Aim : **Data Warehouse Construction a) Real life Problem to be defined for Warehouse Design b) Construction of star schema and snow flake schema c) ETL Operations.**

THEORY:

Problem Statement

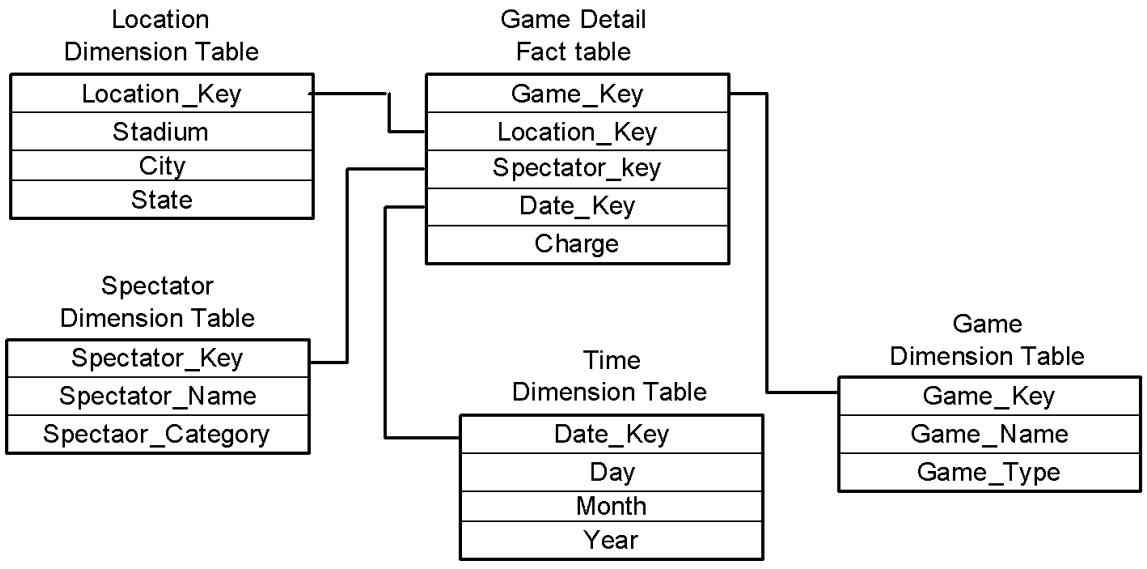
GAME COMPANY (Write the problem statement here)

TYPES OF TABLE:

FACT TABLE:
– (Game_Key,Location_Key,Spectator_Key,Date_Key,Charge)

DIMENSION TABLES:
– Location (Location_Key)
– Spectator (Spectator_Key)
-- Time (Date_Key,Day,Month,Year)
--Game (Game_Key,Game_Name,Game_Type)

Star Schema



ETL PROCESS(Theory about ETL)

QUERIES:

A) Implementation of the STAR Schema

SPECTATOR TABLE:

```
SQL> create table spectator(
2       spec_id number(3) primary key,
3       spec_name varchar(15),
4       category varchar(15) constraint c check(category in('student','senior','adult')),
5       charge number(5));

Table created.

SQL> insert into spectator values(1,'mayuri',student,500);

1 row created.

SQL> select * from spectator;

   SPEC_ID  SPEC_NAME CATEGORY CHARGE
---------- ---------- ---------- ----------
         1 mayuri     student    500
         2 rucha      senior     600
         3 shraddha   adult      1000
         4 sharvari   student    500
         5 pallavi    senior     600
         6 chandrika  student    500
         7 nikita     adult      1000

7 rows selected.
```

LOCATION TABLE:

```
SQL> create table location(
2       loc_id number primary key,
```

```
3       stadium varchar(15),
4       city varchar(15),
5       state varchar(15));

Table created.


SQL> select * from location;

   LOC_ID        STADIUM      CITY          STATE
--------------   ---------------   ----------------  ---------------
        1            vankhede       mumbai      maharashtra
        2            d y patil         vashi         maharashtra
        3            abc               cityxyz       stateabc
        4            pqr               citypqr       statepqr

4 rows selected.
```

GAME TABLE:

```
SQL> create table game(
2       game_id number(3) primary key,
3       game_name varchar(15),
4       game_type varchar(15));

Table created.

SQL> select * from game;

  GAME_ID   GAME_NAME   GAME_TYPE
---------------   ------------------   ---------------
        1            cricket             outdoor
        2            boxcricket       indoor
        3             basketball      outdoor
        4            football          outdoor
        5             tennis           indoor

5 rows selected.
```

G_DATE TABLE:

```
SQL> create table g_date(
2       dt_id number(3) primary key,
```

```
3       day varchar(15),
4       month varchar(15),
5       year varchar(15));

Table created.

SQL> select * from g_date;

    DT_ID      DAY          MONTH        YEAR
---------- ---------- ---------- ----------
        1          2          feb          2004
        2          6          march        2004
        3          20         may          2010
        4          13         july         2006

4 rows selected.
```

GAME_STAT TABLE:

```
SQL> create table game_stat(spec_id number(3) references spectator(spec_id),
  2     loc_id number references location(loc_id),game_id number(3) references game(game_id),
  3     dt_id number(3) references g_date(dt_id),
  4     charge number(5));

Table created.

SQL> insert into game_stat values(1,2,3,4,500);

1 row created.

SQL> select * from game_stat;

   SPEC_ID   LOC_ID   GAME_ID  DT_ID   CHARGE
---------- ---------- ---------- ---------- ----------
        1          2          3          4        500
        3          2          4          4       1000
        2          3          4          2        600
        5          1          3          4        500

4 rows selected.
```

QUERY on STAR SCHEMA:

```
SQL> select sum(gs.charge),spec.category,l.stadium,d.year
  2  from game_stat gs,spectator spec,location l,g_date d
  3  where gs.spec_id=spec.spec_id and
  4   gs.dt_id=d.dt_id and
  5    gs.dt_id=d.dt_id
  6   group by spec.category,l.stadium,d.year
  7   having spec.category='student'
  8   and d.year='2004'
  9   and l.stadium='vankhede';

no rows selected.
```

B) Implementation of ETL

CREATING TABLE SAELST:

```
SQL> create table salest
  2  (s_person varchar2(15),
  3  item_sold number(5),
  4  price number(10),
  5  s_dt date,
  6  item_category varchar2(10),
  7  num_item_sold number(10));

Table created.
```

INSERTING VALUES INTO SALEST:

```
SQL> insert into salest values('Mayuri','2','12000','1-feb-2012','hardware','5');
1 row created.

SQL> select * from salest;

S_PERSON     ITEM_SOLD   PRICE       S_DT         ITEM_CATEG  NUM_ITEM_SOLD
---------------  ----------------  ------------  -------------  ---------------  ----------------------
Mayuri                2          12000      01-FEB-12     hardware          5
```

Chandrika	1	1500	02-FEB-12	software	3
Pallavi	1	15000	02-FEB-12	hardware	3
Shraddha	2	10000	01-FEB-12	hardware	4
Shweta	2	8000	01-FEB-12	software	7
Siya	3	6000	01-FEB-12	software	10
Sharvari	1	2500	01-FEB-12	software	5
Nikki	3	5000	02-FEB-12	hardware	13
Rucha	2	7000	2-FEB-12	hardware	2
Neha	3	7500	02-FEB-12	hardware	4

10 rows selected.

CREATING TABLE SAELS_DETAILS:

```
SQL> create table sales_details
  2     (sales_person varchar(30),
  3     item_name varchar(30),
  4    item_price number,
  5     sales_date date,
  6    branch varchar(30),
  7      no_of_item_sold number);
```

Table created.

INSERTING VALUES INTO SALES_DETAILS:

```
SQL> insert into sales_details values('Mayuri','HardDisk',12000,'1-feb-2012','Thane',2);
```

1 row created.

```
SQL> select * from sales_details;
```

SALES_PERSON	ITEM_NAME	ITEM_PRICE	S ALES_DAT	BRANCH	NO_OF_ITEM_SOLD
Mayuri	HardDisk	12000	01-FEB-12	Thane	2
Chandrika	AntiVirus	1500	02-FEB-12	CST	3
Pallavi	CPU	1500	02-FEB-12	Thane	2
Shraddha	RAM	10000	01-FEB-12	Andheri	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	4
Rucha	AntiVirus	1500	11-MAR-12	CST	2

6 rows selected.

EXTRACTION:

CREATING TABLE SAELS_RECORDS:

```
SQL>  create table sales_records
  2    (supplier varchar2(15),
  3    item_name varchar2(10),
  4     price number(10),
  5     sales_dt date,
  6     branch varchar2(12),
  7     item_category varchar2(10),
  8     num_item_sold number(6));
```

Table created.

INSERTING VALUES INTO SAELS_RECORDS:

```
SQL> insert into sales_records
  2     (select s_person,item_category,price,s_dt,null,
  3    item_category,num_item_sold from salest);
```

10 rows created.

```
SQL> select * from sales_records;
```

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD
Mayuri	hardware	12000	01-FEB-12		hardware	5
Chandrika	software	1500	02-FEB-12		software	3
Pallavi	hardware	15000	02-FEB-12		hardware	3
Shraddha	hardware	10000	01-FEB-12		hardware	4
Shweta	software	8000	01-FEB-12		software	7
Siya	software	6000	01-FEB-12		software	10
Sharvari	software	2500	01-FEB-12		software	5
Nikki	hardware	5000	02-FEB-12		hardware	13
Rucha	hardware	7000	02-FEB-12		hardware	2
Neha	hardware	7500	02-FEB-12		hardware	4

10 rows selected.

```
SQL>  insert into sales_records
  2     (select sales_person,item_name,item_price,sales_date,branch,
  3    null,no_of_item_sold from sales_details);

6 rows created.


SQL> select * from sales_records;

SUPPLIER      ITEM_NAME  PRICE      SALES_DT  BRANCH    ITEM_CATEG  NUM_ITEM_SOLD
---------------  ----------------- -----------  ------------  ------------ ------------------ -------------
Mayuri        hardware     12000    01-FEB-12              hardware          5
Chandrika     software     1500     02-FEB-12              software          3
Pallavi       hardware     15000    02-FEB-12              hardware          3
Shraddha      hardware     10000    01-FEB-12              hardware          4
Shweta        software     8000     01-FEB-12              software          7
Siya          software     6000     01-FEB-12              software          10
Sharvari      software     2500     01-FEB-12              software          5
Nikki         hardware     5000     02-FEB-12              hardware          13
Rucha         hardware     7000     02-FEB-12              hardware          2
Neha          hardware     7500     02-FEB-12              hardware          4
Mayuri        HardDisk     12000    01-FEB-12  Thane                         2
Chandrika     AntiVirus    1500     02-FEB-12  CST                           3
Pallavi       CPU          1500     02-FEB-12  Thane                         2
Shraddha      RAM          10000    01-FEB-12  Andheri                       5
Sharvari      HardDisk     12000    11-FEB-12  Thane                         4
Rucha         AntiVirus    1500     11-MAR-12  CST                           2

16 rows selected.
```

TRANSFORMATION:

INSERTING VALUES INTO SAELS_DETAILS:

```
SQL> insert into sales_details
values('Gargi','HardDisk',5000,to_date('2012-jan-03','yyyy-mm-dd'),'mumbai',2);

1 row created.

SQL> select * from sales_details;
```

SALES_PERSON	ITEM_NAME	ITEM_PRICE	SALES_DAT	BRANCH	NO_OF_ITEM_SOLD
Mayuri	HardDisk	12000	01-FEB-12	Thane	2
Chandrika	AntiVirus	1500	02-FEB-12	CST	3
Pallavi	CPU	1500	02-FEB-12	Thane	2
Shraddha	RAM	10000	01-FEB-12	Andheri	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	4
Rucha	AntiVirus	1500	11-MAR-12	CST	2
Gargi	HardDisk	5000	03-JAN-12	mumbai	2

7 rows selected.

UPDATING VALUES OF SALES_RECORDS:

```
SQL> update sales_records set branch='unknown' where branch is null;

10 rows updated.

SQL> select * from sales_records;
```

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD
Mayuri	hardware	12000	01-FEB-12	unknown	hardware	5
Chandrika	software	1500	02-FEB-12	unknown	software	3
Pallavi	hardware	15000	02-FEB-12	unknown	hardware	3
Shraddha	hardware	10000	01-FEB-12	unknown	hardware	4
Shweta	software	8000	01-FEB-12	unknown	software	7
Siya	software	6000	01-FEB-12	unknown	software	10
Sharvari	software	2500	01-FEB-12	unknown	software	5
Nikki	hardware	5000	02-FEB-12	unknown	hardware	13
Rucha	hardware	7000	02-FEB-12	unknown	hardware	2
Neha	hardware	7500	02-FEB-12	unknown	hardware	4
Mayuri	HardDisk	12000	01-FEB-12	Thane		2
Chandrika	AntiVirus	1500	02-FEB-12	CST		3

Pallavi	CPU	1500	02-FEB-12	Thane	2
Shraddha	RAM	10000	01-FEB-12	Andheri	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	4
Rucha	AntiVirus	1500	11-MAR-12	CST	2

16 rows selected.

```
SQL> update sales_records set item_category='software' where item_name='AntiVirus';

2 rows updated.

SQL> update sales_records set item_category='hardware' where item_name in
('CPU','RAM','HardDisk');

4 rows updated.

SQL> select * from sales_records;
```

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD
Mayuri	hardware	12000	01-FEB-12	unknown	hardware	5
Chandrika	software	1500	02-FEB-12	unknown	software	3
Pallavi	hardware	15000	02-FEB-12	unknown	hardware	3
Shraddha	hardware	10000	01-FEB-12	unknown	hardware	4
Shweta	software	8000	01-FEB-12	unknown	software	7
Siya	software	6000	01-FEB-12	unknown	software	10
Sharvari	software	2500	01-FEB-12	unknown	software	5
Nikki	hardware	5000	02-FEB-12	unknown	hardware	13
Rucha	hardware	7000	02-FEB-12	unknown	hardware	2
Neha	hardware	7500	02-FEB-12	unknown	hardware	4
Mayuri	HardDisk	12000	01-FEB-12	Thane	hardware	2
Chandrika	AntiVirus	1500	02-FEB-12	CST	software	3
Pallavi	CPU	1500	02-FEB-12	Thane	hardware	2
Shraddha	RAM	10000	01-FEB-12	Andheri	hardware	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	hardware	4
Rucha	AntiVirus	1500	11-MAR-12	CST	software	2

16 rows selected.

```
SQL> update sales_records set num_item_sold=(select avg(num_item_sold)
 2      from sales_records)  where item_name='AntiVirus';
```

2 rows updated.

SQL> select * from sales_records;

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD
Mayuri	hardware	12000	01-FEB-12	unknown	hardware	5
Chandrika	software	1500	02-FEB-12	unknown	software	3
Pallavi	hardware	15000	02-FEB-12	unknown	hardware	3
Shraddha	hardware	10000	01-FEB-12	unknown	hardware	4
Shweta	software	8000	01-FEB-12	unknown	software	7
Siya	software	6000	01-FEB-12	unknown	software	10
Sharvari	software	2500	01-FEB-12	unknown	software	5
Nikki	hardware	5000	02-FEB-12	unknown	hardware	13
Rucha	hardware	7000	02-FEB-12	unknown	hardware	2
Neha	hardware	7500	02-FEB-12	unknown	hardware	4
Mayuri	HardDisk	12000	01-FEB-12	Thane	hardware	2
Chandrika	AntiVirus	1500	02-FEB-12	CST	software	5
Pallavi	CPU	1500	02-FEB-12	Thane	hardware	2
Shraddha	RAM	10000	01-FEB-12	Andheri	hardware	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	hardware	4
Rucha	AntiVirus	1500	11-MAR-12	CST	software	5

16 rows selected.

```
SQL>  update sales_records set num_item_sold=(select avg(num_item_sold)
2       from sales_records) where item_name in('CPU','RAM','HardDisk');
```

4 rows updated.

SQL> select * from sales_records;

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD

Mayuri	hardware	12000	01-FEB-12	unknown	hardware	5
Chandrika	software	1500	02-FEB-12	unknown	software	3
Pallavi	hardware	15000	02-FEB-12	unknown	hardware	3
Shraddha	hardware	10000	01-FEB-12	unknown	hardware	4
Shweta	software	8000	01-FEB-12	unknown	software	7
Siya	software	6000	01-FEB-12	unknown	software	10
Sharvari	software	2500	01-FEB-12	unknown	software	5
Nikki	hardware	5000	02-FEB-12	unknown	hardware	13
Rucha	hardware	7000	02-FEB-12	unknown	hardware	2
Neha	hardware	7500	02-FEB-12	unknown	hardware	4
Mayuri	HardDisk	12000	01-FEB-12	Thane	hardware	5
Chandrika	AntiVirus	1500	02-FEB-12	CST	software	5
Pallavi	CPU	1500	02-FEB-12	Thane	hardware	5
Shraddha	RAM	10000	01-FEB-12	Andheri	hardware	5
Sharvari	HardDisk	12000	11-FEB-12	Thane	hardware	5
Rucha	AntiVirus	1500	11-MAR-12	CST	software	5

LOADING:

CREATING VIEW CATALOG1:

```
SQL>  create view catalog1 as (select sum(price*num_item_sold) as yearly_sale,
  2      to_char(sales_dt,'yyyy') as year from sales_records group by to_char(sales_dt,'yyyy'));

View created.

SQL> select * from catalog1;

YEARLY_SALE YEAR
----------- ----
    519500 2012

SQL> create view catalog2 as (select sum(price*num_item_sold) as yearly_sale,
  2      to_char(sales_dt,'yy') as year from sales_records group by to_char(sales_dt,'yy'));

View created.


SQL> select * from catalog2;

YEARLY_SALE YE
```

---------- --
519500 12

SQL> create view category as(select * from sales_records where item_category='software');

View created.

SQL> select * from category;

SUPPLIER	ITEM_NAME	PRICE	SALES_DT	BRANCH	ITEM_CATEG	NUM_ITEM_SOLD
Chandrika	software	1500	02-FEB-12	unknown	software	3
Shweta	software	8000	01-FEB-12	unknown	software	7
Siya	software	6000	01-FEB-12	unknown	software	10
Sharvari	software	2500	01-FEB-12	unknown	software	5