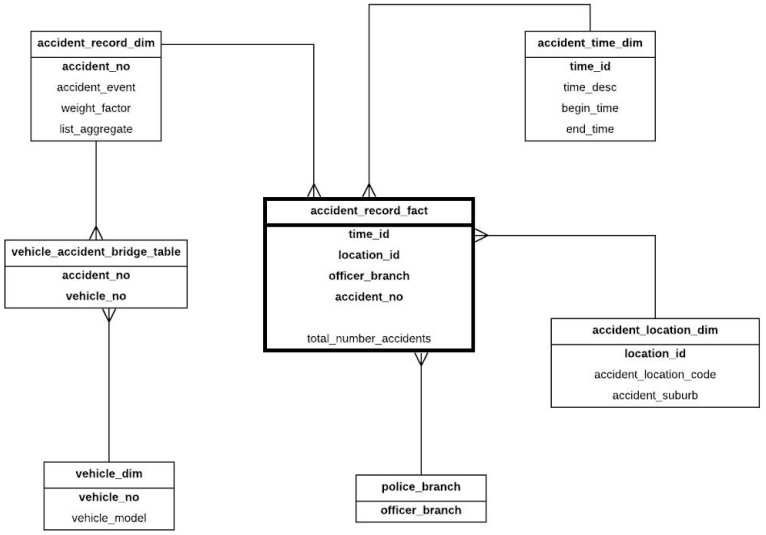
CASE STUDY 2

TASK 1:

Develop an ACCIDENT RECORDS star schema. Identify the fact table, dimensions and attributes required to support the schema. If the star schema consists of a Bridge Table, you have to also include the Weight Factor and List Aggregate. The result of this task is a star schema diagram. You can use any drawing tool, such as Lucidchart, to draw the star schema.



Case Study 2 Star Schema

Gayatri Aniruddha - 30945305

TASK 2:

Validate your star schema using the Two-Column Table Methodology. You are required to illustrate some two-column tables for this task based on your star schema design.

Two Column Table Methodology Validation:

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

- Total number of accidents

Dimensions:

- Accident time period
- Accident location
- Police branch
- Accident record - Further, this dimension is then connected to a vehicle dimension via a bridge table.

Here, the first column would be dimensions and the second column would be the facts.
Following are my two-column tables.

- **Accident period point of view**

accident_time_id	Total Number of accidents
1	14
2	11

- **Accident location point of view**

location_id	Total Number of accidents
CC1CLarinda	2
CC2Clayton	1
CD1Carnegie	2
CD2Clayton	1

The same trend can be observed for other location_ids as well.

- **Accident number point of view**

accident_no	Total Number of accidents
A001	1
A002	1
A003	1
A004	1

The same trend can be observed for other accident_location_codes as well.

- **Officer Branch point of view**

officer_branch	Total Number of accidents
Donvale	8
Ringwood	7
Blackburn	4
Boxhill	4
Mitcham	2

TASK 3:

Write the SQL commands to create the fact and dimension tables. You need to create a script file containing the appropriate SQL commands to create the fact and dimension tables. The operational tables are accessible from the ACCIDENT account. The result of this task is the SQL commands. You will also need to show the contents of the tables that you have created.

```
/* task 3 */

DROP TABLE accident_time_dim CASCADE CONSTRAINTS PURGE;
DROP TABLE accident_location_dim CASCADE CONSTRAINTS PURGE;
DROP TABLE police_branch_dim CASCADE CONSTRAINTS PURGE;
DROP TABLE vehicle_dim CASCADE CONSTRAINTS PURGE;
DROP TABLE accident_record_dim CASCADE CONSTRAINTS PURGE;
DROP TABLE vehicle_accident_bridge_table CASCADE CONSTRAINTS PURGE;
DROP TABLE temp_fact CASCADE CONSTRAINTS PURGE;
DROP TABLE accident_record_fact CASCADE CONSTRAINTS PURGE;

/* Creating dimension tables */

/* 1) accident time dim */
CREATE TABLE accident_time_dim
(
   time_id NUMBER(7),
   time_desc VARCHAR2(50),
   begin_time DATE,
   end_time DATE
);
/* Inserting values for Day time and Night Time */
INSERT INTO accident_time_dim
VALUES ('1','Day
time',TO_DATE('06:00:00','HH24:MI:SS'),TO_DATE('17:59:00','HH24:MI:SS'));
INSERT INTO accident_time_dim
VALUES ('2','Night
time',TO_DATE('18:00:00','HH24:MI:SS'),TO_DATE('05:59:00','HH24:MI:SS'));
```

	TIME_ID	TIME_DESC	TO_CHAR(BEGIN_TIME,'HH24:MI:SS')	TO_CHAR(END_TIME,'HH24:MI:SS')
1	1	Day time	06:00:00	17:59:00
2	2	Night time	18:00:00	05:59:00

```
/* 2) accident location dim */
CREATE TABLE accident_location_dim AS
SELECT DISTINCT
   accident_location_code || accident_suburb AS location_id,
   accident_street,
   accident_suburb
FROM
   ACCIDENT.accident;
```

	LOCATION_ID	ACCIDENT_STREET	ACCIDENT_SUBURB
1	CD1Carnegie	Dandenong Road	Carnegie
2	CW1Chadstone	Warrigal Road	Chadstone
3	CD3Chadstone	Dandenong Road	Chadstone
4	CH1Chadstone	Huntingdale Road	Chadstone
5	CC1Clarinda	Centre Road	Clarinda
6	CD2Clayton	Dandenong Road	Clayton
7	CC2Clayton	Clayton Road	Clayton
8	CW1Clayton	Wellington Road	Clayton
9	GB1Glen Waverley	Blackburn Road	Glen Waverley
10	GH1Glen Waverley	High Street Road	Glen Waverley
11	GW1Glen Waverley	Waverley Road	Glen Waverley
12	HD1Huntingdale	Dandenong Road	Huntingdale
13	MS1Morabbin	South Road	Morabbin
14	MC1Mount Waverley	Clayton Road	Mount Waverley
15	MH1Mount Waverley	Highbury Road	Mount Waverley
16	MS1Mulgrave	Springvale Road	Mulgrave
17	MB1Mulgrave	Blackburn Road	Mulgrave
18	RW1Rowville	Wellington Road	Rowville
19	RN1Rowville	North Road	Rowville
20	SD1Sandown	Dandenong Road	Sandown
21	SS1Springvale	Springvale Road	Springvale
22	SD1Syndal	Blackburn Road	Syndal
23	WC1Westall	Centre Road	Westall

```
/* 3) police branch dimension */
CREATE TABLE police_branch_dim AS
SELECT
   officer_branch
FROM
   ACCIDENT.police_officer;
```

	OFFICER_BRANCH
1	Donvale
2	Donvale
3	Ringwood
4	Ringwood
5	Blackburn
6	Blackburn
7	Box Hill
8	Box Hill
9	Mitcham
10	Mitcham

```
/* 4b) vehicle dimension */
CREATE TABLE vehicle_dim AS
SELECT DISTINCT
   vehicle_no,
   vehicle_model
FROM
   ACCIDENT.vehicle;
```

	VEHICLE_NO	VEHICLE_MODEL
1	VC001	Yaris
2	VT002	Titan
3	VM003	Ninja 1000
4	VC004	Jazz
5	VT005	Dutro
6	VC006	Corolla
7	VC007	X5
8	VC008	Carnival
9	VM009	Volusia
10	VC010	Odyssey
11	VC011	Pulsar
12	VC012	CX5
13	VT013	Gigamax
14	VC014	Golf
15	VC015	Tarago
16	VC016	Estima
17	VC017	Alphard
18	VC018	Serena
19	VC019	Rondo
20	VC020	Camry
21	VC021	Magna
22	VC022	Grandis
23	VC023	Wish
24	VC024	Commodore
25	VC025	Astra

```
/* vehicle accident bridge table*/
CREATE TABLE vehicle_accident_bridge_table AS
SELECT DISTINCT
   vehicle_no,
   accident_no
FROM
   ACCIDENT.accident_record A;
```

	VEHICLE_NO	ACCIDENT_NO
1	VC010	A001
2	VC011	A002
3	VC004	A003
4	VT002	A003
5	VC001	A004
6	VC006	A004
7	VM003	A005
8	VC006	A006
9	VC007	A006
10	VC008	A007
11	VM009	A007
12	VC012	A008
13	VT013	A008
14	VC014	A009
15	VC015	A009
16	VC010	A010
17	VC016	A011
18	VC017	A011
19	VC018	A012
20	VC019	A012
21	VC020	A013
22	VC021	A013
23	VM009	A014
24	VC022	A015
25	VC023	A015
26	VC024	A016
27	VC025	A016
28	VC011	A021

```
/* 4a) accident record dimension */
/* Adding the weight factor and list of attributes as well */
CREATE TABLE accident_record_dim AS
SELECT
   a.accident_no,
   a.accident_event,
   1.0/(COUNT(v.vehicle_no)) AS "weight_factor",
   LISTAGG(v.vehicle_no,'_') WITHIN GROUP( ORDER BY v.vehicle_no) AS
"list_aggregate"
FROM
   ACCIDENT.accident a,
   ACCIDENT.accident_record v
WHERE
   a.accident_no = v.accident_no
GROUP BY
   a.accident_no,
   a.accident_event;
```

	ACCIDENT_NO	ACCIDENT_EVENT	weight_factor	list_aggregate
1	A001	Rollover on/off carriageway	1	VC010
2	A002	Ran off carriageway	1	VC011
3	A003	Collision	0.5	VC004_VT002
4	A004	Collision	0.5	VC001_VC006
5	A005	Fell from vehicle	1	VM003
6	A006	Collision	0.5	VC006_VC007
7	A007	Collision	0.5	VC008_VM009
8	A008	Collision	0.5	VC012_VT013
9	A009	Collision	0.5	VC014_VC015
10	A010	Ran off carriageway	1	VC010
11	A011	Collision	0.5	VC016_VC017
12	A012	Collision	0.5	VC018_VC019
13	A013	Collision	0.5	VC020_VC021
14	A014	Fell from vehicle	1	VM009
15	A015	Collision	0.5	VC022_VC023
16	A016	Collision	0.5	VC024_VC025
17	A021	Ran off carriageway	1	VC011

```
/* Creating a temporary fact table */

CREATE TABLE temp_fact AS
   SELECT
      a.accident_location_code || a.accident_suburb AS location_id,
      a.accident_no,
      a.accident_date_time,
      p.officer_branch
   FROM
      ACCIDENT.accident a,
      ACCIDENT.police_officer p
   WHERE
      a.officer_id = p.officer_id;

ALTER TABLE temp_fact
ADD time_id NUMBER(7);

UPDATE temp_fact
SET time_id = 1
WHERE to_char(accident_date_time,'HH24:MI') >= '06:00'
AND to_char(accident_date_time,'HH24:MI') <= '17:59';

UPDATE temp_fact
SET time_id = 2
WHERE to_char(accident_date_time,'HH24:MI') >= '18:00'
AND to_char(accident_date_time,'HH24:MI') <= '23:59';

UPDATE temp_fact
SET time_id = 2
WHERE to_char(accident_date_time,'HH24:MI') >= '00:00'
AND to_char(accident_date_time,'HH24:MI') <= '05:59';
```

	LOCATION_ID	ACCIDENT_NO	ACCIDENT_DATE_TIME	OFFICER_BRANCH	TIME_ID
1	GH1Glen Waverley	A011	21/APR/20	Donvale	1
2	WC1Westall	A025	01/APR/20	Donvale	2
3	CD2Clayton	A003	24/APR/20	Donvale	1
4	CD1Carnegie	A001	10/APR/20	Donvale	1
5	SS1Springvale	A002	11/APR/20	Donvale	1
6	GB1Glen Waverley	A005	15/APR/20	Donvale	1
7	SD1Syndal	A004	22/APR/20	Donvale	1
8	CC1Clarinda	A013	12/APR/20	Donvale	2
9	CD1Carnegie	A006	21/APR/20	Ringwood	1
10	MS1Mulgrave	A016	05/APR/20	Ringwood	2
11	GW1Glen Waverley	A007	14/APR/20	Ringwood	2
12	CH1Chadstone	A008	14/APR/20	Ringwood	1
13	CC2Clayton	A019	07/APR/20	Ringwood	1
14	RW1Rowville	A020	09/APR/20	Ringwood	2
15	MC1Mount Waverley	A009	12/APR/20	Ringwood	2
16	MB1Mulgrave	A012	15/APR/20	Blackburn	1
17	CD3Chadstone	A010	23/APR/20	Blackburn	2
18	CW1Chadstone	A015	11/APR/20	Blackburn	1
19	HD1Huntingdale	A014	22/APR/20	Blackburn	1
20	MH1Mount Waverley	A018	09/APR/20	Box Hill	1
21	SD1Sandown	A017	11/APR/20	Box Hill	2
22	RN1Rowville	A022	03/APR/20	Box Hill	2
23	CW1Clayton	A021	10/APR/20	Box Hill	2
24	CC1Clarinda	A023	05/APR/20	Mitcham	2
25	MS1Morabbin	A024	04/APR/20	Mitcham	1

/* Creating the actual fact table */

```
CREATE TABLE accident_record_fact AS
SELECT
   t.time_id,
   t.accident_no,
   t.location_id,
   t.officer_branch,
   COUNT(t.accident_no) AS "total_number_of_accidents"
FROM
   temp_fact t
GROUP BY
   t.time_id,
   t.accident_no,
   t.location_id,
   t.officer_branch;
```

	TIME_ID	ACCIDENT_NO	LOCATION_ID	OFFICER_BRANCH	total_number_of_accidents
1	2	A013	CC1Clarinda	Donvale	1
2	2	A023	CC1Clarinda	Mitcham	1
3	1	A019	CC2Clayton	Ringwood	1
4	1	A001	CD1Carnegie	Donvale	1
5	1	A006	CD1Carnegie	Ringwood	1
6	1	A003	CD2Clayton	Donvale	1
7	2	A010	CD3Chadstone	Blackburn	1
8	1	A008	CH1Chadstone	Ringwood	1
9	1	A015	CW1Chadstone	Blackburn	1
10	2	A021	CW1Clayton	Box Hill	1
11	1	A005	GB1Glen Waverley	Donvale	1
12	1	A011	GH1Glen Waverley	Donvale	1
13	2	A007	GW1Glen Waverley	Ringwood	1
14	1	A014	HD1Huntingdale	Blackburn	1
15	1	A012	MB1Mulgrave	Blackburn	1
16	2	A009	MC1Mount Waverley	Ringwood	1
17	1	A018	MH1Mount Waverley	Box Hill	1
18	1	A024	MS1Morabbin	Mitcham	1
19	2	A016	MS1Mulgrave	Ringwood	1
20	2	A022	RN1Rowville	Box Hill	1
21	2	A020	RW1Rowville	Ringwood	1
22	2	A017	SD1Sandown	Box Hill	1
23	1	A004	SD1Syndal	Donvale	1
24	1	A002	SS1Springvale	Donvale	1
25	2	A025	WC1Westall	Donvale	1

TASK 4:

Write the SQL commands to answer the following queries:

a) Show the total number of accidents happening by different locations and by different lighting periods (daytime: 6AM - 5:59PM and nighttime 6PM - 5:59AM).

b) Show the total number of accidents by each vehicle model.

c) Show the number of vehicles involved in every accident event in different locations.

d) Show the number of accidents taken care of by different police officer branches.

```
/* task 4 */

/* 4a */
/*  total number of accidents happening by different locations and by different
lighting periods 8 */

SELECT
   COUNT(DISTINCT f.accident_no) AS "TOTAL_ACCIDENTS",
   l.accident_suburb,
   t.time_desc
FROM
   accident_record_fact f,
   accident_time_dim t,
   accident_location_dim l
WHERE
   f.time_id = t.time_id
   AND
   f.location_id = l.location_id
GROUP BY
   l.accident_suburb,
   t.time_desc
ORDER BY
   l.accident_suburb;
```

	TOTAL_ACCIDENTS	ACCIDENT_SUBURB	TIME_DESC
1	2	Carnegie	Day time
2	2	Chadstone	Day time
3	1	Chadstone	Night time
4	2	Clarinda	Night time
5	2	Clayton	Day time
6	1	Clayton	Night time
7	2	Glen Waverley	Day time
8	1	Glen Waverley	Night time
9	1	Huntingdale	Day time
10	1	Morabbin	Day time
11	1	Mount Waverley	Day time
12	1	Mount Waverley	Night time
13	1	Mulgrave	Day time
14	1	Mulgrave	Night time
15	2	Rowville	Night time
16	1	Sandown	Night time
17	1	Springvale	Day time
18	1	Syndal	Day time
19	1	Westall	Night time

```
/* 4b */
/* e total number of accidents by each vehicle model */
SELECT
   COUNT(DISTINCT a.accident_no) AS "ACCIDENT_COUNT",
   v.vehicle_model
FROM
   accident_record_fact f,
   accident_record_dim a,
   vehicle_accident_bridge_table b,
   vehicle_dim v
WHERE
   f.accident_no = a.accident_no
   AND
   a.accident_no = b.accident_no
   AND
   v.vehicle_no = b.vehicle_no
GROUP BY
   V.vehicle_model;
```

	ACCIDENT_COUNT	VEHICLE_MODEL
1	2	Odyssey
2	1	Alphard
3	1	Estima
4	1	Astra
5	1	Carnival
6	1	Tarago
7	1	Gigamax
8	1	Serena
9	1	Wish
10	1	Golf
11	1	Magna
12	2	Volusia
13	1	Rondo
14	2	Pulsar
15	1	Yaris
16	1	Titan
17	1	Jazz
18	1	Camry
19	1	CX5
20	2	Corolla
21	1	Grandis
22	1	Commodore
23	1	Ninja 1000
24	1	X5

```
/* 4c */
/*  number of vehicles involved in every accident event on different locations */
SELECT
   COUNT(b.vehicle_no) AS "VEHICLE_COUNT",
   a.accident_event
FROM
   accident_record_fact f JOIN accident_record_dim a ON
   f.accident_no = a.accident_no
   JOIN
   vehicle_accident_bridge_table b
   ON
   b.accident_no = a.accident_no
GROUP BY
   A.accident_event;
```

	VEHICLE_COUNT	ACCIDENT_EVENT
1	3	Ran off carriageway
2	22	Collision
3	2	Fell from vehicle
4	1	Rollover on/off carriageway

```
/* 4d */
/*  number of accidents taken care of by different police officer branches */
SELECT
   COUNT(DISTINCT f.accident_no) AS "ACCIDENT_COUNT",
   p.officer_branch
FROM
   accident_record_fact f,
   police_branch_dim p
WHERE
   f.officer_branch = p.officer_branch
GROUP BY
   p.officer_branch
ORDER BY
   p.officer_branch;
```

	ACCIDENT_COUNT	OFFICER_BRANCH
1	4	Blackburn
2	4	Box Hill
3	8	Donvale
4	2	Mitcham
5	7	Ringwood

TASK 5:

You need to come up with additional two more questions and answer these questions using the SQL commands. Also explain the reason for why the management would like to have such information.

Question 1

Names and service time of the police officers who are in charge of the various accidents, grouped by events and officer branches.

Explanation:

Here, we need this data so that management is aware of the police officers who are handling the different cases.

- This is done so that the hard work, consistency, perseverance of the officers is recognised.
- The management can reward the deserving officers who handle many cases effectively.
- Finally, the management can even ensure that the officers are not over-burdened and are not handling multiple cases.
- Thus, this data will help in the smooth functioning of the police department and improve the efficiency.

For this, we will have to add some additional attributes to our star schema as follows.

```
/* police branch dimension */
DROP TABLE police_branch_dim CASCADE CONSTRAINTS PURGE;

CREATE TABLE police_branch_dim AS
SELECT
   officer_id,
   officer_fname,
   officer_lname,
   officer_startdate,
   officer_branch
FROM
   ACCIDENT.police_officer;

/* Query Solution */
SELECT
   COUNT(f."total_number_of_accidents") AS "ACCIDENT COUNT",
   p.officer_id,
   p.officer_fname,
   p.officer_lname,
   p.officer_startdate,
   p.officer_branch
FROM
   accident_record_fact f JOIN police_branch_dim p
   ON
```

```
    p.officer_branch = f.officer_branch
GROUP BY
    p.officer_id,
    p.officer_fname,
    p.officer_lname,
    p.officer_startdate,
    p.officer_branch
ORDER BY
    p.officer_branch;
```

Query Result:

	ACCIDENT COUNT	OFFICER_ID	OFFICER_FNAME	OFFICER_LNAME	OFFICER_STARTDATE	OFFICER_BRANCH
1	4	P126	Ella	Nirmala	22/MAR/18	Blackburn
2	4	P125	Jake	Rahardian	07/MAR/18	Blackburn
3	4	P128	Alice	Indira	13/APR/18	Box Hill
4	4	P127	Isabella	Adiratna	16/APR/18	Box Hill
5	8	P121	Daniel	Raditya	18/MAR/18	Donvale
6	8	P122	Tyler	Prasetyo	03/MAR/18	Donvale
7	2	P129	Aria	Kenanga	31/JAN/18	Mitcham
8	2	P130	Sofia	Naresha	16/APR/18	Mitcham
9	7	P123	Jayden	Perdana	20/APR/18	Ringwood
10	7	P124	Louie	Jayachandra	31/JAN/18	Ringwood

Question 2

A measure of the vehicle damage severity caused by the different accidents */

Explanation:

- Here, the management needs this data so that it can keep a tab on the damage caused by the accidents.
- This data is for creating **public awareness** so that people drive **safely and carefully.**
- When these statistics regarding the accidents and their severity are published by the management, the public will be more aware of the damage their negligence and lack of focus can cause.
- Hence, this will help **prevent further accidents** because of increased public awareness.

For this, we will have to add some additional attributes.

```
DROP TABLE accident_record_dim CASCADE CONSTRAINTS PURGE;

CREATE TABLE accident_record_dim AS
SELECT
    r.accident_no,
    a.accident_event,
    r.vehicle_damage_severity
FROM
    ACCIDENT.accident a,
```

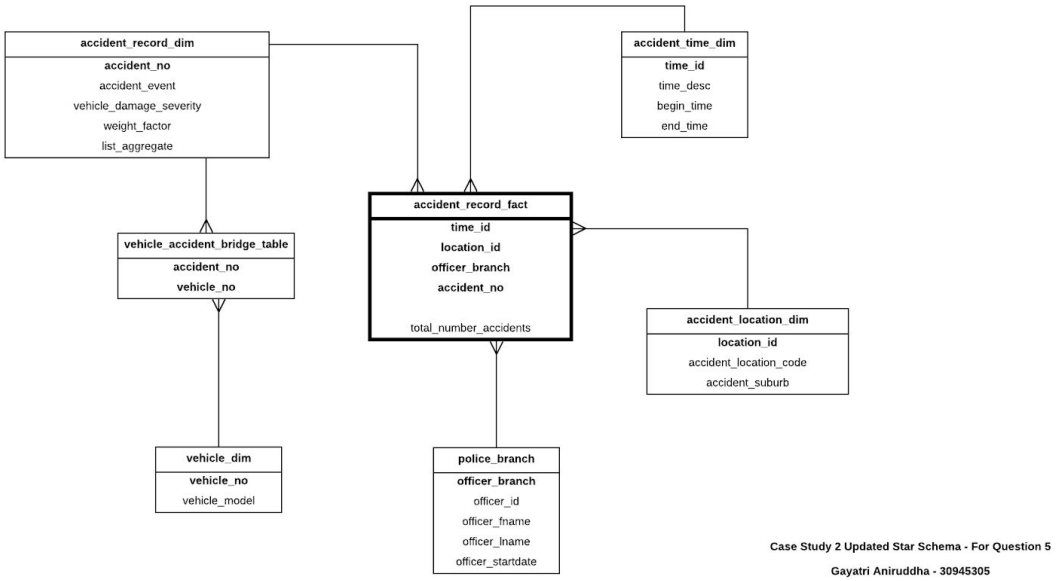
```
    ACCIDENT.accident_record r;

/* Query solution */
SELECT
    COUNT(f."total_number_of_accidents") AS "ACCIDENT COUNT",
    a.vehicle_damage_severity
FROM
    accident_record_fact f JOIN accident_record_dim a ON
    f.accident_no = a.accident_no
GROUP BY
     a.vehicle_damage_severity;
```

Query Result:

	ACCIDENT COUNT	VEHICLE_DAMAGE_SEVERITY
1	425	1
2	150	2
3	125	3

Changes made to the star-schema:

Here, in order to implement the above queries, I made some changes to star schema.
I have attached the picture of my updates star schema as well :



Case Study 2 Updated Star Schema - For Question 5

Gayatri Aniruddha - 30945305