

CASE STUDY 1

The SQL commands to create tables and insert records to the operational database – Task 1

```
/* Task 1*/
DROP TABLE Clinic CASCADE CONSTRAINTS;
DROP TABLE Service CASCADE CONSTRAINTS;
DROP TABLE Doctor CASCADE CONSTRAINTS;
DROP TABLE Assignment CASCADE CONSTRAINTS;
DROP TABLE Patient CASCADE CONSTRAINTS;

/* Creating the tables */
CREATE TABLE Clinic
(hospital_id VARCHAR(5) NOT NULL,
hospital_name VARCHAR2(20) NOT NULL,
hospital_address VARCHAR2(50),
suburb VARCHAR2(50),
postcode NUMBER(4)
);

CREATE TABLE Service
(service_id VARCHAR(5) NOT NULL,
staff_id VARCHAR2(20) NOT NULL,
hospital_id VARCHAR(5) NOT NULL,
service_name VARCHAR2(50),
service_cost NUMBER(5)
);

CREATE TABLE Doctor
(staff_id VARCHAR2(20) NOT NULL,
staff_name VARCHAR(20),
staff_ph NUMBER(10)
);

CREATE TABLE Assignment
(assignment_id VARCHAR(5) NOT NULL,
patient_id VARCHAR(10) NOT NULL,
patient_service_start_date DATE NOT NULL,
patient_service_end_date DATE NOT NULL,
service_id VARCHAR(5) NOT NULL
);

CREATE TABLE Patient
(patient_id VARCHAR(10) NOT NULL,
```

```
patient_name VARCHAR2(20) NOT NULL,  
patient_age NUMERIC(3,1),  
patient_ph_no NUMBER(10),  
patient_address VARCHAR2(50),  
patient_nationality VARCHAR2(20),  
patient_emergency_contact NUMBER(10) NOT NULL  
);
```

```
/* Adding Primary Keys */
```

```
ALTER TABLE Clinic  
ADD CONSTRAINT PK_Clinic PRIMARY KEY (hospital_id);
```

```
ALTER TABLE Service  
ADD CONSTRAINT PK_Service PRIMARY KEY (service_id);
```

```
ALTER TABLE Assignment  
ADD CONSTRAINT PK_Assignment PRIMARY KEY  
(assignment_id,service_id,patient_id);
```

```
ALTER TABLE Doctor  
ADD CONSTRAINT PK_Doctor PRIMARY KEY (staff_id);
```

```
ALTER TABLE Patient  
ADD CONSTRAINT PK_Patient PRIMARY KEY (patient_id);
```

```
/* Adding Foreign Key constraints */
```

```
ALTER TABLE Service  
ADD FOREIGN KEY (hospital_id) REFERENCES Clinic(hospital_id);
```

```
ALTER TABLE Service  
ADD FOREIGN KEY (staff_id) REFERENCES Doctor(staff_id);
```

```
ALTER TABLE Assignment  
ADD FOREIGN KEY (service_id) REFERENCES Service(service_id);
```

```
ALTER TABLE Assignment  
ADD FOREIGN KEY (patient_id) REFERENCES Patient(patient_id);
```

```
/* Inserting values into Clinic*/
```

```
INSERT INTO Clinic  
VALUES ('H1','Clayton Clinic','21 Clayton Road','Clayton',3168);  
INSERT INTO Clinic  
VALUES ('H2','Caulfield Clinic','22 Caulfield Road','Caulfield',3162);  
INSERT INTO Clinic  
VALUES ('H3','Carnegie Clinic','23 Carnegie Road','Carnegie',3163);  
INSERT INTO Clinic
```

```
VALUES ('H4','Oakleigh Clinic','24 Oakleigh Road','Oakleigh',3166);
INSERT INTO Clinic
VALUES ('H5','Westall Clinic','25 Westall Road','Westall',3169);
```

/* Inserting values into Patient*/

```
INSERT INTO Patient
VALUES ('P0','Albus Severus Potter',0.5,420452371,'31 Clayton Road, Clayton
3168','Australian',420450001);
```

```
INSERT INTO Patient
VALUES ('P1','Harry Potter',11,420452371,'31 Clayton Road, Clayton
3168','Australian',420450001);
INSERT INTO Patient
VALUES ('P2','Hermione Granger',22,420452372,'32 Caulfield Road, Clayton
3162','British',420450002);
INSERT INTO Patient
VALUES ('P3','Ronald Weasley',33,420452373,'33 Carnegie Road, Carnegie
3163','American',420450003);
```

```
INSERT INTO Patient
VALUES ('P4','Severus Snape',14,420452374,'34 Oakleigh Road, Oakleigh
3168','Canadian',420450004);
```

```
INSERT INTO Patient
VALUES ('P5','Draco Malfoy',25,420452375,'35 Westall Road, Westall
3169','French',420450005);
```

```
INSERT INTO Patient
VALUES ('P6','Daniel Radcliffe',31,420452371,'36 Clayton Road, Clayton
3168','Australian',420450001);
```

```
INSERT INTO Patient
VALUES ('P7','Emma Watson',12,420452372,'37 Caulfield Road, Clayton
3162','British',420450002);
```

```
INSERT INTO Patient
VALUES ('P8','Rupert Grint',23,420452373,'38 Carnegie Road, Carnegie
3163','American',420450003);
```

```
INSERT INTO Patient
VALUES ('P9','Alan Rickman',34,420452374,'39 Oakleigh Road, Oakleigh
3168','Canadian',420450004);
```

```
INSERT INTO Patient
VALUES ('P10','Tom Felton',15,420452375,'40 Westall Road, Westall
3169','French',420450005);
```

```
INSERT INTO Patient
VALUES ('P11','Arthur Weasley',70,420452376,'41 Oakleigh Road, Oakleigh
3168','Canadian',420450006);
```

```
INSERT INTO Patient
VALUES ('P12','Molly Weasley',65,420452377,'42 Westall Road, Westall
3169','French',420450007);
```

```
/* Inserting values into Doctor*/
```

```
INSERT INTO Doctor VALUES ('D1','James Potter',420451000);
INSERT INTO Doctor VALUES ('D2','Lily Evans',420451002);
INSERT INTO Doctor VALUES ('D3','Serius Black',420451003);
INSERT INTO Doctor VALUES ('D4','Remus Lupin',420451004);
INSERT INTO Doctor VALUES ('D5','Albus Dumbledore',420451005);
INSERT INTO Doctor VALUES ('D6','Minerva McGonagall',420451006);
INSERT INTO Doctor VALUES ('D7','Godric Gryffindor',420451007);
INSERT INTO Doctor VALUES ('D8','Salazar Slytherin',420451008);
INSERT INTO Doctor VALUES ('D9','Helga Hufflepuff',420451009);
INSERT INTO Doctor VALUES ('D10','Rowena Ravenclaw',420451010);
```

```
/* Inserting values into Service*/
```

```
INSERT INTO Service VALUES ('S1','D1','H1','General Medicine',10);
INSERT INTO Service VALUES ('S2','D2','H1','Mental Health',20);
INSERT INTO Service VALUES ('S3','D3','H2','Skin Diseases',30);
INSERT INTO Service VALUES ('S4','D4','H3','Paediatric Health',40);
INSERT INTO Service VALUES ('S5','D5','H3','Sexual Health',50);
INSERT INTO Service VALUES ('S6','D6','H3','Paediatric Health',40);
INSERT INTO Service VALUES ('S7','D7','H4','General Medicine',10);
INSERT INTO Service VALUES ('S8','D8','H4','Mental Health',20);
INSERT INTO Service VALUES ('S9','D9','H5','Skin Diseases',30);
INSERT INTO Service VALUES ('S10','D10','H5','Paediatric Health',40);
```

```
/* Inserting values into Assignment */
```

```
INSERT INTO Assignment
VALUES
('A0','P0',TO_DATE('01/04/2020','DD/MM/YYYY'),TO_DATE('01/05/2020','DD/MM/YYYY'),'S1');
```

```
INSERT INTO Assignment
VALUES
('A1','P1',TO_DATE('01/01/2020','DD/MM/YYYY'),TO_DATE('01/02/2020','DD/MM/YYYY'),'S1');
```

```
INSERT INTO Assignment
VALUES
('A2','P2',TO_DATE('02/01/2020','DD/MM/YYYY'),TO_DATE('02/02/2020','DD/MM/YYYY'),'S2');
```

```
INSERT INTO Assignment
VALUES
('A3','P3',TO_DATE('03/02/2020','DD/MM/YYYY'),TO_DATE('03/03/2020','DD/MM/YYYY'),'S3');
```

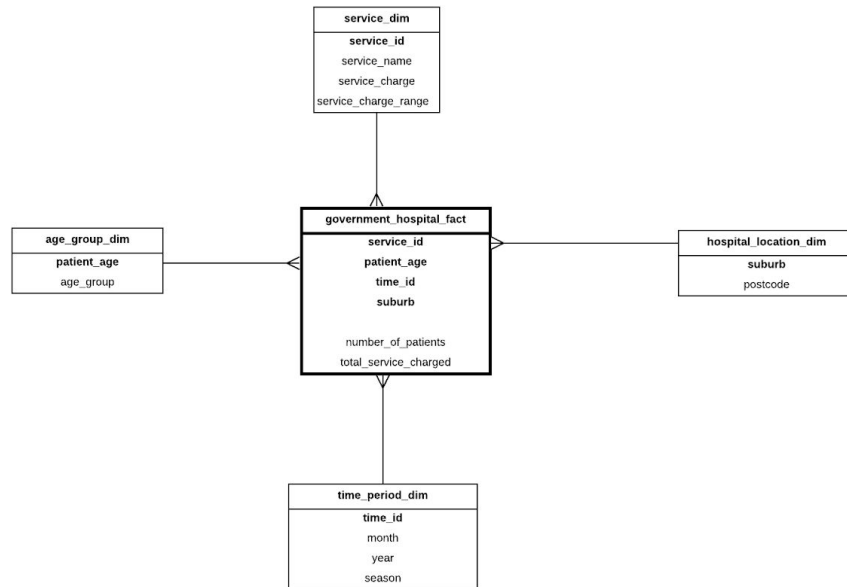
```
INSERT INTO Assignment
```

```
VALUES
('A4','P4',TO_DATE('04/02/2020','DD/MM/YYYY'),TO_DATE('04/03/2020','DD/MM/Y
YYY'),'S4');
INSERT INTO Assignment
VALUES
('A5','P5',TO_DATE('05/03/2020','DD/MM/YYYY'),TO_DATE('05/04/2020','DD/MM/Y
YYY'),'S5');
INSERT INTO Assignment
VALUES
('A6','P6',TO_DATE('06/03/2020','DD/MM/YYYY'),TO_DATE('06/04/2020','DD/MM/Y
YYY'),'S6');
INSERT INTO Assignment
VALUES
('A7','P7',TO_DATE('07/04/2020','DD/MM/YYYY'),TO_DATE('01/05/2020','DD/MM/Y
YYY'),'S7');
INSERT INTO Assignment
VALUES
('A8','P8',TO_DATE('08/04/2020','DD/MM/YYYY'),TO_DATE('01/05/2020','DD/MM/Y
YYY'),'S8');
INSERT INTO Assignment
VALUES
('A9','P9',TO_DATE('09/04/2020','DD/MM/YYYY'),TO_DATE('01/05/2020','DD/MM/Y
YYY'),'S9');
INSERT INTO Assignment
VALUES
('A10','P10',TO_DATE('10/04/2020','DD/MM/YYYY'),TO_DATE('01/05/2020','DD/MM/
YYYY'),'S10');

INSERT INTO Assignment
VALUES
('A11','P11',TO_DATE('11/04/2020','DD/MM/YYYY'),TO_DATE('02/05/2020','DD/MM/
YYYY'),'S9');
INSERT INTO Assignment
VALUES
('A12','P12',TO_DATE('12/04/2020','DD/MM/YYYY'),TO_DATE('02/05/2020','DD/MM/
YYYY'),'S10');

commit;
```

The star schema diagram – Task 2



Case Study 1 Star Schema
Gayatri Aniruddha - 30945305

The Two-Column Table Methodology illustration – Task 3

Two Column Table Methodology Validation:

In my star schema, my facts and dimensions are given below:

Facts:

- Number of Patients
- Total Service Charge

Dimensions:

- Age Group
- Time Period
- Service
- Hospital Locations

Here, the first column would be dimensions and the second column would be the facts.

Following are my two-column tables.

- **Patient age point of view**

patient_age	Number of Patients	Total Service Charge
<1	0	0
<18	5	1900
18+	4	1299
65+	1	400

- **Time point of view**

time_id	Number of Patients	Total Service Charge
01/2020	4	1000
02/2020	3	900
03/2020	3	1200

- **Service point of view**

service_id	Number of Patients	Total Service Charge
S1	2	20
S2	1	20
S3	1	30
S4	1	40
S5	1	50
S6	1	40
S7	1	10

S8	1	20
S9	1	30
S10	2	80

- Suburb point of view

suburb	Number of Patients	Total Service Charge
Clayton	2	300
Caulfield	2	700
Oakleigh	2	1100
Carnegie	2	300
Westall	2	700

The SQL commands to create the dimension and fact tables, as well as the contents of these tables – Task 4

```
/* Task 4 */
```

```
DROP TABLE age_group_dim CASCADE CONSTRAINTS PURGE;  
DROP TABLE service_dim CASCADE CONSTRAINTS PURGE;  
DROP TABLE time_period_dim CASCADE CONSTRAINTS PURGE;  
DROP TABLE hospital_location_dim CASCADE CONSTRAINTS PURGE;  
DROP TABLE temporary_fact CASCADE CONSTRAINTS PURGE;  
DROP TABLE government_hospital_fact CASCADE CONSTRAINTS PURGE;
```

```
/* Creating dimension tables */
```

```
/* 1) age group dimension */
```

```
CREATE TABLE age_group_dim as  
SELECT patient_age  
FROM patient;
```

```
/* Adding an age group to age_group_dim */
```

```
ALTER TABLE age_group_dim  
ADD age_group VARCHAR2(50);
```

```
UPDATE age_group_dim  
SET age_group = 'Infant'  
WHERE patient_age < 1;
```

```
UPDATE age_group_dim  
SET age_group = 'Children'  
WHERE patient_age >= 1 AND patient_age < 18;
```

```
UPDATE age_group_dim  
SET age_group = 'Adult'  
WHERE patient_age >= 18 AND patient_age < 65;
```

```
UPDATE age_group_dim  
SET age_group = 'Senior'  
WHERE patient_age >= 65;
```

	PATIENT_AGE	AGE_GROUP
1	0.5	Infant
2	11	Children
3	22	Adult
4	33	Adult
5	14	Children
6	25	Adult
7	31	Adult
8	12	Children
9	23	Adult
10	34	Adult
11	15	Children
12	70	Senior
13	65	Senior

/* 2) service dimension */

```
CREATE TABLE service_dim as
SELECT DISTINCT service_id, service_name, service_cost
FROM Service;
```

/* Adding a service range */

```
ALTER TABLE service_dim
ADD service_charge_range VARCHAR2(50);
```

```
UPDATE service_dim
SET service_charge_range = 'Low Price'
WHERE service_cost < 20;
```

```
UPDATE service_dim
SET service_charge_range = 'Medium Price'
WHERE service_cost <=50 AND service_cost >= 20;
```

```
UPDATE service_dim
SET service_charge_range = 'High Price'
WHERE service_cost > 50;
```

	SERVICE_ID	SERVICE_NAME	SERVICE_COST	SERVICE_CHARGE_RANGE
1	S1	General Medicine	10	Low Price
2	S2	Mental Health	20	Medium Price
3	S3	Skin Diseases	30	Medium Price
4	S4	Paediatric Health	40	Medium Price
5	S5	Sexual Health	50	Medium Price
6	S6	Paediatric Health	40	Medium Price
7	S7	General Medicine	10	Low Price
8	S8	Mental Health	20	Medium Price
9	S9	Skin Diseases	30	Medium Price
10	S10	Paediatric Health	40	Medium Price

```
/* 3) time period dim */
```

```
CREATE TABLE time_period_dim as
SELECT DISTINCT
    TO_CHAR(patient_service_start_date,'MM/YYYY') as time_id,
    TO_CHAR(patient_service_start_date,'MM') as month,
    TO_CHAR(patient_service_start_date,'YYYY') as year
FROM Assignment;
```

```
/* Adding month names */
```

```
ALTER TABLE time_period_dim
ADD month_name VARCHAR2(50);
```

```
UPDATE time_period_dim
SET month_name = 'January'
WHERE to_char(month) = '01';
```

```
UPDATE time_period_dim
SET month_name = 'February'
WHERE to_char(month) = '02';
```

```
UPDATE time_period_dim
SET month_name = 'March'
WHERE to_char(month) = '03';
```

```
UPDATE time_period_dim
SET month_name = 'April'
WHERE to_char(month) = '04';
```

```
/* Adding seasons */
```

```
ALTER TABLE time_period_dim
ADD season VARCHAR2(50);
```

```
UPDATE time_period_dim
SET season = 'Summer'
WHERE time_id LIKE '01%';
```

```
UPDATE time_period_dim
SET season = 'Autumn'
WHERE time_id LIKE '02%' OR time_id LIKE '03%';
```

```
UPDATE time_period_dim
SET season = 'Winter'
WHERE time_id LIKE '04%' OR time_id = '05%' OR time_id LIKE '06%' OR time_id = '07%';
```

```
UPDATE time_period_dim
```

```
SET season = 'Spring'
```

```
WHERE time_id LIKE '08%' OR time_id = '09%' OR time_id LIKE '10%';
```

	TIME_ID	MONTH	YEAR	MONTH_NAME	SEASON
1	01/2020	01	2020	January	Summer
2	04/2020	04	2020	April	Winter
3	03/2020	03	2020	March	Autumn
4	02/2020	02	2020	February	Autumn

```
/* 4) hospital_location dimension */
```

```
CREATE TABLE hospital_location_dim as
```

```
SELECT suburb, postcode
```

```
FROM Clinic;
```

	SUBURB	POSTCODE
1	Clayton	3168
2	Caulfield	3162
3	Carnegie	3163
4	Oakleigh	3166
5	Westall	3169

```
/* Creating a temporary fact tables */
```

```
CREATE TABLE temporary_fact as
```

```
SELECT DISTINCT
```

```
    TO_CHAR(A.patient_service_start_date,'MM/YYYY') as time_id,
```

```
    S.service_id,
```

```
    S.service_cost,
```

```
    H.hospital_id,
```

```
    H.suburb,
```

```
    P.patient_id,
```

```
    P.patient_age
```

```
FROM Service S, Clinic H, Patient P, Assignment A
```

```
WHERE
```

```
    H.hospital_id = S.hospital_id AND
```

```
    S.service_id = A.service_id AND
```

```
    A.patient_id = P.patient_id
```

```
GROUP BY
```

```
    TO_CHAR(A.patient_service_start_date,'MM/YYYY'),
```

```
    S.service_id,
```

```
    s.service_cost,
```

```
    H.hospital_id,
```

```
    H.suburb,
```

```
    P.patient_id,
```

```
    P.patient_age;
```







	TIME_ID	SERVIC...	SERVICE_COST	HOSPITAL_ID	SUBURB	PATIENT_ID	PATIENT_AGE
1	01/2020	S1	10	H1	Clayton	P1	11
2	04/2020	S1	10	H1	Clayton	P0	0.5
3	04/2020	S10	40	H5	Westall	P12	65
4	04/2020	S10	40	H5	Westall	P10	15
5	01/2020	S2	20	H1	Clayton	P2	22
6	02/2020	S3	30	H2	Caulfield	P3	33
7	02/2020	S4	40	H3	Carnegie	P4	14
8	03/2020	S5	50	H3	Carnegie	P5	25
9	03/2020	S6	40	H3	Carnegie	P6	31
10	04/2020	S7	10	H4	Oakleigh	P7	12
11	04/2020	S8	20	H4	Oakleigh	P8	23
12	04/2020	S9	30	H5	Westall	P11	70
13	04/2020	S9	30	H5	Westall	P9	34

/* Creating the actual fact table */

```

CREATE TABLE government_hospital_fact AS
SELECT
  t.time_id,
  t.patient_age,
  t.service_id,
  t.suburb,
  COUNT(t.patient_id) AS "NUMBER_OF_PATIENTS",
  SUM(t.service_cost) AS "TOTAL_SERVICE_CHARGED"
FROM temporary_fact t
GROUP BY
  t.time_id,
  t.patient_age,
  t.service_id,
  t.suburb;

Commit;
```

	 TIME_ID	 PATIENT_AGE	 SERVICE_ID	 SUBURB	 NUMBER_OF_PATIENTS	 TOTAL_SERVICE_CHARGED
1	01/2020	11	S1	Clayton	1	10
2	01/2020	22	S2	Clayton	1	20
3	02/2020	14	S4	Carnegie	1	40
4	02/2020	33	S3	Caulfield	1	30
5	03/2020	31	S6	Carnegie	1	40
6	03/2020	25	S5	Carnegie	1	50
7	04/2020	70	S9	Westall	1	30
8	04/2020	15	S10	Westall	1	40
9	04/2020	65	S10	Westall	1	40
10	04/2020	23	S8	Oakleigh	1	20
11	04/2020	12	S7	Oakleigh	1	10
12	04/2020	0.5	S1	Clayton	1	10
13	04/2020	34	S9	Westall	1	30

The SQL commands to answer the queries in Task 5 and the query results

```
/* task 5 */
```

```
/* 5a */
```

```
/* Total number of patients in Winter */
```

```
/* Clarification - Here, while calculating the total number of patients who have made appointments - I am considering only the patient_service_start_date. */
```

```
/* This is because a patient may make an appointment in Autumn and can still be there in the hospital during Winter. */
```

```
/* Thus, patients making appointments during Autumn i.e patients with patient_service_start_date in Autumn will be considered as Autumn patients and not Winter patients */
```

```
SELECT
    t.season,
    COUNT(NUMBER_OF_PATIENTS)
FROM
    government_hospital_fact g JOIN time_period_dim t
    ON
    g.time_id = t.time_id
WHERE
    t.season = 'Winter'
GROUP BY
    t.season;
```

SEASON	COUNT(NUMBER_OF_PATIENTS)
1 Winter	7

```
/* 5b */
```

```
/* Total service charged for each service cost type. */
```

```
SELECT
    s.service_charge_range AS "SERVICE COST TYPE",
    SUM(total_service_charged) AS "TOTAL SERVICE CHARGED"
FROM
    government_hospital_fact g JOIN service_dim s
    ON
    g.service_id = s.service_id
GROUP BY
    s.service_charge_range;
```

SERVICE COST TYPE	TOTAL SERVICE CHARGED
1 Medium Price	340
2 Low Price	30

```
/* 5c */
```

```
/* total number of patients by each age group in April 2020. */
```

```
SELECT
    COUNT(number_of_patients) AS "TOTAL NUMBER OF PATIENTS",
    a.age_group AS "AGE GROUP"
FROM
    government_hospital_fact g JOIN age_group_dim a
    ON
    g.patient_age = a.patient_age
    JOIN
    time_period_dim t
    ON
    g.time_id = t.time_id
WHERE
    t.month_name = 'April'
    AND
    TO_CHAR(t.year) = '2020'
GROUP BY
    a.age_group;
```

	⚡ TOTAL NUMBER OF PATIENTS	⚡ AGE GROUP
1	2	Children
2	2	Adult
3	1	Infant
4	2	Senior

```
/* 5d */
```

```
/* e total service charged for general medical consultations in each suburb. */
```

```
SELECT
    SUM(total_service_charged) AS "Total Service Charged",
    l.suburb,
    l.postcode
FROM
    government_hospital_fact g JOIN
    service_dim s ON
    g.service_id = s.service_id
    JOIN
    hospital_location_dim l
    ON
    g.suburb = l.suburb
WHERE
    s.service_name = 'General Medicine'
GROUP BY
    l.suburb,
    l.postcode;
```


	Total Service Charged	SUBURB	POSTCODE
1		20 Clayton	3168
2		10 Oakleigh	3166