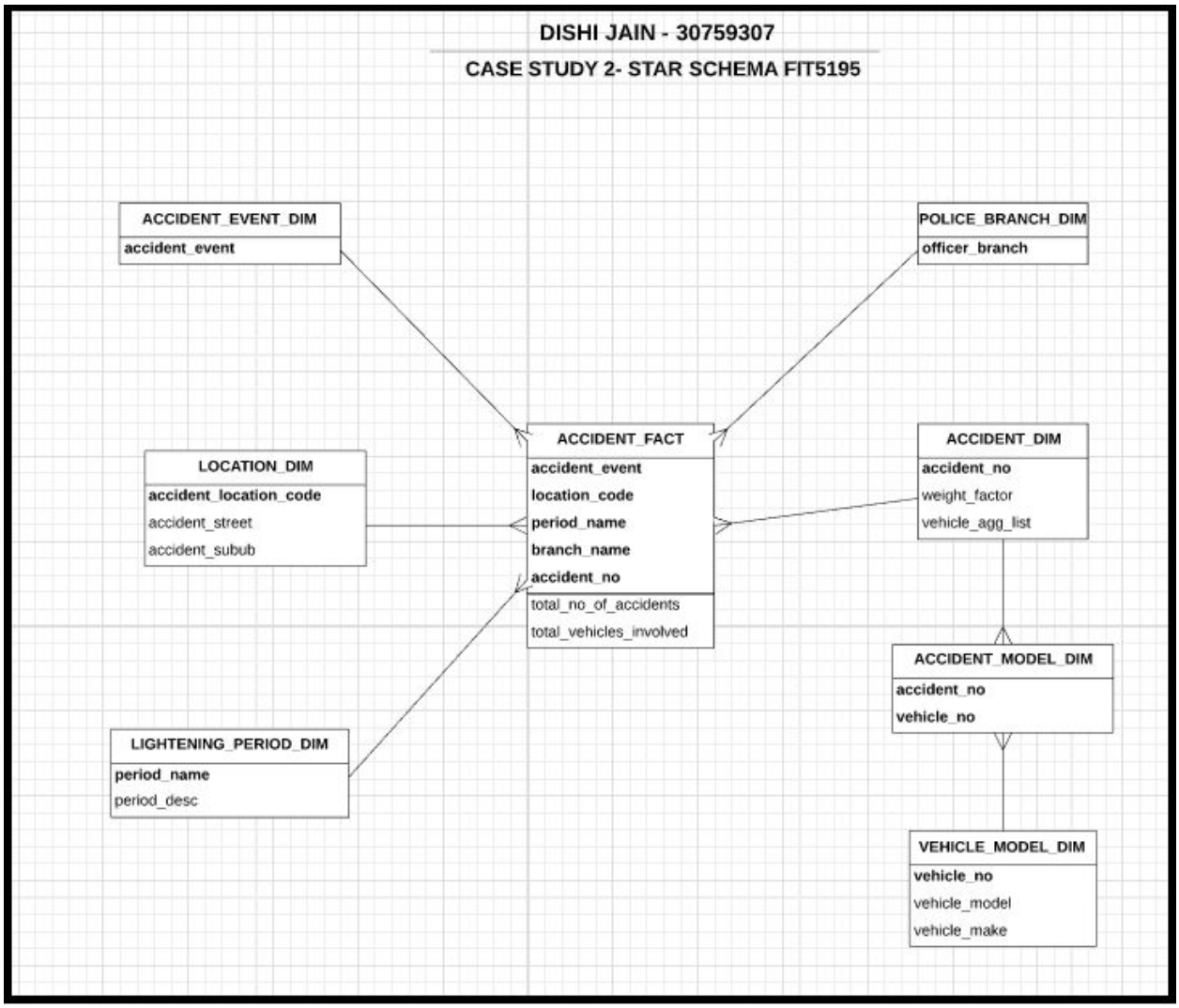
Case Study 2: Accident Records Data Warehouse

Tasks [1]:



Tasks [2]:

ACCIDENT_EVENT_DIM	Total_no_of_accidents	total_vehicles_involved
Fell from vehicle	7	2
Collision	11	22
Ran off carriageway	3	3
Rollover on/off carriageway	1	1

LOCATION_DIM	Total_no_of_accidents	total_vehicles_involved
Melbourne	6	12
Sydney	9	10
Richmond	3	5

LIGHTENING_PERIOD_DIM	Total_no_of_accidents	total_vehicles_involved
daytime	20	49
nighttime	23	30

POLICE_BRANCH_DIM	Total_no_of_accidents	total_vehicles_involved
Glen Huntly	6	12
Clayton	5	19
Glen waverley	7	12

THIS VEHICLE_MODEL_DIM doesn't make sense while calculating the fact measure - total_vehicles_involved. Hence this dim violates the two column table methodology.

VEHICLE_MODEL_DIM	Total_no_of_accidents	total_vehicles_involved
Yaris	3	?
Titan	7	?
Ninja 1000	4	?

Hence creating another dimension as ACCIDENT_DIM and connecting it with VEHICLE_MODEL_DIM through a bridge table ACCIDENT_MODEL_DIM.

Tasks [3]:

The script for creating dimensions and fact tables for the star schema are -

```
DROP TABLE accident_event_dim CASCADE CONSTRAINTS;

DROP TABLE location_dim CASCADE CONSTRAINTS;

DROP TABLE lightening_period_dim CASCADE CONSTRAINTS;

DROP TABLE police_branch_dim CASCADE CONSTRAINTS;

DROP TABLE accident_dim CASCADE CONSTRAINTS;

DROP TABLE accident_model_dim CASCADE CONSTRAINTS;

DROP TABLE vehicle_model_dim CASCADE CONSTRAINTS;

DROP TABLE temp_accident_fact CASCADE CONSTRAINTS;

DROP TABLE accident_fact CASCADE CONSTRAINTS;

CREATE TABLE accident_event_dim
   AS
      SELECT DISTINCT
         accident_event
```

```
    FROM
        accident.accident;

CREATE TABLE location_dim
    AS
        SELECT DISTINCT
            accident_location_code,
            accident_street,
            accident_suburb
        FROM
            accident.accident;

CREATE TABLE lightening_period_dim (
    period_name   VARCHAR2(50) NOT NULL,
    period_desc   VARCHAR2(50) NOT NULL
);

INSERT INTO lightening_period_dim VALUES (
    'daytime',
    '6AM - 5:59PM'
);

INSERT INTO lightening_period_dim VALUES (
    'nighttime',
    '6PM - 5:59AM'
);

CREATE TABLE police_branch_dim
    AS
        SELECT DISTINCT
            officer_branch
        FROM
            accident.police_officer;

CREATE TABLE accident_dim
    AS
        SELECT
            accident.a.accident_no,
            1.0 / COUNT(accident.v.vehicle_no) AS weight_factor,
            LISTAGG(accident.v.vehicle_no, '_') WITHIN GROUP(
                ORDER BY
                    accident.v.vehicle_no
            ) AS vehicle_agg_list
        FROM
            accident.accident          a,
            accident.accident_record   ar,
            accident.vehicle           v
        WHERE
```

```
        a.accident_no = ar.accident_no
        AND ar.vehicle_no = v.vehicle_no
    GROUP BY
        accident.a.accident_no;

CREATE TABLE accident_model_dim
  AS
    SELECT
        accident.a.accident_no,
        accident.v.vehicle_no
    FROM
        accident.accident         a,
        accident.accident_record  ar,
        accident.vehicle          v
    WHERE
        a.accident_no = ar.accident_no
        AND ar.vehicle_no = v.vehicle_no
    ORDER BY
        a.accident_no;

CREATE TABLE vehicle_model_dim
  AS
    SELECT DISTINCT
        vehicle_no,
        vehicle_model,
        vehicle_make
    FROM
        accident.vehicle;

CREATE TABLE temp_accident_fact
  AS
    SELECT
        a.accident_event,
        a.accident_location_code   AS location_code,
        TO_CHAR(a.accident_date_time, 'HH:MM:SS AM') AS accident_time,
        p.officer_branch          AS branch_name,
        a.accident_no,
        v.vehicle_no
    FROM
        accident.accident         a,
        accident.police_officer   p,
        accident.vehicle          v,
        accident.accident_record  ar
    WHERE
        a.accident_no = ar.accident_no
        AND ar.vehicle_no = v.vehicle_no
        AND a.officer_id = p.officer_id
    ORDER BY
```

```
        a.accident_no;

ALTER TABLE temp_accident_fact ADD (
   period_name VARCHAR2(50)
);

UPDATE temp_accident_fact
SET
   period_name =
      CASE
         WHEN CAST(accident_time AS time) BETWEEN CAST('06:00:00 am' AS time) AND
CAST('05:59:59 pm' AS time) THEN
            'Daytime'
         ELSE
            'Nighttime'
      END;

CREATE TABLE accident_fact
   AS
      SELECT
         accident_event,
         location_code,
         period_name,
         branch_name,
         accident_no,
         COUNT(accident_no) AS total_no_of_accidents,
         COUNT(vehicle_no) AS total_vehicles_involved
      FROM
         temp_accident_fact
      GROUP BY
         accident_event,
         location_code,
         period_name,
         branch_name,
         accident_no;



COMMIT;
```

CONTENTS OF THE DIMENSIONS AND FACT TABLES LOOK LIKE THIS-

ACCIDENT_EVENT_DIM

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

ACCIDENT_MODEL_DIM

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

ACCIDENT_DIM

	ACCIDENT_NO	WEIGHT_FACTOR	VEHICLE_AGG_LIST
1	A001	1	VC010
2	A002	1	VC011
3	A003	0.5	VC004_VT002
4	A004	0.5	VC001_VC006
5	A005	1	VM003
6	A006	0.5	VC006_VC007
7	A007	0.5	VC008_VM009
8	A008	0.5	VC012_VT013
9	A009	0.5	VC014_VC015
10	A010	1	VC010
11	A011	0.5	VC016_VC017
12	A012	0.5	VC018_VC019
13	A013	0.5	VC020_VC021
14	A014	1	VM009
15	A015	0.5	VC022_VC023
16	A016	0.5	VC024_VC025
17	A021	1	VC011

POLICE_BRANCH_DIM

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

ACCIDENT_FACT

	ACCIDENT_EVENT	LOCATION_CODE	PERIOD_NAME	BRANCH_NAME	ACCIDENT_NO	TOTAL_NO_OF_ACCIDENTS	TOTAL_VEHICLES_INVOLVED
1	Collision	CD2	Daytime	Donvale	A003	2	2
2	Collision	SD1	Daytime	Donvale	A004	2	2
3	Collision	CH1	Daytime	Ringwood	A008	2	2
4	Collision	MB1	Daytime	Blackburn	A012	2	2
5	Collision	MS1	Nighttime	Ringwood	A016	2	2
6	Ran off carriageway	SS1	Daytime	Donvale	A002	1	1
7	Collision	CC1	Nighttime	Donvale	A013	2	2
8	Rollover on/off carriageway	CD1	Daytime	Donvale	A001	1	1
9	Fell from vehicle	GB1	Daytime	Donvale	A005	1	1
10	Collision	CD1	Daytime	Ringwood	A006	2	2
11	Fell from vehicle	HD1	Daytime	Blackburn	A014	1	1
12	Collision	MC1	Nighttime	Ringwood	A009	2	2
13	Collision	GH1	Daytime	Donvale	A011	2	2
14	Collision	CW1	Daytime	Blackburn	A015	2	2
15	Ran off carriageway	CW1	Nighttime	Box Hill	A021	1	1
16	Collision	GW1	Nighttime	Ringwood	A007	2	2
17	Ran off carriageway	CD3	Nighttime	Blackburn	A010	1	1

LIGHTENING_PERIOD_DIM

	PERIOD_NAME	PERIOD_DESC
1	daytime	6AM - 5:59PM
2	nighttime	6PM - 5:59AM

LOCATION_DIM

	ACCIDENT_LOCATION_CODE	ACCIDENT_STREET	ACCIDENT_SUBURB
1	SS1	Springvale Road	Springvale
2	MS1	Springvale Road	Mulgrave
3	MS1	South Road	Morabbin
4	CH1	Huntingdale Road	Chadstone
5	HD1	Dandenong Road	Huntingdale
6	SD1	Blackburn Road	Syndal
7	GB1	Blackburn Road	Glen Waverley
8	SD1	Dandenong Road	Sandown
9	WC1	Centre Road	Westall
10	MC1	Clayton Road	Mount Waverley
11	CD3	Dandenong Road	Chadstone
12	CC2	Clayton Road	Clayton
13	RW1	Wellington Road	Rowville
14	RN1	North Road	Rowville
15	CD1	Dandenong Road	Carnegie
16	CD2	Dandenong Road	Clayton
17	CC1	Centre Road	Clarinda
18	MH1	Highbury Road	Mount Waverley
19	GW1	Waverley Road	Glen Waverley
20	MB1	Blackburn Road	Mulgrave
21	CW1	Warrigal Road	Chadstone
22	CW1	Wellington Road	Clayton
23	GH1	High Street Road	Glen Waverley

VEHICLE_MODEL_DIM

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

TEMP_ACCIDENT_FACT

	ACCIDENT_EVENT	LOCATION_CODE	ACCIDENT_TIME	BRANCH_NAME	ACCIDENT_NO	VEHICLE_NO	PERIOD_NAME
1	Rollover on/off carriageway	CD1	06:04:00 AM	Donvale	A001	VC010	Daytime
2	Ran off carriageway	SS1	11:04:00 AM	Donvale	A002	VC011	Daytime
3	Collision	CD2	01:04:00 PM	Donvale	A003	VC004	Daytime
4	Collision	CD2	01:04:00 PM	Donvale	A003	VT002	Daytime
5	Collision	SD1	10:04:00 AM	Donvale	A004	VC001	Daytime
6	Collision	SD1	10:04:00 AM	Donvale	A004	VC006	Daytime
7	Fell from vehicle	GB1	10:04:00 AM	Donvale	A005	VM003	Daytime
8	Collision	CD1	04:04:00 PM	Ringwood	A006	VC006	Daytime
9	Collision	CD1	04:04:00 PM	Ringwood	A006	VC007	Daytime
10	Collision	GW1	07:04:00 PM	Ringwood	A007	VC008	Nighttime
11	Collision	GW1	07:04:00 PM	Ringwood	A007	VM009	Nighttime
12	Collision	CH1	01:04:00 PM	Ringwood	A008	VC012	Daytime
13	Collision	CH1	01:04:00 PM	Ringwood	A008	VT013	Daytime
14	Collision	MC1	02:04:00 AM	Ringwood	A009	VC014	Nighttime
15	Collision	MC1	02:04:00 AM	Ringwood	A009	VC015	Nighttime
16	Ran off carriageway	CD3	02:04:00 AM	Blackburn	A010	VC010	Nighttime
17	Collision	GH1	09:04:00 AM	Donvale	A011	VC016	Daytime
18	Collision	GH1	09:04:00 AM	Donvale	A011	VC017	Daytime
19	Collision	MB1	07:04:00 AM	Blackburn	A012	VC018	Daytime
20	Collision	MB1	07:04:00 AM	Blackburn	A012	VC019	Daytime
21	Collision	CC1	04:04:00 AM	Donvale	A013	VC020	Nighttime
22	Collision	CC1	04:04:00 AM	Donvale	A013	VC021	Nighttime
23	Fell from vehicle	HD1	01:04:00 PM	Blackburn	A014	VM009	Daytime
24	Collision	CW1	09:04:00 AM	Blackburn	A015	VC022	Daytime
25	Collision	CW1	09:04:00 AM	Blackburn	A015	VC023	Daytime
26	Collision	MS1	10:04:00 PM	Ringwood	A016	VC024	Nighttime
27	Collision	MS1	10:04:00 PM	Ringwood	A016	VC025	Nighttime
28	Ran off carriageway	CW1	08:04:00 PM	Box Hill	A021	VC011	Nighttime

Tasks [4]:

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).
ANSWER -

```
SELECT
    l.accident_suburb,
    f.period_name,
    SUM(total_no_of_accidents) AS total_accidents
FROM
    accident_fact   f,
    location_dim    l
```

```
WHERE
   l.accident_location_code = f.location_code
GROUP BY
   l.accident_suburb,
   f.period_name
ORDER BY
   l.accident_suburb;
```

	ACCIDENT_SUBURB	PERIOD_NAME	TOTAL_ACCIDENTS
1	Carnegie	Daytime	3
2	Chadstone	Daytime	4
3	Chadstone	Nighttime	2
4	Clarinda	Nighttime	2
5	Clayton	Daytime	4
6	Clayton	Nighttime	1
7	Glen Waverley	Daytime	3
8	Glen Waverley	Nighttime	2
9	Huntingdale	Daytime	1
10	Morabbin	Nighttime	2
11	Mount Waverley	Nighttime	2
12	Mulgrave	Daytime	2
13	Mulgrave	Nighttime	2
14	Sandown	Daytime	2
15	Springvale	Daytime	1
16	Syndal	Daytime	2

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

```
SELECT
   v.vehicle_model,
   COUNT(f.total_no_of_accidents) AS total_accidents
FROM
   accident_fact       f,
```

```
    accident_dim         a,
    vehicle_model_dim    v,
    accident_model_dim   am
WHERE
    f.accident_no = a.accident_no
    AND a.accident_no = am.accident_no
    AND am.vehicle_no = v.vehicle_no
GROUP BY
    v.vehicle_model
ORDER BY
    total_accidents DESC;
```

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

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

```
SELECT
    f.accident_event,
    l.accident_suburb,
    SUM(f.total_vehicles_involved) as vehicles_involved
FROM
    accident_fact   f,
    location_dim    l
```

```
WHERE
   f.location_code = l.accident_location_code
GROUP BY
   f.accident_event,
   l.accident_suburb
ORDER BY
   f.accident_event;
```

	ACCIDENT_EVENT	ACCIDENT_SUBURB	VEHICLES_INVOLVED
1	Collision	Carnegie	2
2	Collision	Chadstone	4
3	Collision	Clarinda	2
4	Collision	Clayton	4
5	Collision	Glen Waverley	4
6	Collision	Morabbin	2
7	Collision	Mount Waverley	2
8	Collision	Mulgrave	4
9	Collision	Sandown	2
10	Collision	Syndal	2
11	Fell from vehicle	Glen Waverley	1
12	Fell from vehicle	Huntingdale	1
13	Ran off carriageway	Chadstone	2
14	Ran off carriageway	Clayton	1
15	Ran off carriageway	Springvale	1
16	Rollover on/off carriageway	Carnegie	1

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

```
SELECT
   branch_name,
   SUM(total_no_of_accidents) AS total_accidents
FROM
   accident_fact
GROUP BY
   branch_name
ORDER BY
```

branch_name;

	BRANCH_NAME	TOTAL_ACCIDENTS
1	Blackburn	6
2	Box Hill	1
3	Donvale	11
4	Ringwood	10

Tasks [5]:

--Q1 Show the number of vehicles involved by different lighting periods (daytime: 6AM - 5:59PM and nighttime 6PM - 5:59AM).
The management would like to see the number of vehicles involved during the time period of the day. This would help them to find a relationship between the time period and vehicles involved to find during which time most number of vehicles are present. From the query it can be seen that during daytime 18 vehicles were involved for the accidents and during night time 10 vehicles were involved.

```
SELECT
   period_name,
   SUM(total_vehicles_involved) AS total_vehicles
FROM
   accident_fact
GROUP BY
   Period_name;
```

	PERIOD_NAME	TOTAL_VEHICLES
1	Daytime	18
2	Nighttime	10

--Q1 Show the total number of accidents for each accident event
The management would like to see the total accidents present for the different accident_event types. It would help them to determine which kind of accidents are seen the most of the time. From the query it can be seen that collision type accidents are most frequently seen and rollover on/off carriageway are least frequently seen.

```
SELECT
   accident_event,
   COUNT(total_no_of_accidents) AS total_accidents
FROM
   accident_fact
GROUP BY
   accident_event;
```

	ACCIDENT_EVENT	TOTAL_ACCIDENTS
1	Collision	11
2	Fell from vehicle	2
3	Ran off carriageway	3
4	Rollover on/off carriageway	1