

Practical 1

DIMENSIONAL MODELING FOR BUSINESS PROCESS OF

***"STUDENT ADMISSION FOR GRADUATION AFTER 12TH SCIENCE"
ON
RESHUFFLING***

STEPS TO BE FOLLOWED FOR DIMENSIONAL MODELING:

- 1) Identify the Business Process
- 2) Identify the Grain
- 3) Identify the Dimensions
- 4) Identify the Facts

Step 1

Identifying the business process is the process of determining business process that the data warehouse represents.

Here the data warehouse is going to represent number of admission with average of percentage across branches (Eng-CS, Eng-EC, Med-MMBS, Med-Dental), across type of students (Male-Open, Male-SC, Female-Open, Female-SC), across period (2004-1RS, 2004-2RS, 2005-1RS, 2005-2RS), Where RS means reshuffling.

Step 2

Identifying the Grain is the process of identifying the level of detail of the fact table.

Here the grain is branch wise students with their average percentage per kind of student, per period of admission.

Step 3

Identifying the dimensions for the business process of interest is the process of representing characteristics such as who, what, where, when, how of a measurement.

In this case, the measurement is admissions with average percentage.

Characteristics:

Who is going to be admitted?

Male/Female student having cast OPEN/SC.

When is going to be admitted?

During 1st Reshuffling of year 2004

During 2nd Reshuffling of year 2004

During 1st Reshuffling of year 2005


During 2nd Reshuffling of year 2005.

Where is going to be admitted?

To the category of Engineering/Medical within stream CS/EC or MBBS/Dental accordingly.


Hence, here three dimension tables as per following:

Dimension 1: JD_BRANCH_ADM Table **(WHERE)**
Represents different branch types available for admission.

Field Name	Type	Primary key
Branch_ID	Integer	
Branch_CATEGORY	Varchar2	
Branch_STREAM	Varchar2	


```
create table JD_Branch_ADM (Branch_ID integer not null,Branch_CATEGORY  
varchar2(15),Branch_STREAM varchar2(20),  
PRIMARY KEY(Branch_ID));
```

Dimension 2: JD_FELLOW_ADM Table **(WHO)**
Represents different fellow types admitted.

Field Name	Type	Primary key
Fellow_ID	Integer	
Fellow_GENDER	Varchar2	
Fellow_CATEGORY	Varchar2	

```
create table JD_Fellow_ADM (Fellow_ID integer not null,Fellow_GENDER  
char,Fellow_CATEGORY varchar2(5),  
PRIMARY KEY(Fellow_ID));
```

Dimension 2: JD_PERIOD_ADM Table **(WHEN)**
Represents different period of admission

Field Name	Type	Primary key
Period_ID	Integer	
Period_TYPE	Varchar2	
Period_YEAR	Varchar2	

```
create table JD_Period_ADM (Period_ID integer not null, Period_TYPE.  
integer,Period_YEAR varchar2(5),  
PRIMARY KEY(Duration_ID));
```

Step 4


Identifying the facts is the process of determining measurements for the business process of interest. **(what)**

The facts must be confirm to the grain defined in step2.

Here number of admission with average of percentage are the measurements and hence are the facts which are numeric.

Students admitted, with their average percentage, of same fellow type and branch during specific period.

FACT TABLE SCHEMA: JF_ADM

Field Name	Type	Primary key
Branch_ID	Integer	
Fellow_ID	Integer	
Period_ID	Integer	
Fact_No_Adm	Integer	
Fact_Avg_Per	Integer	

create table JF_ADM (Branch_ID integer references JD_Branch_ADM,
Fellow_ID integer references JD_Fellow_ADM,
Period_ID integer references JD_Period_ADM,
Fact_No_Adm integer,
Fact_Avg_Per integer,primary key (Branch_Id,Fellow_Id,Duration_Id));

→ Notice that the Fact table contains the composite primary key consisting of each dimension table's simple - non-composite primary key.

DATA ENTRY FOR THE DIMENSIONS & FACT TABLE(S)

BRANCH_ID BRANCH_CATEGORY BRANCH_STREAM

1	ENG	CS
2	ENG	EC
3	MED	MBBS
4	MED	DENTAL

FELLOW_ID FELLOW_GENDER FELLO_CATEGORY

1	M	OPEN
2	M	SC
3	F	OPEN
4	F	SC

PERIOD_ID PERIOD_TYPE PERIOD_YEAR

1	1RS	2004
2	2RS	2004
3	1RS	2005
4	2RS	2005

BRANCH_ID FELLOW_ID PERIOD_ID FACT_NO_ADM FACT_AVG_PER

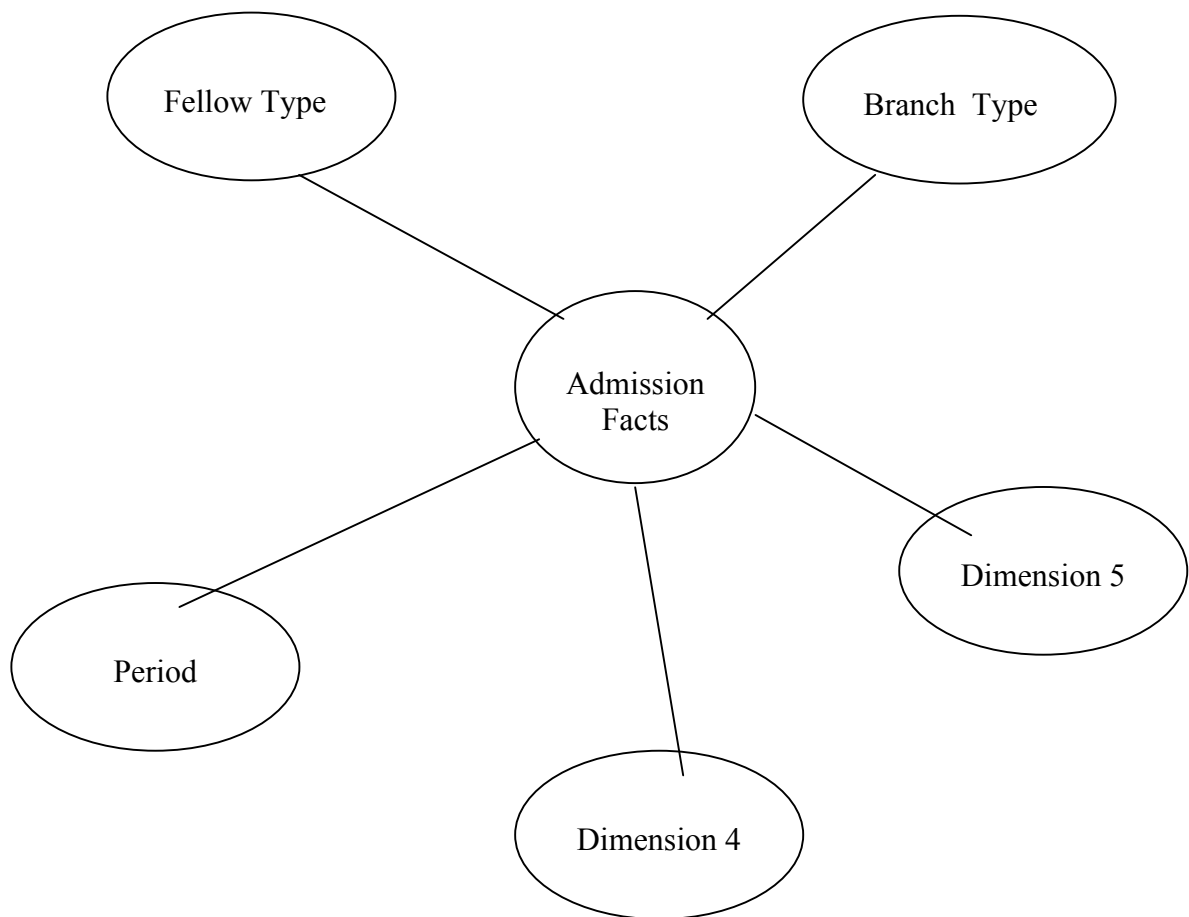
1	1	1	1000	60
1	1	2	1100	70
1	1	3	1200	80
1	1	4	1300	90
1	2	1	1400	60
1	2	2	1500	70
1	2	3	1600	80
1	2	4	1700	90
1	3	1	1800	60
1	3	2	1900	70
1	3	3	2000	80
1	3	4	2100	90
1	4	1	2200	60
1	4	2	2300	70
1	4	3	2400	80
1	4	4	2500	90
2	1	1	2600	60
2	1	2	2700	70
2	1	3	2800	80
2	1	4	2900	90
2	2	1	3000	60

BRANCH_ID	FELLOW_ID	PERIOD_ID	FACT_NO_ADM	FACT_AVG_PER
2	2	2	3100	70
2	2	3	3200	80
2	2	4	3300	90
2	3	1	3400	60
2	3	2	3500	70
2	3	3	3600	80
2	3	4	3700	90
2	4	1	3800	60
2	4	2	3900	70
2	4	3	4000	80
2	4	4	4100	90
3	1	1	4200	60
3	1	2	4300	70
3	1	3	4400	80
3	1	4	4500	90
3	2	1	4600	60
3	2	2	4700	70
3	2	3	4800	80
3	2	4	4900	90
3	3	1	5000	60
3	3	2	5100	70
3	3	3	5200	80
3	3	4	5300	90
3	4	1	5400	60
3	4	2	5500	70
3	4	3	5600	80
3	4	4	5700	90
4	1	1	5800	60
4	1	2	5900	70
4	1	3	6000	80
4	1	4	6100	90
4	2	1	6200	60
4	2	2	6300	70
4	2	3	6400	80
4	2	4	6500	90
4	3	1	6600	60
4	3	2	6700	70
4	3	3	6800	80
4	3	4	6900	90
4	4	1	7000	60
4	4	2	7100	70
4	4	3	7200	80
4	4	4	7300	90

64 rows selected.

Here 4 types are considered in each of the dimensions. This limits total number of records in the fact table to 64. (4 * 4 *4).

Star Schema of the business process



Above drawing looks like a star. Hence it is known as star schema.

Practical 2

OLAP OPERATIONS IN MULTIDIMENSIONAL MODEL FOR BUSINESS PROCESS OF

*“STUDENT ADMISSION FOR GRADUATION AFTER 12TH SCIENCE”
ON
RESHUFFLING*

Basic OLAP Operations:

- Slice
- Dice
- Roll-Up
- Drill Down

Definition:

Slice:

Performing a selection on one dimension of the given cube is known as Slicing.

Dice:

Defines a subcube by performing a selection on two or more dimensions.

Roll- Up:

Performs aggregation on a data cube, either by climbing up a concept hierarchy for a dimension or by dimension reduction.

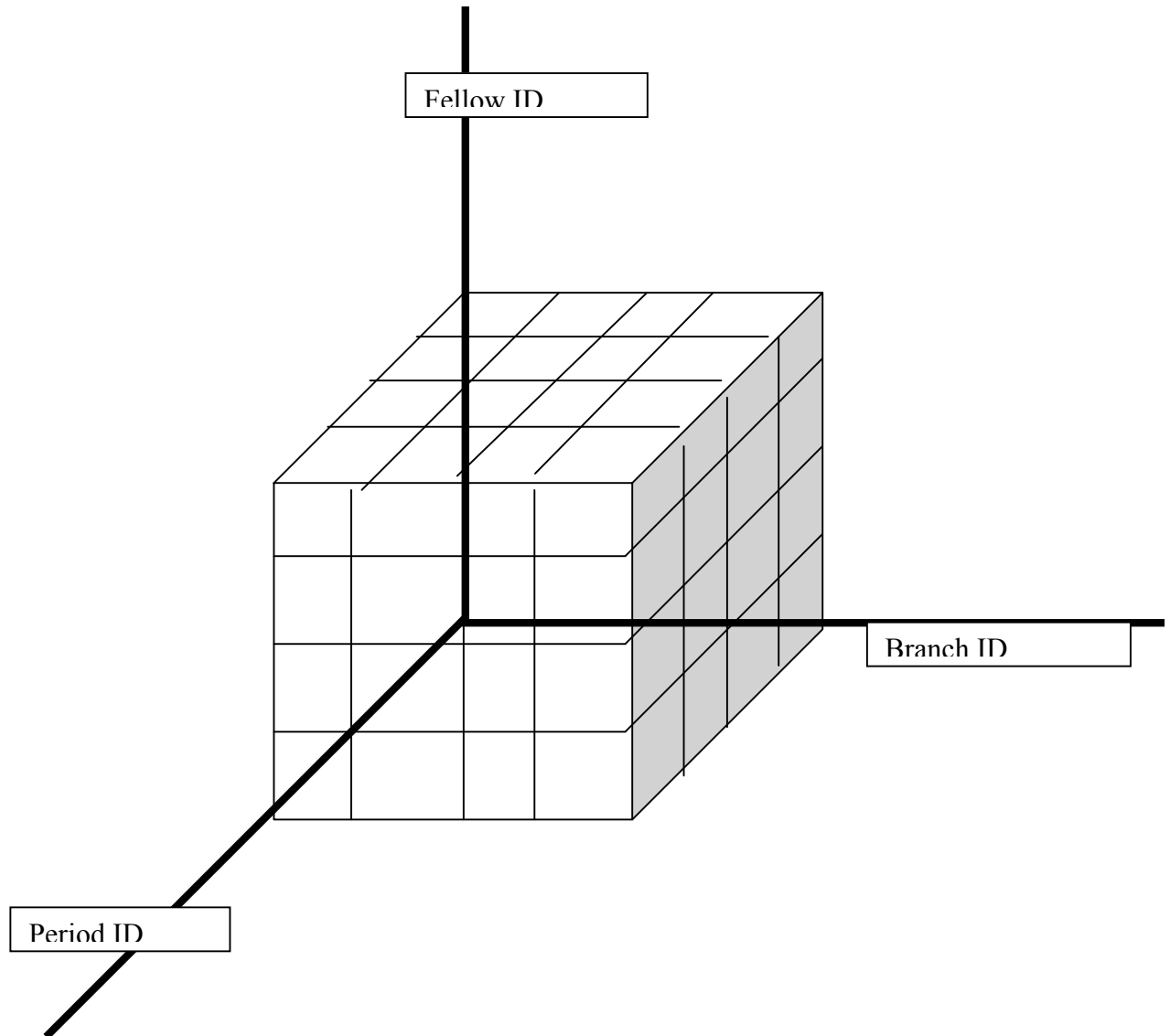
Drill-Down:

Achieved by either stepping down a concept hierarchy for a dimension or introducing additional dimensions. Reverse of roll-up.

Other OLAP Operations:

Rotate, Drill across, Drill through.

OLAP CUBE



1) SLICE

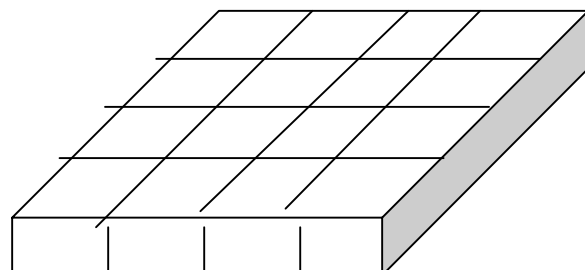
-> **Find and display number of students only those Male & Open type of fellow with their percentage average for each period and for each type of branch separately.**

-> Select Fact_No_Adm, Fact_Avg_Per
from JD_Branch_ADM B,JD_Fellow_ADM F,JD_Period_ADM D,JF_ADM FACT
where (
B.Branch_Id = FACT.Branch_Id and
F.Fellow_Id = FACT.Fellow_Id and
D.Period_Id = FACT.Period_Id and
F.Fellow_Gender = 'M' and
F.Fellow_Category = 'OPEN');

Resulting slice:

FACT_NO_ADM FACT_AVG_PER

1000	60
1100	70
1200	80
1300	90
2600	60
2700	70
2800	80
2900	90
4200	60
4300	70
4400	80
4500	90
5800	60
5900	70
6000	80
6100	90



16 rows selected.

2. DICE

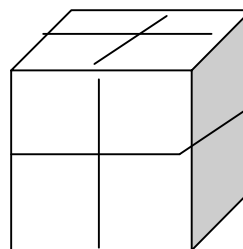
➔ **Find number of students with their average percentage which satisfies following:**

- **Belong to Branch Category of Engineering**
- **Male**
- **Have reshuffled during year 2004**

```
Select Fact_No_Adm, Fact_Avg_Per
from JD_Branch_ADM B,JD_Fellow_ADM F,JD_Period_ADM D,JF_ADM FACT
where (
B.Branch_Id = FACT.Branch_Id and
F.Fellow_Id = FACT.Fellow_Id and
D.Period_Id = FACT.Period_Id and
(F.Fellow_Id = 1 or F.Fellow_Id = 2) and
(B.Branch_Id = 1 or B.Branch_Id = 2) and
(D.Period_Id = 1 or D.Period_Id = 2));
```

FACT_NO_ADM FACT_AVG_PER

3100	70
3000	60
1500	70
1400	60
2700	70
2600	60
1100	70
1000	60



8 rows selected.

Other Way of finding the same dice:

```
Select Fact_No_Adm, Fact_Avg_Per
from JD_Branch_ADM B,JD_Fellow_ADM F,JD_Period_ADM D,JF_ADM FACT
where (
B.Branch_Id = FACT.Branch_Id and
F.Fellow_Id = FACT.Fellow_Id and
D.Period_Id = FACT.Period_Id and
F.Fellow_Gender = 'M' and
B.Branch_Category = 'ENG' and
D.Period_Year = 2004);
```

FACT_NO_ADM FACT_AVG_PER

1000	60
1100	70

1400	60
1500	70
2600	60
2700	70
3000	60
3100	70

8 rows selected.

Roll UP:

Find number of total students admitted to branch category of Engineering.

```

Select sum(Fact_No_Adm)
from JD_Branch_ADM B,JD_Fellow_ADM F,JD_Period_ADM D,JF_ADM FACT
where
B.Branch_Id = FACT.Branch_Id and
F.Fellow_Id = FACT.Fellow_Id and
D.Period_Id = FACT.Period_Id and
B.Branch_Category = 'ENG'
group by B.Branch_Category;

```

Result of the query:
SUM(FACT_NO_ADM)

```

-----
      81600

```