

Database Design

Spring 2015

Data Warehousing

Points: 10 points

Submission Format: Online on D2L Dropbox as a PDF document

Objective: Upon successful completion of this lab, you should be able to create queries to access a dimensional database/data warehouse that answer interesting questions.

For each item in this lab, paste the SQL you used to complete the task. Ensure you follow the coding standards listed in the ETSU SQL standards guide found on D2L. Save this document as a PDF and name it *your_name_data_warehouse.pdf*

Note: This is worth 1 lab.

You must use the DB named STORE_DW to complete this lab. SALES_FACT is the fact table and PRODUCT, PROMOTION, STORE, and TIME are the dimension tables.

Explore the data first and see what the columns are. Look at some of the actual values. You may not understand all of the columns at first. Explore them and see if you can figure out what they might stand for. You may have to use multiple queries to answer the questions. If so, include all queries you used.

Use descriptive names for the headers of your result sets.

You must comment your queries. Create a header block of comments. See the guide mentioned above (D2L – Extra Materials).

Enter the SQL used to answer the question, the actual answer to your questions, and the result set returned as well as the number of rows returned.

Here is an example.

Do diet products sell better than regular products around Christmas? NO. This works because the data is all 4th quarter data. But you could filter further based on time of year.

```
SELECT TO_CHAR(sql_date, 'YY---MON') AS "Month",
```

```

diet_type,
-- the line below formats the money as a string with a $ and is left padded.
-- See the Oracle documentation for TO_CHAR()
TO_CHAR(SUM(unit_sales), 'L99G999D99MI') AS "Unit Sales"
FROM store_data_warehouse.sales_fact
JOIN store_data_warehouse.time
  USING (time_key)
JOIN store_data_warehouse.product
  USING (product_key)
WHERE diet_type != 'NA'
GROUP BY to_char(sql_date, 'YY---MON'), diet_type
ORDER BY diet_type, "Month";

```

Month	DIET_TYPE	Unit Sales
-----	-----	-----
94---DEC	Diet	\$28,434.00
94---NOV	Diet	\$26,659.00
94---OCT	Diet	\$26,463.00
95---DEC	Diet	\$27,076.00
95---NOV	Diet	\$26,483.00
95---OCT	Diet	\$26,846.00
94---DEC	Regular	\$51,131.00
94---NOV	Regular	\$49,197.00
94---OCT	Regular	\$51,965.00
95---DEC	Regular	\$50,802.00
95---NOV	Regular	\$51,186.00
95---OCT	Regular	\$51,259.00

12 rows selected

1. Create a list of products and their total sales. Sort by sales. Which product has the highest total unit sales for each year?
2. Using the answer from the query above, do these products sell better during a particular quarter of each year and is it the same quarter for each year?
3. Create a list of product categories and their total sales. Sort by sales. Does the product category for the highest total unit sales correspond to the answer in 1 above?

4. Which promotion has brought in the most unit sales?
5. Using the highest and lowest unit sales promotion, is there a pattern in the products between the promotions?
6. Create an interesting query by joining tables together, aggregating data in the fact table, and using the dimensions for drill downs and roll ups. You must use GROUP BY. This query must use 3 or more dimensions.

Write out the question your query answers, followed by the SQL used to answer the question, the actual answer to your questions, and the result set returned. This one must be your own work.