**Product Dimension Population**

Consider following table’s product related oltp tables. This is in normalized model. Typically when you work as an ETL developer, you have to use these three table as the source and populate the dimension table.

**DDL Statements to create the source tables**

create table prod\_family (prod\_fam\_id integer, prod\_fam\_desc varchar(30), prod\_fam\_long\_desc varchar(30), crt\_date date, upd\_date date);

insert into prod\_family (1,'Electronics','Electronics and Gadgets',sysdate - 20, null);

create table prod\_cat (prod\_cat\_id integer not null, prod\_fam\_id integer,prod\_cat\_desc varchar(30), prod\_cat\_long\_desc varchar(60), crt\_dt date, mod\_dt date);

insert into prod\_cat values(100,1,'Computers','Computers and Supporting Devices',sysdate - 10, null);

|  |
| --- |
| create table product (prod\_id integer, prod\_name varchar(30), prod\_price integer, prod\_cat\_id integer, crt\_dt date, upd\_dt date); |
| insert into product values(100,'Monitor',8000,10,sysdate,null); |
| insert into product values(101,'Keyboard',800,10,sysdate,null); |

|  |  |  |
| --- | --- | --- |
| **prod\_family** |  |  |
| prod\_fam\_id |  |  |
| prod\_fam\_desc |  |  |
| prod\_fam\_long\_desc |  |  |
| crt\_date |  |  |
| mod\_date | **prod\_cat** |  |
|  | prod\_cat\_id |  |
|  | prod\_cat\_desc | **product** |
|  | prod\_cat\_long\_desc | prod\_id |
|  | prod\_fam\_id | prod\_name |
|  | crt\_date | prod\_price |
|  | mod\_date | prod\_cat\_id |
|  |  | crt\_date |
|  |  | mod\_date |

Use the above set of three tables and populate the prod\_dim table. Prod\_sur\_id is the PK for this dimension. When there is a change in the source system, you should insert a record into product dim table. This way we know the current data and the data which got changed in the past.

**DDL Statements to create the source tables**

Create table prod\_dim (prod\_sur\_id number(8) primary key, prod\_id number(8), prod\_nm varchar(30), prod\_cat varchar(30), prod\_family varchar(30), prod\_price number(7,2), start\_Date date, end\_date date);

|  |
| --- |
| PROD\_DIM |
| PROD\_SUR\_ID |
| PROD\_ID |
| PROD\_NM |
| PROD\_CAT |
| PROD\_FAMILY |
| PROD\_PRICE |
| START\_DATE |
| END\_DATE |

**Points to remember**

In DW, dimension tables are used to store the categorical / dimensional data for analysis. There are three types of dimension tables. Those are TYPE1, TYPE2 and TYPE3. TYPE1 / SCD1 is used to store only the current data in DW, TYPE2 / SCD2 is used to store the current data + the changes happen in the product data, TYPE3 / SCD3 is used to store the current and previous version of the data.