



CS442 – Programming Assignment 2

Due Date: 5/5/2023 (Fri.)

Objectives:

- You will continue with evaluating simple report queries and produce the output. As with the assignment #1, you will also express the queries in SQL. The reports below are similar in nature with the reports from the assignment #1; however, there are two main differences between the two: (1) the new reports will require aggregation “outside” the groups (in assignment #1, all of the aggregates were computed for the rows within the groups); (2) some of the aggregates in the new reports will be computed based on other aggregates of the same reports – they are known as “dependent aggregates”.

Description:

- Generate reports based on the following queries:
 1. For each *product* and *month*, *count* the number of sales transactions that were between the *previous* and the *following* month's average sales quantities. For January and December, display <NULL> or 0.
 2. For *customer* and *product*, show the average sales *before*, *during* and *after* each quarter (e.g., for Q2, show average sales of Q1 and Q3. For “before” Q1 and “after” Q4, display <NULL>. The “YEAR” attribute is not considered for this query – for example, both Q1 of 2007 and Q1 of 2008 are considered Q1 regardless of the year.
 3. For each *customer*, *product* and *state* combination, compute (1) the product's average sale of this customer for the state (i.e., the simple AVG for the group-by attributes – this is the easy part), (2) the average sale of the product and the state but for *all of the other customers* and (3) the customer's average sale for the given state, but for *all of the other products*.
 4. For each product, find the median sales quantity (assume an odd number of sales for simplicity of presentation). (NOTE – “**median**” is defined as “*denoting or relating to a value or quantity lying at the midpoint of a frequency distribution of observed values or quantities, such that there is an equal probability of falling above or below it.*” E.g., Median value of the list {13, 23, 12, 16, 15, 9, 29} is 15.

For example, given the following sales transactions for Bread, the median quant for Bread is 3.

PRODUCT	QUANT
=====	=====
Bread	1
Bread	1
Bread	1
Bread	2
Bread	2
Bread	3
Bread	4
Bread	5
Bread	6
Bread	7
Bread	7

The following are sample report output (NOTE: the numbers shown below are not the actual aggregate values. You can write simple SQL queries to find the actual aggregate values).

You are only allowed to standard SQL syntax covered in class – do not use any other functions other than the 5 aggregate functions (sum, count, avg, max & min); and use only simple syntax of ‘agg(x)’ – i.e., do not use features such as CASE statement inside (such features hide implicit JOINS).



Report #1:

PRODUCT	MONTH	SALES_COUNT	BETWEEN_AVGS
====	=====	=====	=====
Cookies	1		<NULL>
Yogurt	3		19
. . . .			

Report #2:

CUSTOMER	PRODUCT	QTR	BEFORE_AVG	DURING_AVG	AFTER_AVG
=====	=====	=====	=====	=====	=====
Bloom	Coke	1	<NULL>	1539	2434
Sam	Eggs	3	254	539	325
. . . .					

Report #3:

CUSTOMER	PRODUCT	STATE	PROD_AVG	OTHER_CUST_AVG	OTHER_PROD_AVG
=====	=====	=====	=====	=====	=====
Helen	Bread	NY	243	268	1493
Emily	Milk	NJ	1426	478	926
. . . .					

Report #4:

PRODUCT	MEDIAN	QUANT
=====	=====	=====
Bread		422
Milk		1976
. . . .		

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

Submission: Submit all of the queries in a single TXT file – do NOT submit separate files for the queries or a ZIP file.

Please don't forget to include your name and CWID, and include a "README" section in the file if any special instructions are required.

I encourage you to discuss the "ideas" with your TAs (rather than your classmates, esp, if you have any specific questions), 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.