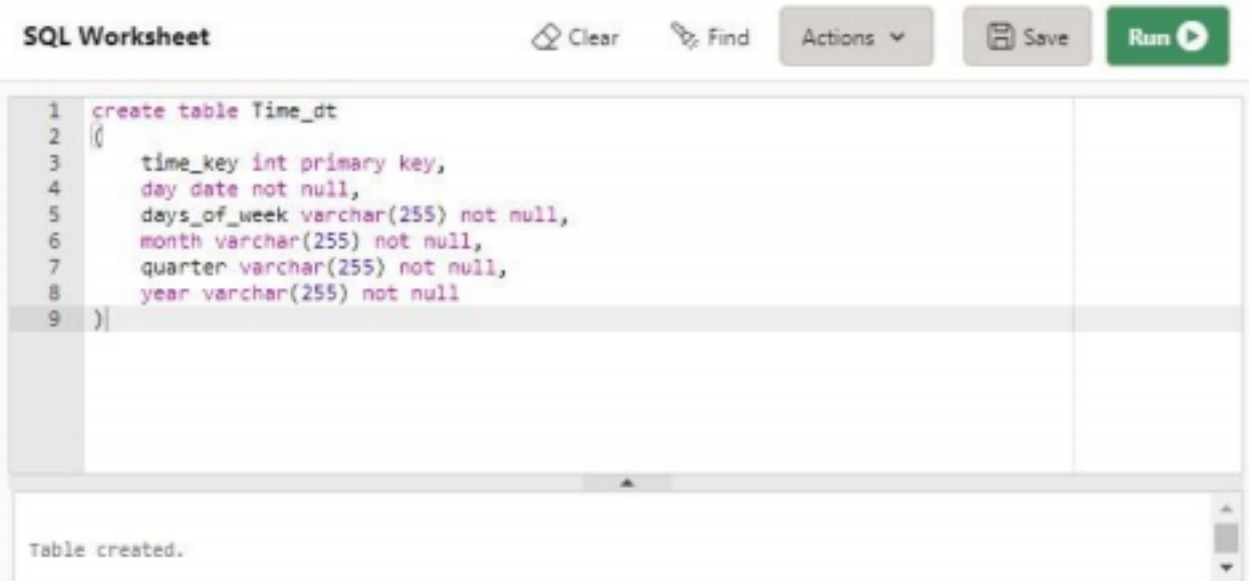
Experiment 1: Datawarehouse and Schemas

Problem Statement:

Creation of Data Warehouse for a hospital having 3 dimensions.
1) Doctor 2) Patient 3) Time and two measures 1) Charge and 2) Count, where charge is for that a doctor charges a patient for a visit. Draw the Star Schema and Snowflake Schema for the same.

Output:

1. Creating Dimension Tables – Fact, Time, Patient and Doctor



SQL Worksheet

Clear Find Actions Save Run

```
1  create table Fact_table
2  (
3      time_key int references Time_dt(time_key),
4      patient_id int references Patient_dt(patient_id),
5      doctor_id int references Doctor(doctor_id),
6      charge int not null,
7      count int not null
8  )
```

```
Table created.
```

2. Inserting Values in Fact, Doctor, Patient and Time dimension tables

SQL Worksheet

Clear Find Actions Save Run

```
1  insert into Time_dt values (101, DATE '2019-08-04' , 'Wednesday' , 'August' , 'Q3' , '2019')
2  insert into Time_dt values (102, DATE '2020-03-20' , 'Saturday' , 'March' , 'Q1' , '2020')
3  insert into Time_dt values (103, DATE '2018-06-12' , 'Monday' , 'June' , 'Q2' , '2018')
4  insert into Time_dt values (104, DATE '2019-11-28' , 'Friday' , 'November' , 'Q4' , '2019')
5  insert into Time_dt values (105, DATE '2020-09-16' , 'Thursday' , 'September' , 'Q3' , '2020')
6
7  select * from Time_dt
```

TIME_KEY	DAY	DAYS_OF_WEEK	MONTH	QUARTER	YEAR
103	12-JUN-18	Monday	June	Q2	2018
104	28-NOV-19	Friday	November	Q4	2019
101	04-AUG-19	Wednesday	August	Q3	2019
102	20-MAR-20	Saturday	March	Q1	2020
105	16-SEP-20	Thursday	September	Q3	2020

Download CSV

5 rows selected.

SQL Worksheet

Clear Find Actions Save Run

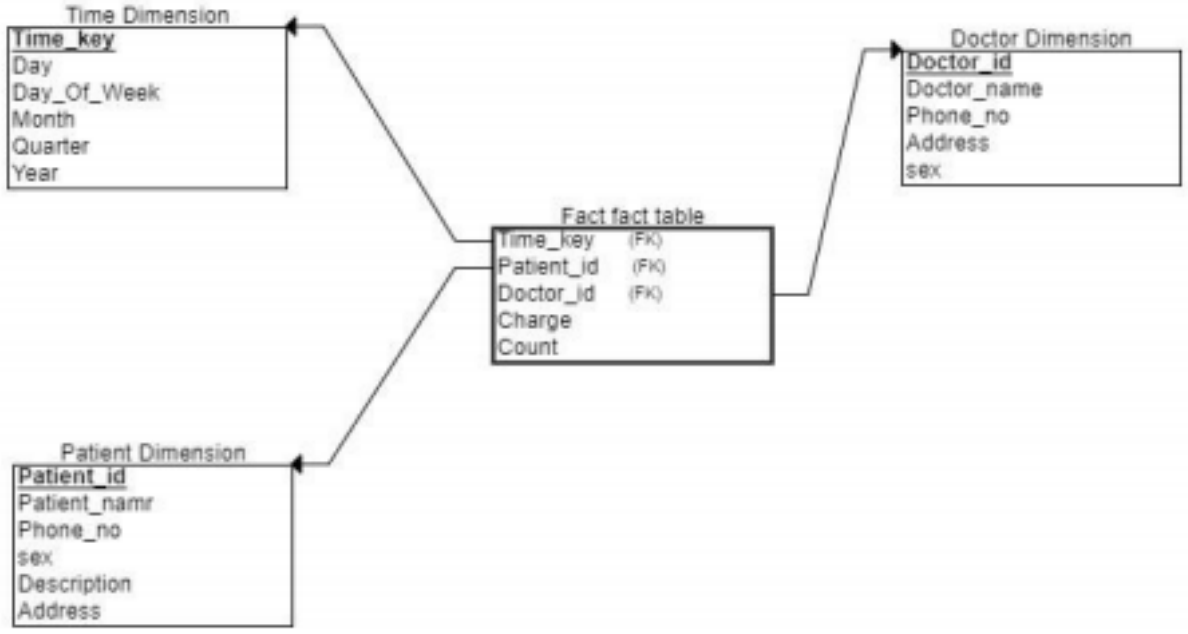
```
1  insert into Fact_table values (101, 203, 502, 1000, 3);
2  insert into Fact_table values (102, 205, 504, 1500, 2);
3  insert into Fact_table values (103, 201, 505, 2500, 5);
4  insert into Fact_table values (104, 202, 501, 2000, 4);
5  insert into Fact_table values (105, 204, 503, 4200, 1);
6  Select * from Fact_table
7
8
```

TIME_KEY	PATIENT_ID	DOCTOR_ID	CHARGE	COUNT
101	203	502	1000	3
102	205	504	1500	2
103	201	505	2500	5
104	202	501	2000	4
105	204	503	4200	1

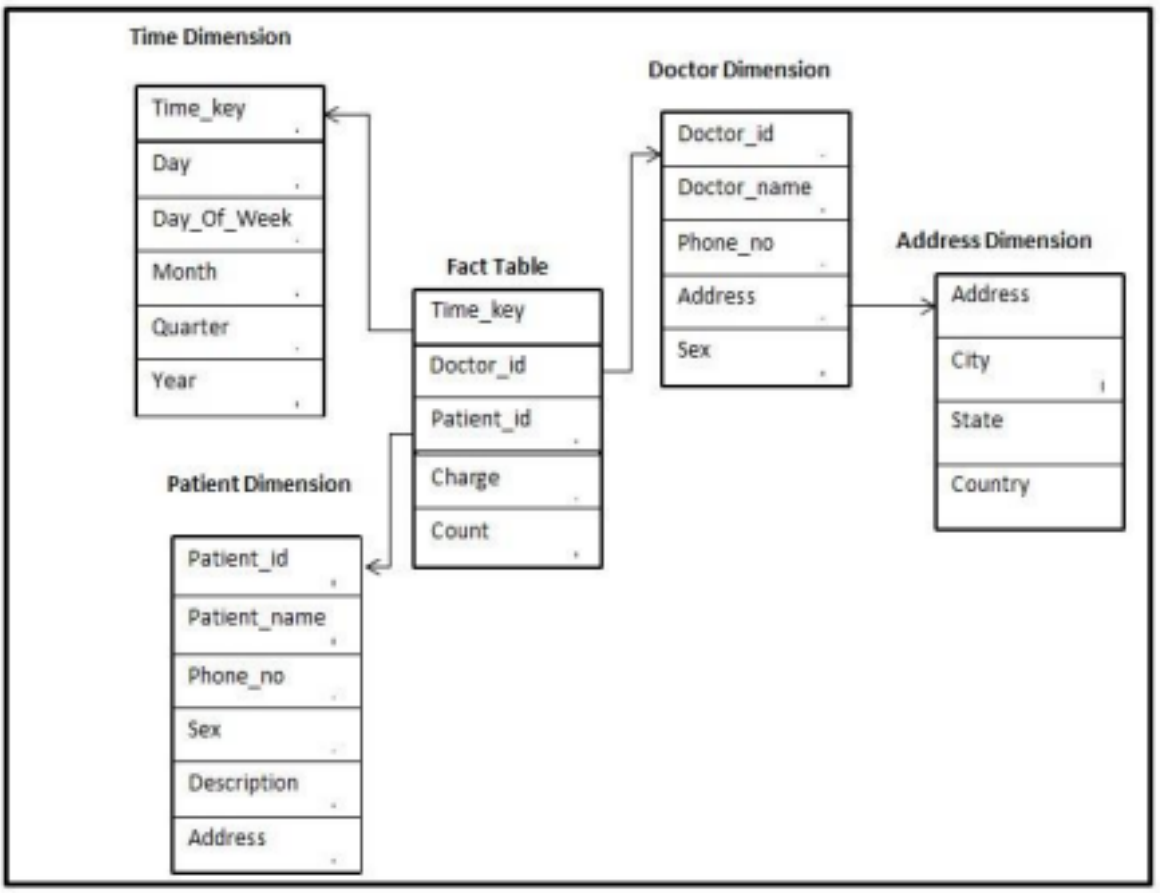
Download CSV

5 rows selected.

Star Schema:



Snowflake Schema:


Time Dimension
Time_key
Day
Day_Of_Week
Month
Quarter
Year
Doctor Dimension
Doctor_id
Doctor_name
Phone_no
Address
Sex
Address Dimension
Address
City
State
Country
Fact Table
Time_key
Doctor_id
Patient_id
Charge
Count
Patient Dimension
Patient_id
Patient_name
Phone_no
Sex
Description
Address