Conestoga College Doon Campus

PROG8650 - Section 2

Group Project 3 (Group 5)

Kanishka Kaushalya Makewita Appuhamilage (8803608)

Jayraj Anil Sali (8774136)

Sharmini Sathis Kumar Pillai (8808970)

Anju Arun (8811685)

Jimmy Jose (8811979)

Mohammed Yaqoob Atif (8720021)

Naveen Yaragarla (8796043)

Section B – Q3 – Email Response

To: Matt Kozi (Solutions Architect)

Subject: Benefits of using stored procedure in script

Hello Matt Kozi (Solutions Architect),

The main role of DBA in any database management is to maintain the database updates with more readability, reusability and reduce the redundant SQL code while automating the task. Changing the table names or changing the column datatypes is one of the most common tasks to be performed as the new features coming in or modifying the existing requirements. We can perform those tasks in more conventional way with redundant code, but the best alternative is to automate those tasks with the help of stored procedures will be best choice for maintainability of any database. The following two paragraphs explains the benefit of using stored procedures and the last paragraph will explain how it is being implemented in our project for Salary field in Employee table.

There are ample of benefit of having stored procedure in our SQL code. Stored procedures are a potential tool for accelerating code and improving performance. The major benefits include the ability to automate tasks requiring numerous SQL statements, the ability to be readily updated, and the ability to be reused. In a relational database management system, a stored procedure is a collection of SQL statements and multiple PL/SQL expressions that can be saved and used repeatedly. The most SQL queries like SELECT, INSERT, UPDATE, or DELETE can make up a stored procedure. These sentences operate collectively to address a certain issue or carry out a group of connected actions. (TechStrongGroupINC, 2022, p. 1)

We believe that in some circumstances, dynamic SQL is unavoidably required, but whenever possible, parameterized SQL or stored procedures should be beneficial. Since stored procedures are typically precompiled, you may be sure that the statements are correct when they are written in the database and avoid SQL syntax issues. Knowing these things now should prevent future hassles. It may possible sometimes that we cannot be able to debug stored procedures in the same environment as we do the rest of your development, but I believe that by using stored procedures, you would be less likely to need to debug the SQL. Additionally, it may be unit-tested if you used saved procedures. (CodingHorror, 2022, p. 1)

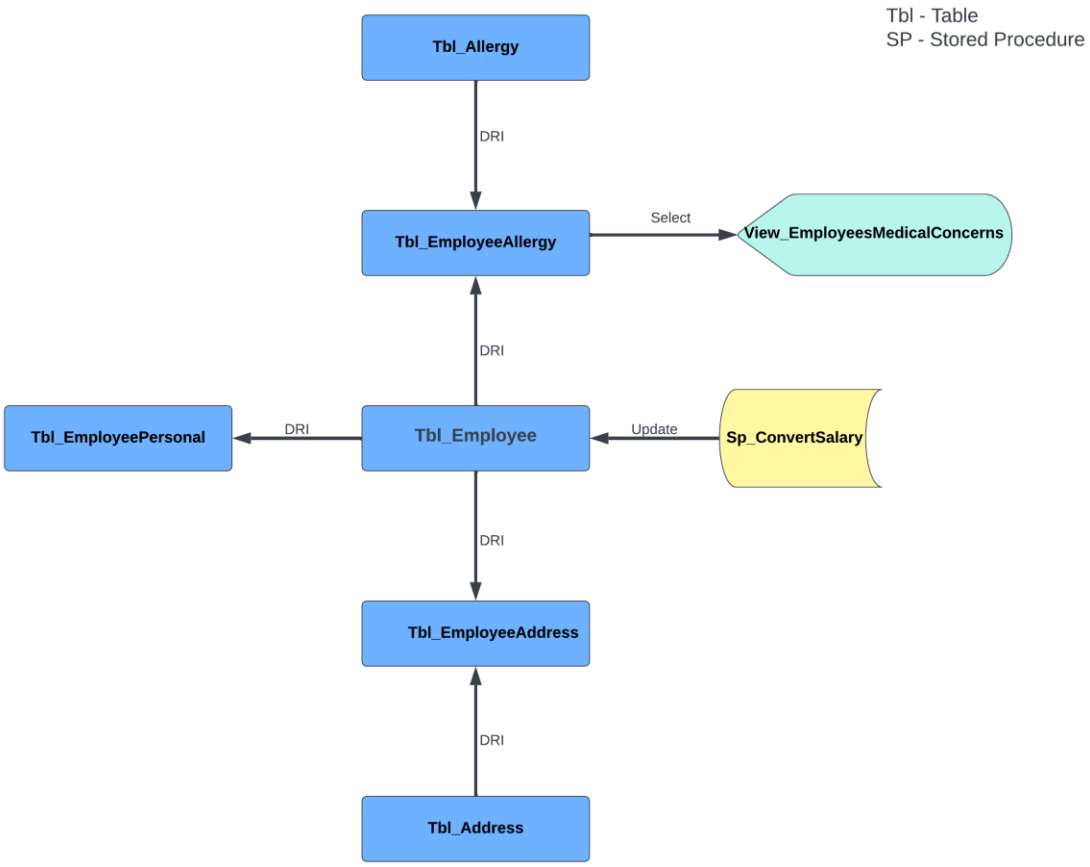
For example, in our database scenario, we required to update the salary field of Employee table. There are two possible ways by which we can make this possible. The first way of doing this is conventional way through which we can apply ALTER query on selected column and will update the table. But this not feasible way to do so if you are working on large and complex database. However, in organisation, most of the time we are working on complex database then it might possible that we require to update existing version of multiple entity. So, in these cases stored procedures comes into the picture and it makes our task easier. Here we used loop statement where created logic that will take one row from Employee table at a time and will convert the Salary field datatype from INT to VARCHAR. Meanwhile, loop also executes some statements like it will add commas to number, insert dollar sign before number and add two decimal places to number. And the loop execute until the counter value exceeds the total number of records. Also, after each iteration record has been updated in the EMPLYOEE table. This is how it will redundant our task.

Overall, by implementing the stored procedures way will be best choice for database maintainability and automate these tasks in a better way. Please go through our proposal to partially automate task of cleansing the data types in any entities in the tables and provide your valuable feedback whether it will be better solution to implement in our scenario.

Section C:

Q1) Create a dependency graph that shows dependencies among the updated set of tables, views and stored procedures.

Ans)



Q2) Explain why you need to develop dependency graphs which include views and database objects such as stored procedures when we already have logical and physical data models. How do dependency graphs and data models differentiate?

Ans) According to Pablo Azero of Jalasoft, "A dependency graph is a graph that represents dependencies between objects of some application domain." In an application, this tool will help us to map out the relationship between different components. It helps to figure out both the connections and directions of the dependencies to map which components rely on each other.

Dependency Graphs	Data Models
A graph that represents dependencies between objects of some application domain.	Define how the logical structure of a database is modeled
It's a tool that maps out relationships between the different components of an application	They are fundamental entities to introduce abstraction in a DBMS.
It shows both the connections and directions of the dependencies to help us to visualize which components depend on each other.	They define how data is connected to each other and how they are processed and stored inside the system.

Q3) Through your recent work, you implemented a new stored procedure, while stored procedures can be helpful, they can also present challenges to data stability. Assume the following scenario is true; A stored procedure that exports health data from the EmployeePersonal table was running, and another stored procedure that reformats emergency contact phone data could also be running.

a) Assume we just executed the stored procedure you have created for this assessment while the other two stored procedures were also running, and an error occurred. Give an example of a dirty read, a nonrepeatable read, and a phantom read among this group of stored procedures.

Ans) Dirty Read: When one transaction can view the uncommitted statistics of some other, the Dirty Read Concurrency Problem takes place. That is, information has been modified through every other transaction, but it has no longer but been committed or rolled lower back.

- It could now not be a problem in most cases. Though, if the primary transaction is rolled returned after the second reads the uncommitted facts, the second one transaction could have dirty statistics that does not exist within the database anymore.

- For example, if the second transaction accesses a row in the EmployeePersonal table recorded by the first transaction, it will not be committed to the database. Transaction 1 has non-existent or dirty data that is no longer available after the first transaction was rolled back.

- **Non-repeatable read:** When one transaction reads the same data twice and another transaction updates that data in between the first and second reads of the first transaction, this is known as the Non-Repeatable Read Concurrency Problem.

- For example, in the EmployeePersonal table, consider a transaction that reads a row before beginning with the third process. After the third step, when the same transaction occurs, the modified row will be read which would be different from the initial read.

- **Phantom read:** When one transaction executes a query twice, it gets a different number of rows in the result set each time. This is known as the Phantom Read Concurrency Problem. This typically occurs when a second transaction inserts additional rows between the first and second query executions of the first transaction that match the WHERE clause of the first transaction's query.

- For instance, when a row is reread by a transaction in the EmployeePersonal table after the stored procedure that records health data (AllergyName) has been committed, the transaction will read additional rows because the stored procedure records new data.

b) What concurrency control measures are appropriate for the stored procedure that you are creating?

Ans) The stored procedure that is being created in our example is to convert the Salary column to integer and add the commas, decimals and dollar sign to the data that is being already present. To make this stored procedure works without any issues are deadlocks, it would be better to use the following techniques to avoid it.

Two-Phase-Locking Protocol is most appropriate as we are reading the data from the Employee table and updating each data row in the table to match with stored procedure requirement and once, we have the updated data then we are updating the existing value with new modified value.

The main goal of Two-Phase-Locking Protocol is used to gain the ownership of shared resources without the possibility of creating the deadlock (geeksforgeeks - lemilxavier, 07 July, 2022, p-3). So, with the help of this, the stored procedure that is being created will avoid the deadlock scenario. This control method has three phases like below

- Lock Acquisition
- Modification of data
- Release lock

During first phase of this stored procedure, we can lock the employee resource, soon after that we can modify the employee table data where it is appropriate and once, we have the updated data ready then we can release the lock. This way we never encounter the deadlock for the employee table.

c) What concurrency control measures are appropriate for the two other stored procedures?

Ans) Briefly, the one stored procedure is taking care of exporting the health data from EmployeePersonal table and at the same time another stored procedure formatting the emergency contact phone data. In this scenario the EmployeePersonal table is being used, if we are not properly taking care of the scenario the deadlock scenario might occur because the same resource is being used by two different stored procedures.

To avoid such case, Validation Concurrency Protocol will be better concurrency control method to avoid such deadlock scenarios. The way the Validation Concurrency Protocol works in three phases.

- Read Phase
- Validation Phase
- Write Phase

During read phase, the transaction reads the database, executes the needed computations and makes the updates to a private copy of the database values. All update operations of the transactions are recorded in a temporary update file, which is not accessed by the remaining transactions.

During the validation phase, the transaction is validated to ensure that the changes made will not affect the integrity and consistency of the database. If the validation test is positive, the transaction goes to a write phase. If the validation test is negative, the transaction is restarted, and the changes are discarded.

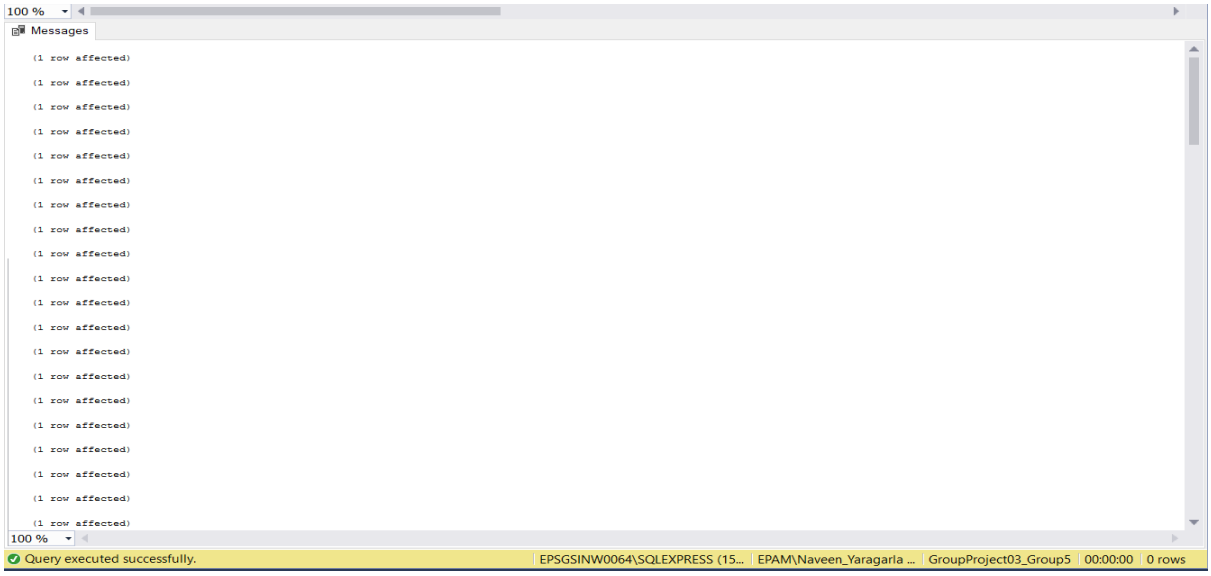
During the write phase, the changes are permanently applied to the database. (Geeksforgeeks - lemilxavier, 07 July, 2022, p-3)

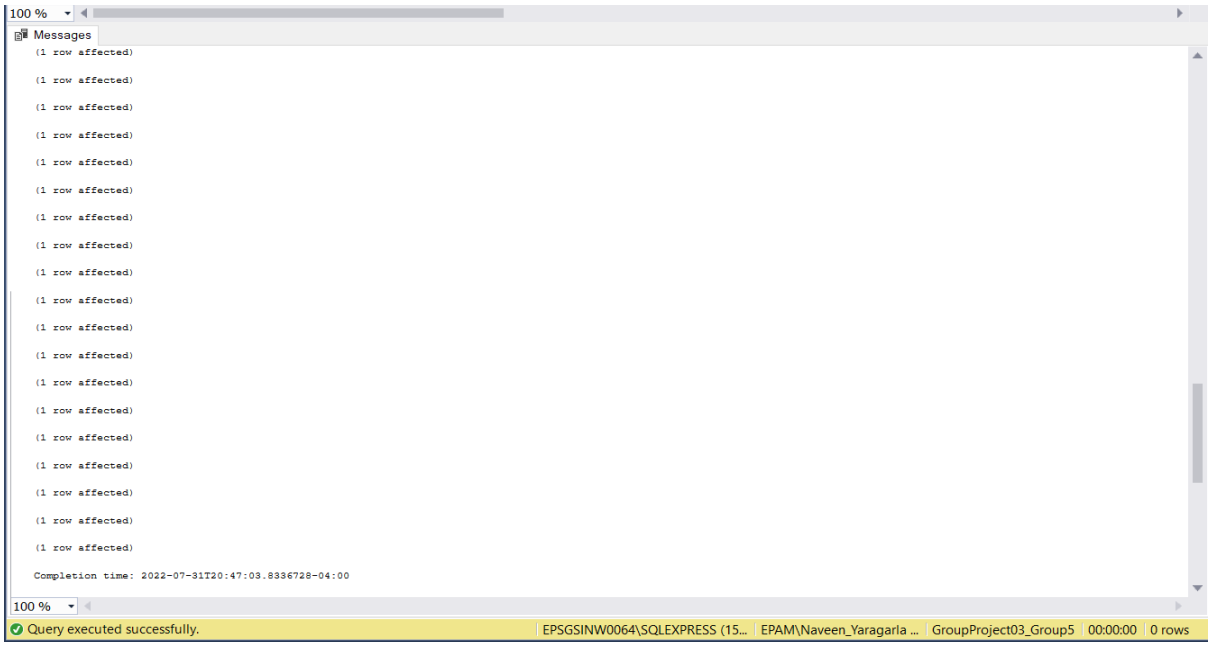
Overall, there were other techniques also be used, but with our scenario the above control mechanism will be a good option to avoid the deadlock scenario.

Results:

SQL Output Screenshots:

Section A, Create and Insert with modified tables:





Section B – Concatenated Last Name and First Name Function:

Results Messages

	EmployeeID	RoleID	ManagerID	DepartmentID	FName	Lname	Salary	HireDate	Probation	PersonaleMail	PersonalPhoneNumber	ConcatedName
1	1117	1	1219	10	Darrel	Berry	1000	2022-07-02	6 Months	darrel.berry@one.com	9876543345	Berry, Darrel
2	1134	2	1236	20	Erma	Lucas	1300	2022-06-02	6 Months	erma.lucas@one.com	9812377345	Lucas, Erma
3	1151	3	1253	30	Jose	Benson	1500	2021-01-02	6 Months	jose.benson@one.com	9111377345	Benson, Jose
4	1168	4	1270	40	Angelina	Estrada	1100	2019-10-11	6 Months	angelina.estr@one.com	9111675432	Estrada, Angelina
5	1185	5	1287	50	Ebony	Gibson	1000	2019-10-11	6 Months	ebony.gibby@one.com	910098732	Gibson, Ebony
6	1202	6	1304	60	Jean	Franklin	1000	2010-12-22	6 Months	jean.frank@one.com	910098002	Franklin, Jean
7	1219	6	1304	60	Jenny	Frank	1000	2010-12-22	6 Months	jenny.f@one.com	919998002	Frank, Jenny
8	1236	6	1304	60	Jenny	Frank	80000000	2010-12-22	6 Months	jenny.f@one.com	919998002	Frank, Jenny

EmployeesMedicalConcerns View output:

100 %

Results Messages

	EmpName	AllergyName
1	Berry, Darrel	Peanut Allergy
2	Berry, Darrel	Drug Allergy
3	Berry, Darrel	Pet Allergy
4	Lucas, Erma	Insect Allergy
5	Lucas, Erma	Latex Allergy

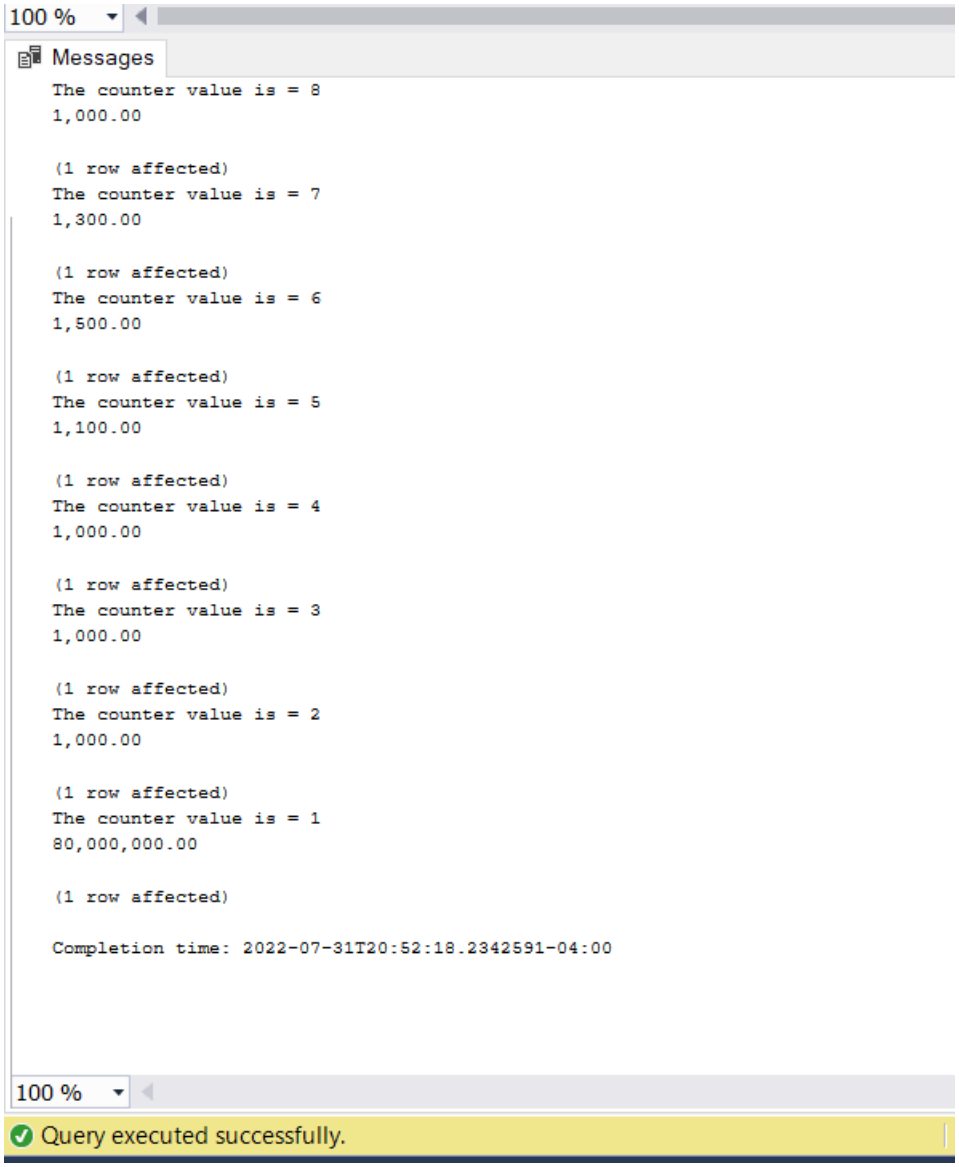
Query executed successfully.

convertSalary Stored Procedure Output:

Results Messages

	EmployeeID	RoleID	ManagerID	DepartmentID	FName	Lname	Salary	HireDate	Probation	PersonaleMail	PersonalPhoneNumber
1	1117	1	1219	10	Darrel	Berry	$1,000.00	2022-07-02	6 Months	darrel.berry@one.com	9876543345
2	1134	2	1236	20	Erma	Lucas	$1,300.00	2022-06-02	6 Months	erma.lucas@one.com	9812377345
3	1151	3	1253	30	Jose	Benson	$1,500.00	2021-01-02	6 Months	jose.benson@one.com	9111377345
4	1168	4	1270	40	Angelina	Estrada	$1,100.00	2019-10-11	6 Months	angelina.estr@one.com	9111675432
5	1185	5	1287	50	Ebony	Gibson	$1,000.00	2019-10-11	6 Months	ebony.gibby@one.com	910098732
6	1202	6	1304	60	Jean	Franklin	$1,000.00	2010-12-22	6 Months	jean.frank@one.com	910098002
7	1219	6	1304	60	Jenny	Frank	$1,000.00	2010-12-22	6 Months	jenny.f@one.com	919998002
8	1236	6	1304	60	Jenny	Frank	$80,000,000.00	2010-12-22	6 Months	jenny.f@one.com	919998002

Query executed successfully. EPSGSINW0064\SQLEXPRESS (15... EPAM\Naveen_Yaragarla ... GroupProject03_Group5



REFERENCES:

- (TechStrongGroupINC, 2022, p. 1) (The power and potential of Stored Procedure, n.d.) **RETRIEVD FROM** https://devops.com/the-power-and-potential-of-stored-procedures/
- (CodingHorror, 2022, p. 1) (Who needs Stored procedres, Anyway?, n.d.) **RETRIVED FROM** https://blog.codinghorror.com/who-needs-stored-procedures-anyways/
- The Essential Guide to Dependency Graphs. (2021, August 29). Argon Security - Holistic Security for Your CI/CD Pipeline. https://www.argon.io/blog/the-essential-guide-to-dependency-graphs/#:~:text=Dependency%20graphs%20are%20essential%20for
- DBMS - Data Models - Tutorialspoint. (2019). Tutorialspoint.com. https://www.tutorialspoint.com/dbms/dbms_data_models.htm
- *Dirty read in SQL*. (2018, September 26). GeeksforGeeks. https://www.geeksforgeeks.org/dbms-dirty-read-in-sql/
- David-Engel. (n.d.). Transaction Isolation Levels (ODBC) - ODBC API Reference. Docs.microsoft.com. Retrieved July 30, 2022, from https://docs.microsoft.com/en-us/sql/odbc/reference/develop-app/transaction-isolation-levels?view=sql-server-ver16
- Progress Documentation. (n.d.). Docs.progress.com. Retrieved July 30, 2022, from https://docs.progress.com/bundle/openedge-crystal-reports-117/page/Phantom-read.html

- DBMS Concurrency Control - javatpoint. (n.d.). Www.javatpoint.com. https://www.javatpoint.com/dbms-concurrency-control
- lemilxavier, Concurrency Control Techniques – geeksforgeeks – July 07th , 2022 (https://www.geeksforgeeks.org/concurrency-control-techniques/)