Department of Computer Science
1 Castle Point Terrace
Hoboken, NJ 07030
stevens.edu

CS561 – SQL Programming Assignment 1

Due Dates: 11/1/2022 (Tue) for Sec. A & 11/3/2022 (Thu) for Sec. B

Objectives

In this assignment, you will learn to *express* "complex" OLAP queries in SQL. The key point of the exercise is to observe the complexity of expressing the type of such queries despite relatively simple ideas of the queries themselves. Your mission (in addition to writing the SQL queries) is to consider the reasons for the complexity of the *expression* of these queries.

Description

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

1. For each *product* compute the *maximum* and *minimum* sales quantities along with the *corresponding customer* (who purchased the product), *dates* (i.e., dates of those maximum and minimum sales quantities) and the *state* in which the sale transaction took place.

   For the same *product*, compute the *average* sales quantity.

2. 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.

3. For *each of the 12 months* (regardless of the year), find the *most "popular"* and *least "popular" states* (those states with most and least total sales quantities) and the corresponding *total sales quantities (i.e., SUMs)*.

4. For each *customer*, find the "*most favorite*" *product* (which the customer purchased the most), the "*least favorite*" *product* (which the customer purchased the least), the "*most favorable state*" (where the most amounts of purchases were made) and the "*least favorable state*" (where the least amounts of purchases were made).

5. Show for each *customer* and *product* combination, 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., 3/11/2001 is considered the same date as 3/11/2002, etc.). Also compute the *average for the "whole" year* (again ignoring the YEAR component, meaning simply compute AVG) along with the *total quantities* (SUM) and the *counts* (COUNT).

The following are sample output reports – quantities displayed are for illustration only (not the actual values.

Report #1:

PRODUCT	MAX_Q	MAX_CUST	MAX_DATE	ST	MIN_Q	MIN_CUST	MIN_DATE	ST	AVG_Q
Pepsi	2893	Bloom	01/01/2006	NJ	12	Emily	09/25/2001	NY	1435
Banana	159	Dan	02/15/2002	NJ	1	Sam	03/23/2004	CT	56
Apple	3087	Helen	07/01/2005	NY	2	Sam	02/02/2001	NJ	1512
. . . .									

Report #2:

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

. . . .

Report #3:

MONTH	MOST_POPULAR_ST	MOST_POP_TOTAL_Q	LEAST_POPULAR_ST	LEAST_POP_TOTAL_Q
1	NJ	497214	CT	55526
2	NJ	1874794	NY	23126
3	CT	974531	NJ	19958

. . . .

Report #4:

CUSTOMER	MOST_FAV_PROD	LEAST_FAV_PROD	MOST_FAV_ST	LEAST_FAV_ST
Claire	Apple	Jellies	NJ	NY
Emily	Eggs	Butter	NJ	CT
Chae	Fish	Ham	PA	NJ

. . . .

Report #5:

CUSTOMER	PRODUCT	Q1_TOT	Q2_TOT	Q3_TOT	Q4_TOT	AVERAGE	TOTAL	COUNT
Sam	Pepsi	1923	4241	2383	1325	759	9872	13
Emily	Milk	239	9872	142	2435	1586	12688	8
Helen	Bread	2534	981	4239	1987	1082	9741	9

. . . .

Grading

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

Submission

Submit **one file** containing all of the 5 queries with your name and CWID on it on Canvas. The file type must be “TXT”.

Please include a “README” section in the same file if any special instructions are required.

You can discuss the “ideas” with your classmates 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.