MODULE 6 HOMEWORK ASSIGNMENT 5 – USING A DIMENSIONAL MODEL WITH THE SALECO DATA WAREHOUSE

Steve Desilets

MSDS 420: Database Systems and Data Preparation

May 7, 2023

1. **List the total sales by region and customer. Your output should be sorted by region name and customer code.**

   The code and output below display the total sales disaggregated by region and customer.

Query | Query History

```
1  SELECT  cus_code,
2          cus_lname,
3          cus_fname,
4          cus_initial,
5          reg_name,
6          SUM(day_subtotal) AS total_sales
7  FROM ksdaysales_cus_reg
8  GROUP BY cus_code,
9           cus_lname,
10          cus_fname,
11          cus_initial,
12          reg_name
13 ORDER BY reg_name, cus_code;
```

	cus_code integer	cus_lname character varying (15)	cus_fname character varying (15)	cus_initial character (1)	reg_name character varying (10)	total_sales numeric
1	10012	Smith	Kathy	W	NE	287.91
2	10013	Olowski	Paul	F	NE	64.32
3	10014	Orlando	Myron	[null]	NW	494.71
4	10019	Smith	Olette	K	NW	39.95
5	10010	Ramas	Alfred	A	SE	180.26
6	10011	Dunne	Leona	K	SE	130.89
7	10015	O'Brian	Amy	B	SE	325.82
8	10016	Brown	James	G	SE	179.22
9	10017	Williams	George	[null]	SW	419.66
10	10018	Farriss	Anne	G	SW	129.32

2. Repeat #1 but produce the output using ROLLUP with region name and customer code.

The code and output below display the total sales disaggregated by region and customer using the ROLLUP function.

Query | Query History

```
SELECT reg_name,
       cus_code,
       CONCAT(cus_fname, ' ', cus_initial, ' ', cus_lname) AS customer_name,
       SUM(day_subtotal) AS total_sales
FROM Ksdaysales_cus_reg
GROUP BY ROLLUP (reg_name,
                 cus_code,
                 customer_name)
ORDER BY reg_name, cus_code;
```

	reg_name character varying (10)	cus_code integer	customer_name text	total_sales numeric
1	NE	10012	Kathy W Smith	287.91
2	NE	10012	[null]	287.91
3	NE	10013	Paul F Olowski	64.32
4	NE	10013	[null]	64.32
5	NE	[null]	[null]	352.23
6	NW	10014	Myron Orlando	494.71
7	NW	10014	[null]	494.71
8	NW	10019	Olette K Smith	39.95
9	NW	10019	[null]	39.95
10	NW	[null]	[null]	534.66
11	SE	10010	Alfred A Ramas	180.26
12	SE	10010	[null]	180.26
13	SE	10011	Leona K Dunne	130.89
14	SE	10011	[null]	130.89
15	SE	10015	Amy B O'Brian	325.82
16	SE	10015	[null]	325.82
17	SE	10016	James G Brown	179.22
18	SE	10016	[null]	179.22
19	SE	[null]	[null]	816.19
20	SW	10017	George Williams	419.66
21	SW	10017	[null]	419.66
22	SW	10018	Anne G Farriss	129.32
23	SW	10018	[null]	129.32
24	SW	[null]	[null]	548.98
25	[null]	[null]	[null]	2252.06

3. Repeat #1 but product the output using CUBE with region name and customer code.

The code and output below display the total sales disaggregated by region and customer using the CUBE function.

Query | Query History

```
1  SELECT reg_name,
2         cus_code,
3         SUM(day_subtotal) AS total_sales
4  FROM Ksdaysales_cus_reg
5  GROUP BY CUBE (reg_name,
6                 cus_code)
7  ORDER BY reg_name, cus_code;
```

	reg_name character varying (10)	cus_code integer	total_sales numeric
1	NE	10012	287.91
2	NE	10013	64.32
3	NE	[null]	352.23
4	NW	10014	494.71
5	NW	10019	39.95
6	NW	[null]	534.66
7	SE	10010	180.26
8	SE	10011	130.89
9	SE	10015	325.82
10	SE	10016	179.22
11	SE	[null]	816.19
12	SW	10017	419.66
13	SW	10018	129.32
14	SW	[null]	548.98
15	[null]	10010	180.26
16	[null]	10011	130.89
17	[null]	10012	287.91
18	[null]	10013	64.32
19	[null]	10014	494.71
20	[null]	10015	325.82
21	[null]	10016	179.22
22	[null]	10017	419.66
23	[null]	10018	129.32
24	[null]	10019	39.95
25	[null]	[null]	2252.06

4. **a) Explain the additional information/intelligence gained when using ROLLUP or CUBE b) Use the output from questions 1, 2 and 3 to explain what the data reveals.**

   The rollup and cube functions instructs SQL to display subtotals and grand totals for each of the columns for which we are aggregating the data. In this case, that means that we are displaying the subtotals and grand total for sales for the various combinations of regions and customers.

   There are a number of findings that we can glean from the outputs of the queries from questions 1, 2, and 3. For example, the total sales were $2,252.06, the top region as measured by total sales was the Southeast region (with $816.19 in sales), and the top customer as measured by total sales was Myron Orlando (with $494.71 in sales).

5. List the total sales by customer code, month, and product code; sort by customer code and month.

The code and output below display the total sales by customer code, month, and product code

Query | Query History

```
1
2  SELECT d.cus_code,
3         t.tm_month,
4         d.p_code,
5         SUM(d.day_subtotal) AS Total_Sales
6  FROM ksdaysales_cus_reg d
7  LEFT JOIN dwtime t
8  ON d.tm_id = t.tm_id
9  GROUP BY d.cus_code, t.tm_month, d.p_code
10 ORDER BY d.cus_code, t.tm_month
11
```

Data Output | Messages | Notifications

	cus_code integer	tm_month integer	p_code character varying (10)	total_sales numeric
1	10010	10	13-Q2/P2	74.95
2	10010	10	23109-HB	19.90
3	10010	10	54778-2T	14.97
4	10010	10	PVC23DRT	70.44
5	10011	10	2232/QTY	109.92
6	10011	10	SM-18277	20.97
7	10012	9	SM-18277	20.97
8	10012	10	23109-HB	9.95
9	10012	10	89-WRE-Q	256.99
10	10013	10	13-Q2/P2	29.98
11	10013	10	54778-2T	4.99
12	10013	10	PVC23DRT	29.35
13	10014	9	13-Q2/P2	14.99
14	10014	9	2232/QTY	109.92
15	10014	9	23109-HB	9.95
16	10014	10	WR3/TT3	359.85
17	10015	9	2238/QPD	38.95
18	10015	9	23109-HB	9.95
19	10015	9	54778-2T	9.98
20	10015	9	89-WRE-Q	256.99
21	10015	10	23109-HB	9.95
22	10016	9	13-Q2/P2	104.93
23	10016	9	1546-QQ2	39.95
24	10016	9	54778-2T	4.99
25	10016	9	PVC23DRT	29.35
26	10017	9	13-Q2/P2	14.99
27	10017	9	23109-HB	29.85
28	10017	9	54778-2T	14.97
29	10017	9	WR3/TT3	359.85
30	10018	9	2238/QPD	38.95
31	10018	9	23109-HB	9.95
32	10018	9	54778-2T	9.98
33	10018	9	PVC23DRT	70.44
34	10019	9	1546-QQ2	39.95

6. **Show all purchases (total sales) in September to show which customer bought the most product in September. Show customer code, customer name and total sales; sort all output by total sales with the highest sales on top.**

   The code and output below display the total sales in September. Using this output, we can determine that George Williams was the customer who bought the most product in September based on total sales.

Query | Query History

```
1  SELECT cus_code,
2         CONCAT(cus_fname, ' ', cus_initial, ' ', cus_lname) AS customer_name,
3         SUM(day_subtotal) AS total_sales
4  FROM Ksdaysales_cus_reg
5  WHERE tm_id IN (SELECT tm_id FROM dwtime WHERE tm_month = 9)
6  GROUP BY cus_code, customer_name
7  ORDER BY total_sales DESC;
```

	cus_code integer	customer_name text	total_sales numeric
1	10017	George Williams	419.66
2	10015	Amy B O'Brian	315.87
3	10016	James G Brown	179.22
4	10014	Myron Orlando	134.86
5	10018	Anne G Farriss	129.32
6	10019	Olette K Smith	39.95
7	10012	Kathy W Smith	20.97

7. **List the total sales by month and product category. Your output should be sorted by month and product category.**

The code and output below display the total sales disaggregated by month and product category.

Query | Query History

```
1  SELECT t.tm_month,
2         p.p_category,
3         SUM(d.day_subtotal) AS total_sales
4  FROM Ksdaysales_cus_reg d
5  LEFT JOIN dwtime t
6      ON d.tm_id = t.tm_id
7  LEFT JOIN dwproduct p
8      ON d.p_code = p.p_code
9  GROUP BY t.tm_month, p.p_category
10 ORDER BY t.tm_month, p.p_category
```

Data Output | Messages | Notifications

	tm_month integer	p_category character varying (5)	total_sales numeric
1	9	CAT1	174.83
2	9	CAT2	446.81
3	9	CAT3	537.54
4	9	CAT4	80.67
5	10	CAT1	124.89
6	10	CAT2	366.91
7	10	CAT3	459.64
8	10	CAT4	60.77

8. **List the number of product sales (number of rows) and total sales by month. Your output should be sorted by month and should show one row per month.**

   The code and output below display the number of product sales and total sales disaggregated by month.

Query | Query History

```
1  SELECT t.tm_month,
2         SUM(d.sale_units) AS total_sale_units,
3         SUM(d.day_subtotal) AS total_sales
4  FROM Ksdaysales_cus_reg d
5  LEFT JOIN dwtime t
6      ON d.tm_id = t.tm_id
7  GROUP BY t.tm_month
8  ORDER BY t.tm_month
9
```

Data Output | Messages | Notifications

	tm_month integer	total_sale_units bigint	total_sales numeric
1	9	52	1239.85
2	10	40	1012.21

9. **Show product category, product code, product description and units sold (sum). Which product is the best seller based on units sold?**

The code and output below display the total units sold for each product (across all time periods). As displayed in the table, the best-selling product based on units sold (across all time periods) is product PVC23DRT – the 3.5 inch by 8 foot PVC pipe.

Query | Query History

```
1  SELECT
2      d.p_code,
3      p.p_category,
4      p.p_descript,
5      SUM(d.sale_units) AS total_units_sold
6  FROM ksdaysales_cus_reg d
7  LEFT JOIN dwproduct p
8      ON d.p_code = p.p_code
9  GROUP BY d.p_code, p.p_category, p.p_descript
10 ORDER BY total_units_sold DESC;
11
```

Data Output | Messages | Notifications

	p_code character varying (10)	p_category character varying (5)	p_descript character varying (35)	total_units_sold bigint
1	PVC23DRT	CAT3	PVC pipe, 3.5-in., 8-ft	34
2	13-Q2/P2	CAT1	7.25-in. pwr. saw blade	16
3	54778-2T	CAT1	Rat-tail file, 1/8-in. fine	12
4	23109-HB	CAT4	Claw hammer	10
5	SM-18277	CAT4	1.25-in. metal screw, 25	6
6	WR3/TT3	CAT3	Steel matting, 4'x8'x1/6", .5" mesh	6
7	89-WRE-Q	CAT2	Hicut chain saw, 16 in.	2
8	1546-QQ2	CAT2	Hrd. cloth, 1/4-in., 2x50	2
9	2238/QPD	CAT3	B\&D cordless drill, 1/2-in.	2
10	2232/QTY	CAT2	B\&D jigsaw, 12-in. blade	2

9a) Show units sold for September

The code and output below display the total units sold for each product in September. As displayed in the table, the best-selling product based on units sold in September is product PVC23DRT – the 3.5 inch by 8 foot PVC pipe.

Query | Query History

```
1  SELECT
2      d.p_code,
3      p.p_category,
4      p.p_descript,
5      SUM(d.sale_units) AS total_units_sold
6  FROM ksdaysales_cus_reg d
7  LEFT JOIN dwproduct p
8      ON d.p_code = p.p_code
9  WHERE d.tm_id IN (
10     SELECT tm_id FROM dwtime WHERE tm_month = 9)
11 GROUP BY d.p_code, p.p_category, p.p_descript
12 ORDER BY total_units_sold DESC;
```

Data Output | Messages | Notifications

	p_code character varying (10)	p_category character varying (5)	p_descript character varying (35)	total_units_sold bigint
1	PVC23DRT	CAT3	PVC pipe, 3.5-in., 8-ft	17
2	13-Q2/P2	CAT1	7.25-in. pwr. saw blade	9
3	54778-2T	CAT1	Rat-tail file, 1/8-in. fine	8
4	23109-HB	CAT4	Claw hammer	6
5	SM-18277	CAT4	1.25-in. metal screw, 25	3
6	WR3/TT3	CAT3	Steel matting, 4'x8'x1/6", .5" mesh	3
7	2238/QPD	CAT3	B\&D cordless drill, 1/2-in.	2
8	1546-QQ2	CAT2	Hrd. cloth, 1/4-in., 2x50	2
9	89-WRE-Q	CAT2	Hicut chain saw, 16 in.	1
10	2232/QTY	CAT2	B\&D jigsaw, 12-in. blade	1

9b) Show units sold for October

The code and output below display the total units sold for each product in September. As displayed in the table, the best-selling product based on units sold in September is product PVC23DRT – the 3.5 inch by 8 foot PVC pipe.

Query Query History

```
1  SELECT
2      d.p_code,
3      p.p_category,
4      p.p_descript,
5      SUM(d.sale_units) AS total_units_sold
6  FROM ksdaysales_cus_reg d
7  LEFT JOIN dwproduct p
8      ON d.p_code = p.p_code
9  WHERE d.tm_id IN (
10     SELECT tm_id FROM dwtime WHERE tm_month = 10)
11 GROUP BY d.p_code, p.p_category, p.p_descript
12 ORDER BY total_units_sold DESC;
```

Data Output Messages Notifications

	p_code character varying (10)	p_category character varying (5)	p_descript character varying (35)	total_units_sold bigint
1	PVC23DRT	CAT3	PVC pipe, 3.5-in., 8-ft	17
2	13-Q2/P2	CAT1	7.25-in. pwr. saw blade	7
3	23109-HB	CAT4	Claw hammer	4
4	54778-2T	CAT1	Rat-tail file, 1/8-in. fine	4
5	WR3/TT3	CAT3	Steel matting, 4'x8'x1/6", .5" mesh	3
6	SM-18277	CAT4	1.25-in. metal screw, 25	3
7	2232/QTY	CAT2	B\&D jigsaw, 12-in. blade	1
8	89-WRE-Q	CAT2	Hicut chain saw, 16 in.	1

10. **List the number of product sales (number of rows) and total sales by month, product category, and product. Your output should be sorted by month, product category and product.**

The code and output below display the number of product sales and total sales disaggregated by month and product.

Query | Query History

```
SELECT t.tm_month,
       p.p_category,
       p.p_code,
       COUNT(d.sale_units) AS number_of_sales,
       SUM(d.day_subtotal) AS total_sales
FROM Ksdaysales_cus_reg d
LEFT JOIN dwtime t
    ON d.tm_id = t.tm_id
LEFT JOIN dwproduct p
    ON d.p_code = p.p_code
GROUP BY t.tm_month, p.p_category, p.p_code
ORDER BY t.tm_month, p.p_category, p.p_code
```

Data Output | Messages | Notifications

	tm_month integer	p_category character varying (5)	p_code character varying (10)	number_of_sales bigint	total_sales numeric
1	9	CAT1	13-Q2/P2	4	134.91
2	9	CAT1	54778-2T	4	39.92
3	9	CAT2	1546-QQ2	2	79.90
4	9	CAT2	2232/QTY	1	109.92
5	9	CAT2	89-WRE-Q	1	256.99
6	9	CAT3	2238/QPD	2	77.90
7	9	CAT3	PVC23DRT	2	99.79
8	9	CAT3	WR3/TT3	1	359.85
9	9	CAT4	23109-HB	5	59.70
10	9	CAT4	SM-18277	1	20.97
11	10	CAT1	13-Q2/P2	2	104.93
12	10	CAT1	54778-2T	2	19.96
13	10	CAT2	2232/QTY	1	109.92
14	10	CAT2	89-WRE-Q	1	256.99
15	10	CAT3	PVC23DRT	2	99.79
16	10	CAT3	WR3/TT3	1	359.85
17	10	CAT4	23109-HB	3	39.80
18	10	CAT4	SM-18277	1	20.97

11. **List the top 5 vendors based on the total sales of their products. Show both the vendors' names and the total sales of their product. Sort by total sales.**

The code and output below display the top 5 vendors based on total sales of their products.

Query | Query History

```
1  CREATE TABLE q11_base AS (
2  SELECT p.p_code,
3         p.v_code,
4         SUM(d.day_subtotal) AS total_sales
5  FROM dwproduct p
6  LEFT JOIN ksdaysales_cus_reg d
7      ON p.p_code = d.p_code
8  GROUP BY p.p_code, p.v_code);
9
10 SELECT e.v_code,
11        v.v_name,
12        e.total_sales
13 FROM q11_base e
14 LEFT JOIN dwvendor v
15     ON e.v_code = v.v_code
16 WHERE e.total_sales IS NOT NULL
17 ORDER BY e.total_sales DESC
18 LIMIT 5;
19
```

Data Output | Messages | Notifications

	v_code integer	v_name character varying (35)	total_sales numeric
1	25595	Rubicon Systems	719.70
2	24288	ORDVA, Inc.	513.98
3	21344	Gomez Bros.	239.84
4	24288	ORDVA, Inc.	219.84
5	21225	Bryson, Inc.	199.58

12. **List the products that have not been sold in the year 2015. Show the product code, the product description and the product category.**

The code and output below display information about the products that have not been sold in the year 2015.

Query Query History

```
1  SELECT p.p_code,
2         p.p_category,
3         p.p_descript
4  FROM dwproduct p
5  WHERE p_code NOT IN (
6      SELECT d.p_code
7      FROM ksdaysales_cus_reg d
8      WHERE d.tm_id IN (
9      SELECT t.tm_id
10     FROM dwtime t
11     WHERE tm_year = 2015) )
```

Data Output Messages Notifications

	p_code [PK] character varying (10)	p_category character varying (5)	p_descript character varying (35)
1	11QER/31	CAT1	Power painter, 15 psi., 3-nozzle
2	14-Q1/L3	CAT1	9.00-in. pwr. saw blade
3	1558-QW1	CAT2	Hrd. cloth, 1/2-in., 3x50
4	2232/QWE	CAT3	B\&D jigsaw, 8-in. blade
5	23114-AA	CAT4	Sledge hammer, 12 lb.
6	SW-23116	CAT2	2.5-in. wd. screw, 50

13. **Find the top-selling products in each region based on the number of units sold. Show the region names, product descriptions, and total units sold. Order by region name and total units sold (from largest to smallest).**

The code and output below display the top-selling products in each region based on the number of units sold. In the Southeast, Southwest, and Northeast regions, the top-selling product is product PVC23DRT – the 3.5 inch by 8 foot PVC pipe. In the Northwest region, the top selling product is product WR3/TT3 – the steel matting.

Query | Query History

```
1  CREATE TABLE q13_base AS (
2  SELECT reg_name,
3         p_code,
4         SUM(sale_units) AS total_units_sold
5  FROM ksdaysales_cus_reg
6  GROUP BY reg_name, p_code);
7
8  WITH added_row_number AS (
9    SELECT
10     reg_name,
11     p_code,
12     total_units_sold,
13     ROW_NUMBER() OVER(PARTITION BY reg_name ORDER BY total_units_sold DESC) AS row_number
14   FROM q13_base
15 )
16 SELECT
17   a.*,
18   p.p_descript
19 FROM added_row_number a
20 LEFT JOIN dwproduct p
21     ON a.p_code = p.p_code
22 WHERE row_number = 1
23 ORDER BY total_units_sold DESC, reg_name;
24
```

Data Output | Messages | Notifications

	reg_name character varying (10)	p_code character varying (10)	total_units_sold bigint	row_number bigint	p_descript character varying (35)
1	SE	PVC23DRT	17	1	PVC pipe, 3.5-in., 8-ft
2	SW	PVC23DRT	12	1	PVC pipe, 3.5-in., 8-ft
3	NE	PVC23DRT	5	1	PVC pipe, 3.5-in., 8-ft
4	NW	WR3/TT3	3	1	Steel matting, 4'x8'x1/6", .5" mesh