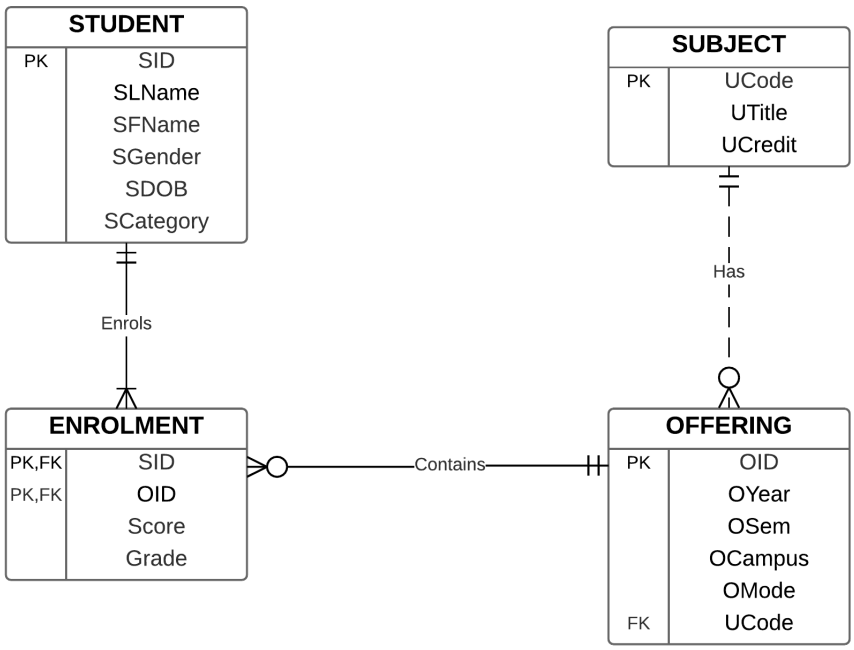
Case Study #1
Student Enrollment

1. The Operational Database

The E/R diagram of the operational database is as follows:



It consists of the following four tables:

```
SQL> desc student;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 SID                                       NOT NULL VARCHAR2(10)
 SLNAME                                    NOT NULL VARCHAR2(20)
 SFNAME                                    NOT NULL VARCHAR2(20)
 SGENDER                                            CHAR(1)
 SDOB                                               DATE
 SCATEGORY                                          NUMBER(5)

SQL> desc subject;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 UCODE                                     NOT NULL VARCHAR2(10)
 UTITLE                                    NOT NULL VARCHAR2(20)
 UCREDIT                                            NUMBER(2)

SQL> desc offering;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 OID                                       NOT NULL NUMBER
 OYEAR                                     NOT NULL NUMBER
 OSEM                                               NUMBER
 OCAMPUS                                            VARCHAR2(20)
 OMODE                                              VARCHAR2(2)
 UCODE                                              VARCHAR2(20)
```

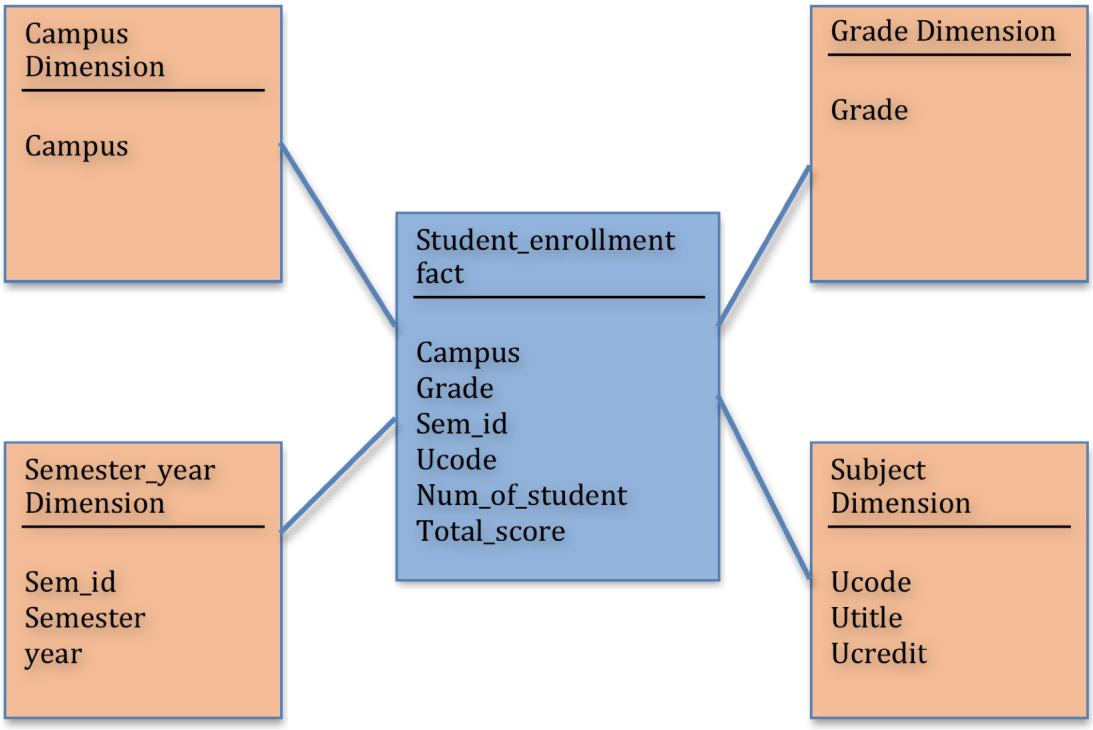
```
SQL> desc enrollment;
 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 OID                                                NUMBER
 SCORE                                              NUMBER
 SID                                                VARCHAR2(20)
 GRADE                                              VARCHAR2(2)
```

The requirements of the data warehouse are to answer the following queries:

1. How many students enrolled in the Database unit offered in the Main campus?
2. What is the total score of students taking the Database unit in the Main campus?
3. How many students enrolled in the Java unit offered in Semester 2, 2009?
4. What is the total score of students taking the Java unit offered in Semester 2, 2009?
5. How many students received HD in the SAP unit offered in Semester 1, 2009?

2. Star Schema

The star schema of of the Student-Enrollment data warehouse is shown as follows.



The fact measures included in the star schema are: (*i*) Number of students, and (*ii*) Total scores.

Four dimensions are included in the star schema, which are (*i*) Campus, (*ii*) Semester-Year, (*iii*) Subject, and (*iv*) Grade.

3. One-attribute Dimensions

The above star schema has one dimension, namely: Campus Dimension; that has only one attribute (also called campus).

The question is whether it would be better to just include the campus attribute as an attribute in the fact, and remove the Campus Dimension altogether.

The answer is that in the real world, there won't be any one-attribute dimension. In this case study, Campus Dimension has only attribute because the case study (e.g. the operational database) is very simple. It is very likely that in the real operational database, there is some information about campus (e.g. campus address, contact number, etc), and consequently, this information will be included as additional attributes in the Campus Dimension.

Even when the operational database does not have information about campus, the management would likely require the data warehouse to include additional information about the dimension. In this case, you would often be required to search for public information on the Internet.

4. Average measure in the Fact

see the next lecture notes