

## Assignment Day3 –SQL: Comprehensive practice (Tommy Park)

### Answer following questions

1. In SQL Server, assuming you can find the result by using both joins and subqueries, which one would you prefer to use and why?

I prefer to use subqueries. While joins can only be used in FROM, subqueries can be used in SELECT, WHERE, HAVING, FROM, ORDER BY. Despite joins having better performance than subqueries, subqueries are easier to understand.

2. What is CTE and when to use it?

CTE stands for common table expression. Often used for temporary result set to reference. One of the common methods is to use CTE recursively to call itself or queries repetitively.

3. What are Table Variables? What is their scope and where are they created in SQL Server?

Table variables are variables that are stored in temp. The scope of table variables is batch is that is declared.

4. What is the difference between DELETE and TRUNCATE? Which one will have better performance and why?

DELETE can be used to delete specific rows with usage of WHERE clause. TRUNCATE will delete all rows and cannot use WHERE clause. Overall, TRUNCATE has better performance as it deletes faster as it requires less transaction log space than DELETE.

5. What is Identity column? How does DELETE and TRUNCATE affect it?

Identity column refers to columns or fields in a database that are made up of values generated by the database, like auto-generated key ID. For identity columns, TRUNCATE resets the sequence but DELETE does not. Therefore, Rollback can be used for DELETE while TRUNCATE cannot.

6. What is the difference between “delete from table\_name” and “truncate table table\_name”?

DELETE will remove specific rows or data so the table will maintain integrity. However, TRUNCATE removes all the data which does not maintain the integrity of the table.

### Write queries for following scenarios (Uploaded in Github)

All scenarios are based on Database NORTHWND.

1. List all cities that have both Employees and Customers.
2. List all cities that have Customers but no Employee.
  - a. Use sub-query
  - b. Do not use sub-query
3. List all products and their total order quantities throughout all orders.
4. List all Customer Cities and total products ordered by that city.

5. List all Customer Cities that have at least two customers.
  - a. Use union
  - b. Use sub-query and no union
6. List all Customer Cities that have ordered at least two different kinds of products.
7. List all Customers who have ordered products, but have the 'ship city' on the order different from their own customer cities.
8. List 5 most popular products, their average price, and the customer city that ordered most quantity of it.
9. List all cities that have never ordered something but we have employees there.
  - a. Use sub-query
  - b. Do not use sub-query
10. List one city, if exists, that is the city from where the employee sold most orders (not the product quantity) is, and also the city of most total quantity of products ordered from. (tip: join sub-query)
11. How do you remove the duplicates record of a table? (Answered in Github as short answer)
12. Sample table to be used for solutions below- Employee (empid integer, mgrid integer, deptid integer, salary money) Dept (deptid integer, deptname varchar(20))  
  
Find employees who do not manage anybody.
13. Find departments that have maximum number of employees. (solution should consider scenario having more than 1 departments that have maximum number of employees). Result should only have - deptname, count of employees sorted by deptname.
14. Find top 3 employees (salary based) in every department. Result should have deptname, empid, salary sorted by deptname and then employee with high to low salary.

GOOD LUCK.