Tutorial 8
Simple OLAP

We are going to use Robcor's Aircraft CHARTER data warehouse in this exercise. If you have your Robcor's Data Warehouse in your account which you have created during the Robcor Case study, you can use those tables to do the following exercise.

Alternatively, you can use the CHARTER data warehousing in the `dw` account. In this case, the tables that are relevant to the Robcor's CHARTER data warehouse are: `dw.charter_fact`, `dw.time`, `dw.pilot`, and `dw.model`.

A. Explore the data warehouse

1. Describe the structure of the three dimension tables. What are the attributes of each of the dimension tables?

```
-- Time dimension

Describe dw.Time;

 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 TIME_ID                                            CHAR(6)
 TIME_YEAR                                          CHAR(4)
 TIME_MONTH                                         CHAR(2)

-- Pilot dimension

Describe dw.Pilot;

 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 EMP_NUM                                            NUMBER(10)
 PIL_LICENSE                                        CHAR(25)
 PIL_RATINGS                                        CHAR(25)
 PIL_MED_TYPE                                       CHAR(1)
 PIL_MED_DATE                                       DATE
 PIL_PT135_DATE                                     DATE

-- Model dimension

Describe dw.Model;

 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 MOD_CODE                                           CHAR(10)
 MOD_MANUFACTURER                                   CHAR(15)
 MOD_NAME                                           CHAR(20)
 MOD_SEATS                                          FLOAT(126)
 MOD_CHG_MILE                                       NUMBER(19,4)
 MOD_CRUISE                                         FLOAT(126)
 MOD_FUEL                                           FLOAT(126)
```

2. Describe the structure of the fact table – what attributes does it have?

```
--Describe dw.Charter_Fact;

 Name                                      Null?    Type
 ----------------------------------------- -------- ----------------------------
 TIME_ID                                            VARCHAR2(6)
 MOD_CODE                                           CHAR(10)
 EMP_NUM                                            NUMBER(10)
 TOT_CHAR_HOURS                                     NUMBER
 TOT_FUEL                                           NUMBER
 REVENUE                                            NUMBER
```

3. Display the contents of each of the dimension tables; some of the dimension tables are not that big.

```
Select * from dw.Time;
Select * from dw.Pilot;
Select * from dw.Model;
```

4. Display the contents of the fact table.

```
Select * from dw.Charter_Fact;
```

B. Simple aggregate exercises using GROUP BY

1. What is the total hours flown by each pilot?

```
select emp_num, sum(tot_char_hours) as
Total_hours_flown
from dw.charter_fact
group by emp_num;
```

```
   EMP_NUM TOTAL_HOURS_FLOWN
---------- -----------------
       101             672.7
       104             716.5
       109             721.5
       105               739
       106             965.1
```

2. Display the total hours flown by each pilot in a descending order.

```
select emp_num, sum(tot_char_hours) as Total_hours_flown
from dw.charter_fact
group by emp_num
order by sum(tot_char_hours) desc;
```

```
   EMP_NUM TOTAL_HOURS_FLOWN
---------- -----------------
       106             965.1
       105               739
       109             721.5
       104             716.5
       101             672.7
```

3. What is the total hours flown by each category of pilot license?

```
select p.pil_license, sum(c.tot_char_hours) as Total_hours_flown
from dw.pilot p, dw.charter_fact c
where p.emp_num = c.emp_num
group by p.pil_license;

PIL_LICENSE                TOTAL_HOURS_FLOWN
-------------------------- -----------------
COM                                   2425.6
ATP                                   1389.2
```

4. What is the total revenue generated by each pilot? Sort the results based on the Pilot ID

```
select emp_num, sum(revenue) as Total_revenue
from dw.charter_fact
group by emp_num
order by emp_num;

   EMP_NUM TOTAL_REVENUE
---------- -------------
       101     280972.05
       104      299024.4
       105     306901.79
       106     379493.77
       109     326897.14
```

5. What is the total fuel consumption of the aircrafts manufactured by each manufacturer?

```
select m.mod_manufacturer, sum(c.tot_fuel) as Total_Fuel_used
from dw.model m, dw.charter_fact c
where m.mod_code = c.mod_code
group by mod_manufacturer;

MOD_MANUFACTURE TOTAL_FUEL_USED
--------------- ---------------
Beechcraft               61708.4
Piper                   115055.5
```

6. What is the total revenue generated in each year?

```
select t.time_year, sum(c.revenue) as Total_revenue
from dw.time t, dw.charter_fact c
where t.time_id=c.time_id
group by t.time_year;

TIME TOTAL_REVENUE
```

```
---- -------------
1997     217264.37
1994     452579.08
1995     472881.69
1996     450564.01
```

C. More complex aggregate exercises using CUBE and ROLLUP

The OLAP queries you need to implement are the following:

1. What is the total fuel used from Oct to Dec 1995 by commercial pilots and airplane model C-90A. Sort the results by the month. How many rows of records do you get?

```
-- Group By
SELECT
   time_id As Period,
   c.emp_num AS Pilot,
   mod_code As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY time_id, c.emp_num, mod_code
ORDER BY time_id;
```

```
PERIOD      PILOT MODEL      SUM(TOT_FUEL)
------ ---------- ---------- -------------
199510         105 C-90A              525.2
199510         109 C-90A              302.9
199511         106 C-90A              272.8
199511         109 C-90A              115.8
199512         105 C-90A              106.6
199512         106 C-90A              575.7
199512         109 C-90A              295.8

7 rows selected.
```

2. Using **cube**, what is the total fuel used from Oct to Dec 1995 by commercial pilots and airplane model C-90A. Sort the results by the month. How many rows of records do you get?

```
-- CUBE (without DECODE)
SELECT
   time_id As Period,
   c.emp_num AS Pilot,
```

```
   mod_code As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY CUBE (time_id, c.emp_num, mod_code)
ORDER BY time_id;

PERIOD      PILOT MODEL      SUM(TOT_FUEL)
------ ---------- ---------- -------------
199510         105 C-90A              525.2
199510         105                    525.2
199510         109 C-90A              302.9
199510         109                    302.9
199510             C-90A              828.1
199510                                828.1
199511         106 C-90A              272.8
199511         106                    272.8
199511         109 C-90A              115.8
199511         109                    115.8
199511             C-90A              388.6

PERIOD      PILOT MODEL      SUM(TOT_FUEL)
------ ---------- ---------- -------------
199511                                388.6
199512         105 C-90A              106.6
199512         105                    106.6
199512         106 C-90A              575.7
199512         106                    575.7
199512         109 C-90A              295.8
199512         109                    295.8
199512             C-90A              978.1
199512                                978.1
               105 C-90A              631.8
               105                    631.8

PERIOD      PILOT MODEL      SUM(TOT_FUEL)
------ ---------- ---------- -------------
               106 C-90A              848.5
               106                    848.5
               109 C-90A              714.5
               109                    714.5
                   C-90A             2194.8
                                     2194.8

28 rows selected.
```

3. Redo question C.2 using Grouping. Notes that “1” and “0” in the TIME, PILOT, and MODEL indicate that they are aggregate values and real values respectively.

```
-- CUBE (with GROUPING)
SELECT
   time_id As Period,
   c.emp_num AS Pilot,
   mod_code As Model,
   SUM(tot_fuel),
   GROUPING(time_id) As PeriodGroup,
   GROUPING(c.emp_num) AS PilotGroup,
   GROUPING(mod_code) As ModelGroup
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY CUBE (time_id, c.emp_num, mod_code)
ORDER BY time_id;
```

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)	PERIODGROUP	PILOTGROUP	MODELGROUP
199510	105	C-90A	525.2	0	0	0
199510	105		525.2	0	0	1
199510	109	C-90A	302.9	0	0	0
199510	109		302.9	0	0	1
199510		C-90A	828.1	0	1	0
199510			828.1	0	1	1
199511	106	C-90A	272.8	0	0	0
199511	106		272.8	0	0	1
199511	109	C-90A	115.8	0	0	0
199511	109		115.8	0	0	1
199511		C-90A	388.6	0	1	0

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)	PERIODGROUP	PILOTGROUP	MODELGROUP
199511			388.6	0	1	1
199512	105	C-90A	106.6	0	0	0
199512	105		106.6	0	0	1
199512	106	C-90A	575.7	0	0	0
199512	106		575.7	0	0	1
199512	109	C-90A	295.8	0	0	0
199512	109		295.8	0	0	1
199512		C-90A	978.1	0	1	0
199512			978.1	0	1	1
	105	C-90A	631.8	1	0	0
	105		631.8	1	0	1

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)	PERIODGROUP	PILOTGROUP	MODELGROUP
	106	C-90A	848.5	1	0	0
	106		848.5	1	0	1
	109	C-90A	714.5	1	0	0
	109		714.5	1	0	1
		C-90A	2194.8	1	1	0
			2194.8	1	1	1

28 rows selected.

4. As like question C.3 above, but instead of using “0” and “1”, it displays “All Periods”, “All Pilots” and “All Models” instead. (Hints: Use DECODE).

```
-- Cube
SELECT
   DECODE(GROUPING(time_id), 1, 'All Periods',
time_id) As Period,
   DECODE(GROUPING(c.emp_num), 1, 'All Pilots',
c.emp_num) AS Pilot,
   DECODE(GROUPING(mod_code), 1, 'All Models',
mod_code) As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY CUBE (time_id, c.emp_num, mod_code)
ORDER BY time_id;
```

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)
199510	105	C-90A	525.2
199510	105	All Models	525.2
199510	109	C-90A	302.9
199510	109	All Models	302.9
199510	All Pilots	C-90A	828.1
199510	All Pilots	All Models	828.1
199511	106	C-90A	272.8
199511	106	All Models	272.8
199511	109	C-90A	115.8
199511	109	All Models	115.8
199511	All Pilots	C-90A	388.6

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)
199511	All Pilots	All Models	388.6
199512	105	C-90A	106.6
199512	105	All Models	106.6
199512	106	C-90A	575.7
199512	106	All Models	575.7
199512	109	C-90A	295.8
199512	109	All Models	295.8
199512	All Pilots	C-90A	978.1
199512	All Pilots	All Models	978.1
All Periods	105	C-90A	631.8
All Periods	105	All Models	631.8

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)
All Periods	106	C-90A	848.5
All Periods	106	All Models	848.5
All Periods	109	C-90A	714.5
All Periods	109	All Models	714.5
All Periods	All Pilots	C-90A	2194.8
All Periods	All Pilots	All Models	2194.8

28 rows selected.

5. Following the results in question C.4, since there is only one aircraft model in the query results (e.g. C-90A), it seems that the "All Models" are redundant. Now, we want to remove them from the report, as there is no point displaying "All Models" when there is only one model (Hints: Use Partial CUBE).

```
-- Partial Cube
SELECT
   DECODE(GROUPING(time_id),    1,    'All    Periods',
time_id) As Period,
   DECODE(GROUPING(c.emp_num),    1,    'All    Pilots',
c.emp_num) AS Pilot,
   DECODE(GROUPING(mod_code),    1,    'All    Models',
mod_code) As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY CUBE (time_id, c.emp_num), mod_code
ORDER BY time_id;
```

```
PERIOD      PILOT                                  MODEL      SUM(TOT_FUEL)
----------- -------------------------------------- ---------- -------------
199510      105                                    C-90A              525.2
199510      109                                    C-90A              302.9
199510      All Pilots                             C-90A              828.1
199511      106                                    C-90A              272.8
199511      109                                    C-90A              115.8
199511      All Pilots                             C-90A              388.6
199512      105                                    C-90A              106.6
199512      106                                    C-90A              575.7
199512      109                                    C-90A              295.8
199512      All Pilots                             C-90A              978.1
All Periods 105                                    C-90A              631.8

PERIOD      PILOT                                  MODEL      SUM(TOT_FUEL)
----------- -------------------------------------- ---------- -------------
All Periods 106                                    C-90A              848.5
All Periods 109                                    C-90A              714.5
All Periods All Pilots                             C-90A             2194.8

14 rows selected.
```

6. Using **rollup** with **decode**, what is the total fuel used from Oct to Dec 1995 by commercial pilots and airplane model C-90A. Sort the results by the month. How many rows of records do you get?

```
-- Roll up
SELECT
   DECODE(GROUPING(time_id),    1,    'All    Periods',
time_id) As Period,
   DECODE(GROUPING(c.emp_num),    1,    'All    Pilots',
c.emp_num) AS Pilot,
```

```
   DECODE(GROUPING(mod_code),    1,    'All    Models',
mod_code) As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY ROLLUP (time_id, c.emp_num, mod_code)
ORDER BY time_id;
```

```
PERIOD      PILOT                                  MODEL      SUM(TOT_FUEL)
----------- -------------------------------------- ---------- -------------
199510      105                                    C-90A              525.2
199510      105                                    All Models         525.2
199510      109                                    C-90A              302.9
199510      109                                    All Models         302.9
199510      All Pilots                             All Models         828.1
199511      106                                    C-90A              272.8
199511      106                                    All Models         272.8
199511      109                                    C-90A              115.8
199511      109                                    All Models         115.8
199511      All Pilots                             All Models         388.6
199512      105                                    C-90A              106.6

PERIOD      PILOT                                  MODEL      SUM(TOT_FUEL)
----------- -------------------------------------- ---------- -------------
199512      105                                    All Models         106.6
199512      106                                    C-90A              575.7
199512      106                                    All Models         575.7
199512      109                                    C-90A              295.8
199512      109                                    All Models         295.8
199512      All Pilots                             All Models         978.1
All Periods All Pilots                             All Models        2194.8

18 rows selected.
```

7. Compare the results in C.2 and C.6. What is the difference?

8. Modify C.6 to use Partial Roll up (exclude “All Models” from the rollup).

```
-- Partial Roll up
SELECT
   DECODE(GROUPING(time_id), 1, 'All Periods', time_id)
As Period,
   DECODE(GROUPING(c.emp_num),    1,    'All    Pilots',
c.emp_num) AS Pilot,
   DECODE(GROUPING(mod_code),    1,    'All    Models',
mod_code) As Model,
   SUM(tot_fuel)
FROM dw.charter_fact c, dw.pilot p
WHERE c.emp_num = p.emp_num
AND time_id LIKE '19951%'
AND mod_code = 'C-90A'
AND p.pil_license = 'COM'
GROUP BY ROLLUP (time_id, c.emp_num), mod_code
ORDER BY time_id;
```

PERIOD	PILOT	MODEL	SUM(TOT_FUEL)
199510	105	C-90A	525.2
199510	109	C-90A	302.9
199510	All Pilots	C-90A	828.1
199511	106	C-90A	272.8
199511	109	C-90A	115.8
199511	All Pilots	C-90A	388.6
199512	105	C-90A	106.6
199512	106	C-90A	575.7
199512	109	C-90A	295.8
199512	All Pilots	C-90A	978.1
All Periods	All Pilots	C-90A	2194.8

11 rows selected.

THE END