey
AND         P.DepartmentKey = D.DepartmentKey
AND         R.StartDateKey = T.TimeKey
AND         Year = '2012'
GROUP BY    DepartmentName;
```

Queries

- By professor, total number of courses delivered during the academic year 2012-2013.

```
SELECT      DepartmentName, COUNT(CourseID)
FROM        Teaching T,
            AcademicSemester S,
            Department D
WHERE       T.SemesterKey = S.SemesterKey
AND         T.DepartmentKey = D.DepartmentKey
AND         AcademicYear = '2012-2013'
GROUP BY    DepartmentName;
```

Queries

- By department and funding agency, total number of projects started in 2012.

```
SELECT      DepartmentName, AgencyName,
            COUNT(ProfessorID)
FROM        Research R,
            Professor P,
            Department D,
            FundingAgency F,
            Time T
WHERE       R.ProfessorKey = P.ProfessorKey
AND         P.DepartmentKey = D.DepartmentKey
AND         R.StartDateKey = T.TImeKey
AND         Year = '2012'
AND         R.AgencyKey = F.AgencyKey
GROUP BY    DepartmentName, AgencyName;
```