

# CS561 – SQL Programming Assignment

Due Date: 6/26/2020 (Fri.)

## Objectives

In this assignment, you will express “complex” OLAP queries in SQL. The key point of the exercise is to observe a large gap between the complexity of expressing the type of such queries and that of evaluating them (such as writing Java programs to produce the same results). Your mission (in addition to writing the SQL queries) is to consider the reasons for the gap (between the *expression* and *evaluation* of such queries) and how to narrow it.

## Description

Generate separate reports/output based on the following queries (one report for each of the queries):

- For each *customer*, compute the minimum and maximum sales quantities along with the corresponding products (purchased), dates (i.e., dates of those minimum and maximum sales quantities) and the states in which the sale transactions took place. If there are >1 occurrences of the min or max, display all.  
  
For the same *customer*, compute the average sales quantity.
- For each combination of *customer* and *product*, output the maximum sales quantities for NY and minimum sales quantities for NJ and CT in 3 separate columns. Like the first report, display the corresponding dates (i.e., dates of those maximum and minimum sales quantities). Furthermore, for CT and NJ, include only the sales that occurred after 2000; for NY, include all sales.
- Show for each product, the total sales quantities for 4 quarters, Q1, Q2, Q3 and Q4 (in four separate columns) – Q1 being the first 3 months of the year (Jan, Feb & Mar), Q2 the next 3 months (Apr, May & Jun), and so on – ignore the YEAR component of the dates (i.e., 6/15/2005 is considered the same date as 6/15/2002, etc.).
- For each product, show the product's average sales before and after each month (e.g., for February, show average sales of January and March). For “before” January and “after” December, display <NULL>. Again, ignore the YEAR component of the dates.
- For each product and for each month, count how many sales of the previous and how many sales of the following months had quantity greater than that month's average sale. Again, ignore the YEAR component of the dates.

The following are sample output reports – quantities displayed are for illustration only (not the actual values). **For dates (e.g., MAX DATE, MIN DATE), you can display ‘month’, ‘day’ and ‘year’ as 3 separate columns.**

### Report #1:

CUSTOMER	MIN_Q	MIN_PROD	MIN_DATE	ST	MAX_Q	MAX_PROD	MAX_DATE	ST	AVG_Q
=====	=====	=====	=====	==	=====	=====	=====	==	=====
Bloom	12	Pepsi	01/01/2006	NJ	2893	Apple	09/25/2001	NY	1435
Sam	1	Milk	02/15/2002	NJ	259	Banana	03/23/2004	CT	56
Emily	2	Bread	07/01/2005	NY	3087	Milk	02/02/2001	NJ	1512

### Report #2:

CUSTOMER	PRODUCT	NY_MAX	DATE	NJ_MIN	DATE	CT_MIN	DATE
=====	=====	=====	=====	=====	=====	=====	=====
Sam	Egg	1908	01/11/2001	2	07/24/2005	2	11/03/2008
Helen	Cookies	392	03/31/2002	42	09/14/2000	11	07/23/2002
Bloom	Butter	7045	09/22/2003	23	03/10/2004	8	09/11/2006

**Report #3:**

PRODUCT	Q1_TOTAL	Q2_TOTAL	Q3_TOTAL	Q4_TOTAL
=====	=====	=====	=====	=====
Pepsi	328923	243241	231873	42325
Milk	14239	9872	12142	2435

. . . .

**Report #4:**

PRODUCT	MONTH	BEFORE_AVG	AFTER_AVG
=====	=====	=====	=====
Bread	1	<NULL>	2434
Milk	3	254	325

. . . .

**Report #5:**

PRODUCT	MONTH	BEFORE_TOT	AFTER_TOT
=====	=====	=====	=====
Bread	4	23	<NULL>
Milk	2	45	35

. . . .

**Grading**

**NOTE: A query with syntax errors will lose 50% of the points for the query.**

**Submission**

Submit one single file of .TXT type containing all of the 5 queries.

Please include a "README" file if any special instructions are required.

You can discuss the "ideas" with your class mates or your friends, but the final queries must be your own work. If I determine that your queries are copies of someone else's, both you and that someone else will be disciplined (you will receive 0 for the entire assignment) and possibly receive additional penalties for the course.