DML stands for Data Manipulation Language and up till now we have introduced only one statement which pertains to this subset of SQL – the SELECT statement. As you should have understood by now the SELECT statement focuses on data retrieval, but there are other statements within DML which perform other important tasks. Other statements that pertain to the DML subset are: INSERT, UPDATE, DELETE, TRUNCATE and MERGE.

Adding Rows of data

In Microsoft SQL Server there are several methods that can be adopted to be able to insert data in an entity. Amongst the available T-SQL statements that we can use to insert data in an entity, one can find; INSERT VALUES, INSERT SELECT, INSERT EXEC, SELECT INTO and BULK INSERT.

INSERT VALUES statement

This statement is used to insert data in an entity which is based on specified values. This statement can be used to insert either a single row or multiple rows of data in a particular entity. Consider the below entity; we can use INSERT VALUES such that a new product is included in the entity. The new product will be Peaches, it will cost 0.70 and it will be supplied by company 10.

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10
3	Banana	0.60	20

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10
3	Banana	0.60	20
4	Peaches	0.70	10

In order to add row/s to a particular entity using the INSERT INTO statement the below syntax needs to be used:

```
INSERT INTO table [(column [, column . . . ])]
VALUES (value [, value . . .])
        [,(value [, value . . ]) . . . ];
```

What do the keywords in this INSERT VALUES syntax statement mean?

- **INSERT INTO table**– chooses the entity in which the data is to be inserted
- **[(column [, column . . .])]** – the names of the columns where the values need to be inserted. This is optional and can be left out (also known as target column names).
- **VALUES** – keyword which indicates that from this point onwards, the actual data to be inserted is to be listed
- **(value [, value . . .])** – the value keyword needs to be replaced by an actual value

NOTE:

- when inserting values which have a date or character/string data type these need to be enclosed in single quotes
- also remember that when inserting dates the format that is to be followed should be **yyyymmdd.**
- when inserting data for columns which have a numeric data type, you just need to include the actual numerical value – do not include any single quotes
- every new row to be inserted must include a value for *each* mandatory (NOT NULL) column in the entity

INSERT VALUES with target column names specified

Although this option is not required, it is highly recommended that you use it as you will have control on the value-column relationships within the table. The order in which the target column names are written is not important as long as the equivalent values follow the order specified after the INSERT INTO keyword.

```
INSERT INTO countries (country_id, country_name, region_id)
VALUES ('ML', 'Malta', 1);
```

Figure 1 - example of a typical INSERT INTO statement

Figure 1 includes an example an INSERT INTO statement. The orange box include the columns that be inserted with data (OPTIONAL). The green box includes the actual data that will be inserted in the countries table.

```
INSERT INTO countries (country_name, country_id, region_id)
VALUES ('Malta','ML', 1);
```

Figure 2 - example of another INSERT INTO statement

Figure 2 will add the same data in the countries table, but it is written differently. Notice that the country_name and country_id have changed position. For this reason it was also important to change the position for the values 'Malta' and 'ML'.

```
INSERT INTO countries (country_id, country_name, region_id)
VALUES ('Malta','ML', 1);
```

Figure 3 - erroneous INSERT INTO statement

NOTE: The statement in Figure 3, will execute normally without errors but the data inserted in the entity will be incorrect. In this case the value 'Malta' would be incorrectly placed in the country_id column and 'ML' in country_name. **Make sure that the value-column relationship is correct**

Example 1: Write a query that will insert a new row of data in the countries table. The new row should have the following data: country_id: ML, country_name: Malta, region_id: 1. The figure below shows that the row was inserted successfully as the number of rows increased by 2.

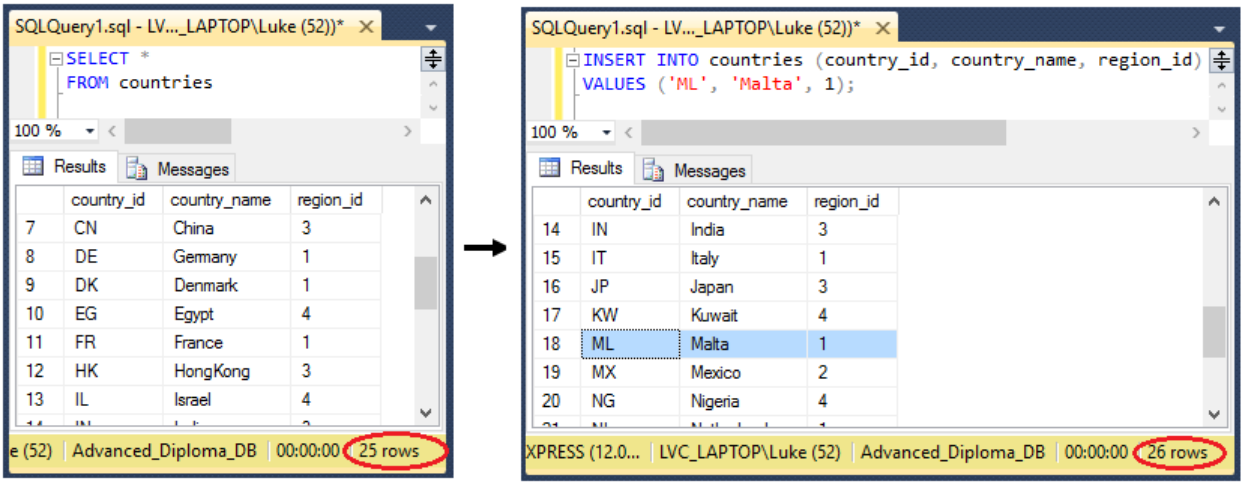


Figure 4 – Result of example 1: adding a new row using target column names

Example 2: Write a query that will input the following information in the employees table: employee_id: 207, first_name: Charles, last_name: Brincat, email: BRINCS, phone_number: 515.123.5555, hire_date: 10th March 2000, job_id: IT_PROG, salary: 4800, commission_pct: null, manager_id:103, department_id: 60

```
INSERT INTO employees (employee_id, first_name, last_name, email,
                       phone_number, hire_date, job_id, salary,
                       commission_pct, manager_id, department_id)
VALUES (207, 'Charles', 'Brincat', 'BRINCS', '515.123.5555',
        '20000310','IT_PROG',4800,null,103,60);
```

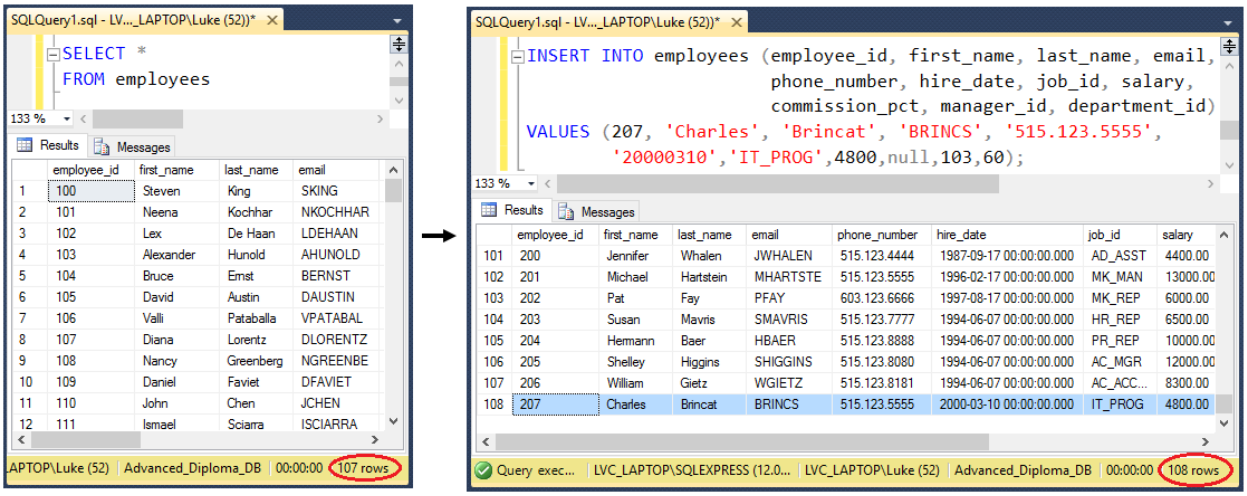


Figure 5 -Result of example2: adding a new employee

INSERT VALUES without target column names specified

While inserting data in the database, it is possible to do without target names soon after the INSERT INTO table. As mentioned in the previous section, target columns are optional and for this reason they can be left out.

If this option is to be used when inserting data it is of ***utmost importance that the order of the columns is equivalent to that used while creating the entity***. To determine the order of the columns used on creation there are two methods:

1. Open the Columns folder of the entity in Object Explorer
2. Use the *sp_columns* stored procedure

From example 3 and 4 in the next page, you can easily determine the order in which the columns have been created (department_id, department_name, manager_id, location_id). Also note that the sp_columns stored procedure includes a number of columns which give us a lot of information about each column in the table.

Example 3: Determine the order of the columns in the departments table using Object Explorer

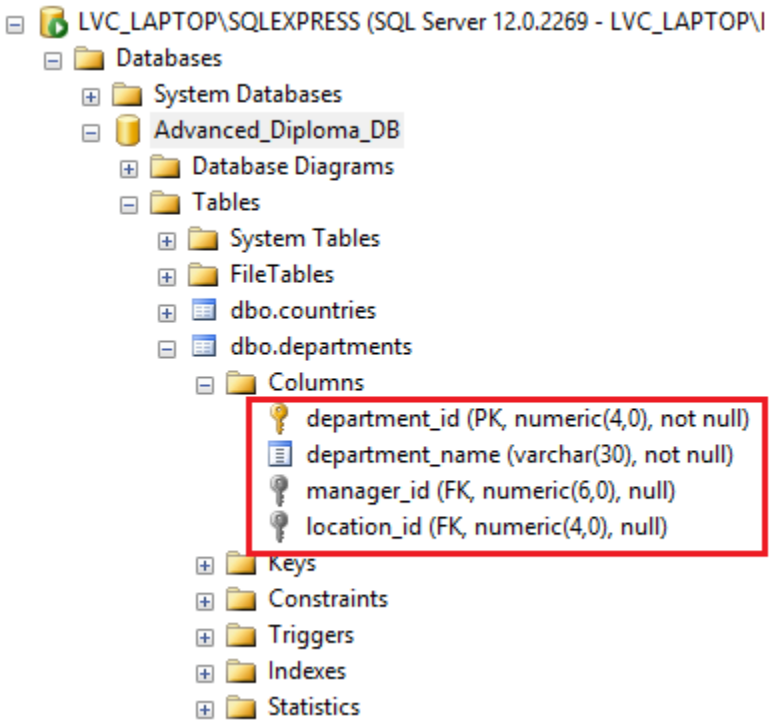


Figure 6 - Screenshot showing the rows in the departments table (order they were created)

Example 4: Determine the order of the columns in the departments table using the *sp_columns* stored procedure

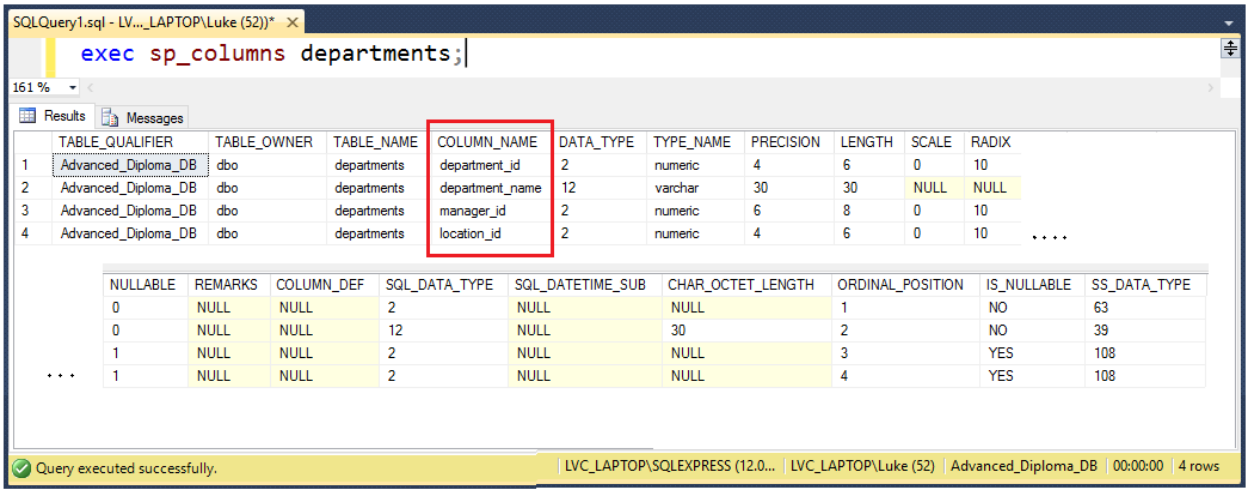


Figure 7 - Screenshot showing all the information related to the columns in the departments table.

Once that the order of the columns in a table is known, INSERT INTO statements can be written without the target columns. When using this method make sure that the values inserted correspond to the order of the columns.

Example 5: Write a query that will insert a new row of data in the countries table. The new row should have the following data: country_id: SP, country_name: Spain, region_id: 1. You are not to include the target columns in the INSERT INTO statement.

Given the fact that the target columns are not to be included the order of the columns needs to be determined. We used the Object explorer in this case. The order of the columns was country_id, country_name, region_id.

Figure 8 - Order of the columns found in the countries table

Below is the answer after inserting the desired row in the countries table. You are to note that the values follow the order of the columns as shown in figure 8. Incorrectly order of values might result in data placed in incorrect columns or data type errors.

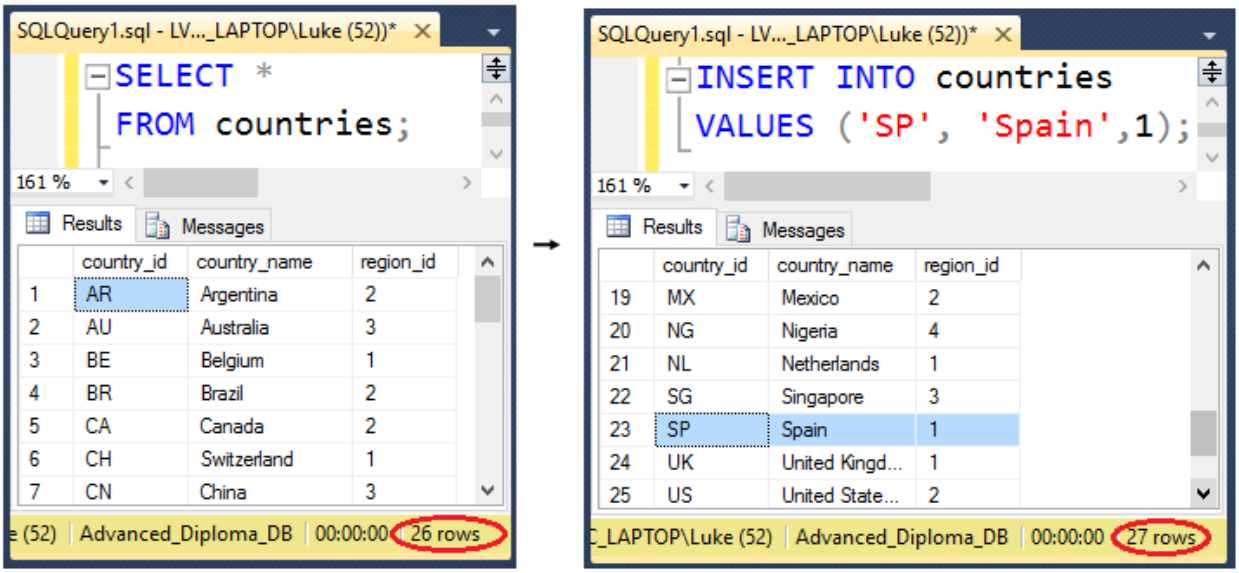


Figure 9 - Result of example 5: adding a new row without target columns being specified

INSERT VALUES for optional columns (which allow NULL values)

As you should have learnt during the first semester, a column/attribute in an entity can be either optional or mandatory. From figure 10, you can easily determine that in the Department entity, the department_id and department_name columns are mandatory while the location_id and manager_id columns are optional.

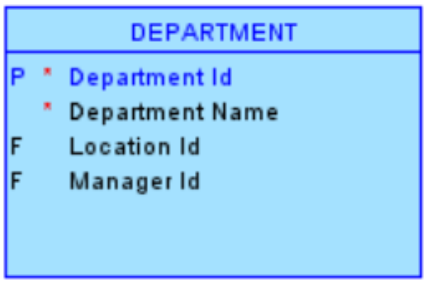


Figure 10 – representation of the Department entity using Information Engineering Notation

While creating the entity in the database, all mandatory attributes should be assigned a NOT NULL constraint. In simple words, a mandatory column should have a value for each row that exists in the entity and an optional column can be left empty for any row in the entity.

Given the fact that you will not always have access to the ERD or the CREATE TABLE code to determine if a column is mandatory or optional, you can use the two methods which were mentioned on page 4. As can be seen in figures 11 and 12, the mandatory fields are (department_id and department_name) and the optional fields are (manager_id and location_id). Note that when the *sp_columns* stored procedure is used to determine of a column is mandatory or not the NULLABLE column needs to be considered. If the value in this column is 0 then the column in mandatory and if the value is 1 then the column is optional (can be left empty)

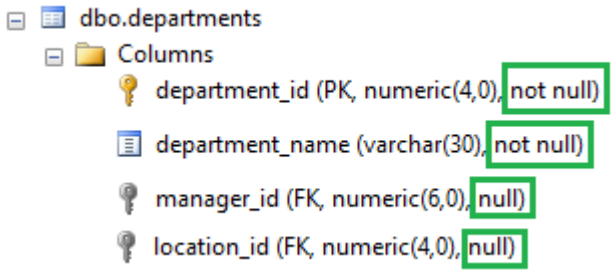


Figure 11 - mandatory or optional columns in the department table (using Object Explorer)

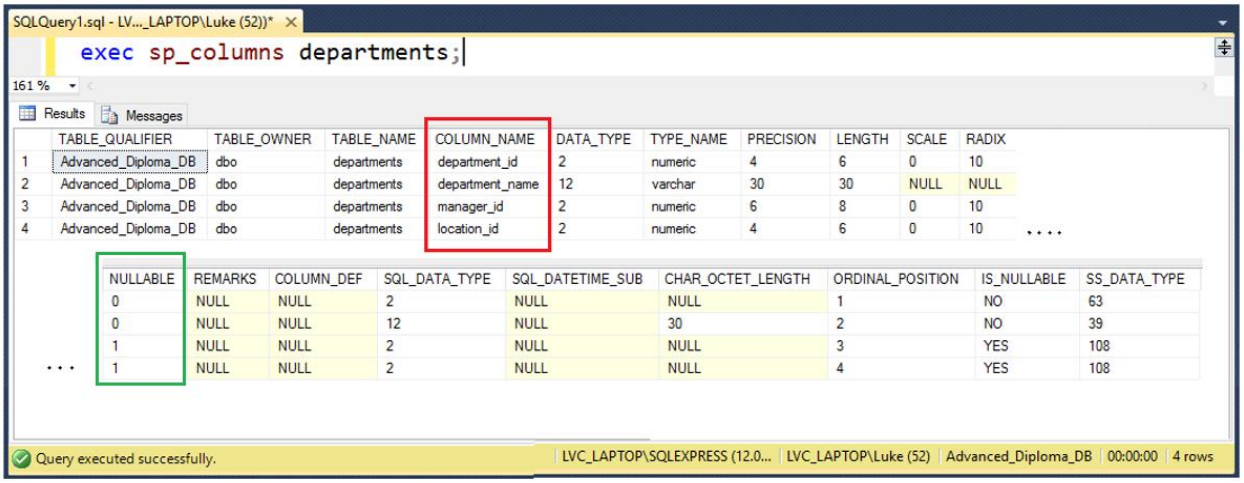


Figure 12 - mandatory or optional columns in department table (using sp_columns)

Once that you manage to determine which columns are optional, you can use either of these methods to add rows of data:

1. *Implicit Method*: remove the optional columns from the target columns and do not include a value for these columns
2. *Explicit Method*: the null keyword is to be specifically included in the values list

Example 6: Write a query that adds a new department (department_id = 31, department_name = Purchasing 2) to the departments table. Note that the manager_id and location_id should be left empty and the implicit method should be used.

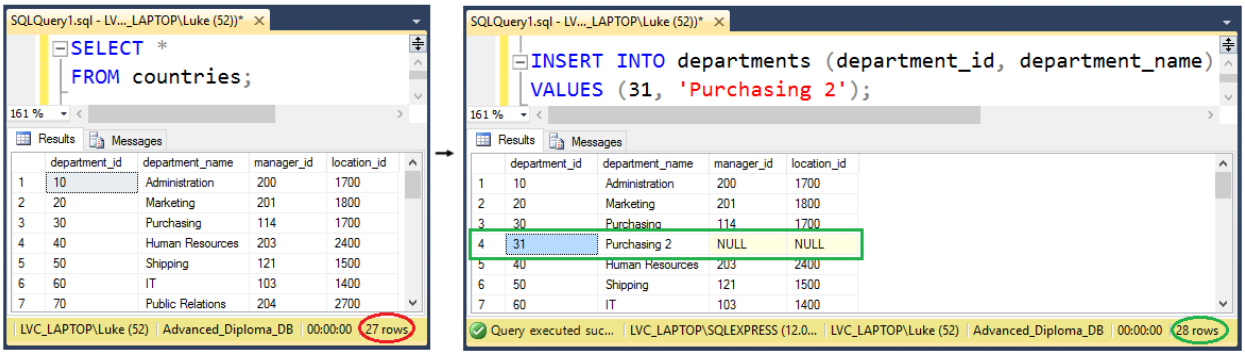


Figure 13 - Result of example 6: Implicit method to add new row

NOTE: since we have used the *Implicit method*, the manager_id and location_id are not included in the target columns as they are optional columns

Example 7: Write a query that adds a new department (department_id = 32, department_name = Purchasing 3) to the departments table. Note that the manager_id and location_id should be left empty and the explicit method should be used.

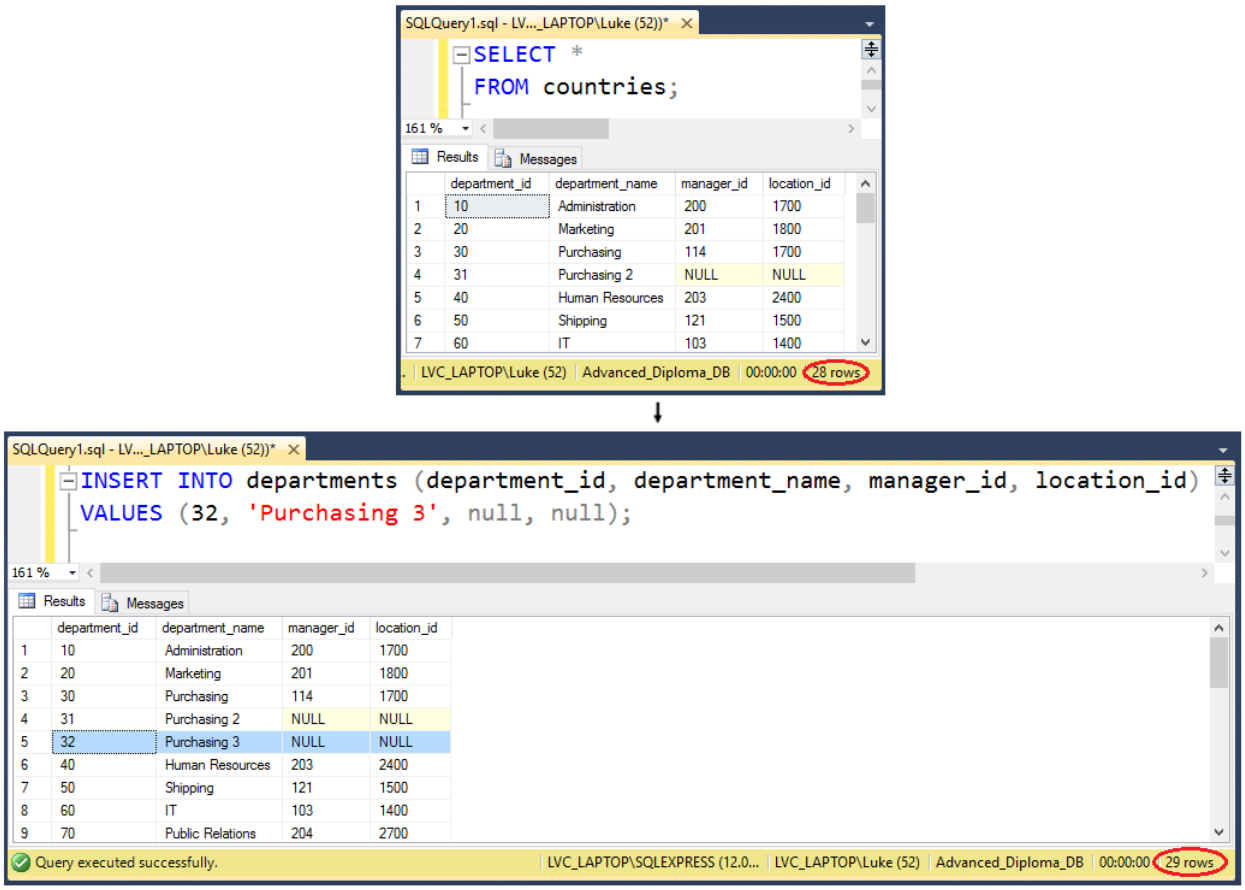


Figure 14 - Result of example 7: Explicit method to add a new row

NOTE: since we have used the *Explicit method*, in the VALUES section we had to include two consecutive 'null' one for manager_id and the other for location_id. Also note that the INSERT statement could be changed to the one in figure 15 (remember that the target columns are optional and can be left out, provided that the actual values are in the column order which was used during table creation).

```
INSERT INTO departments
VALUES (32, 'Purchasing 3', null, null);
```

Figure 15 - The equivalent of the code in the INSERT VALUES used in figure 14

INSERT VALUES and multiple rows

The INSERT VALUES statement can be used to add more than one row at a time using just a single statement. In simple words this means that with one statement you can add multiple rows in the table. Each row of data in the statement is to be separated using a comma.

A very important point to know is that even though multiple rows can be added the statement is still considered as **one** transaction. This means that if a row in the statement fails to be added to the table, the entire statement will fail.

Example 8: Write a query that will add three job titles (Information Technology Officer, Systems Analyst and System Tester) to the jobs table.

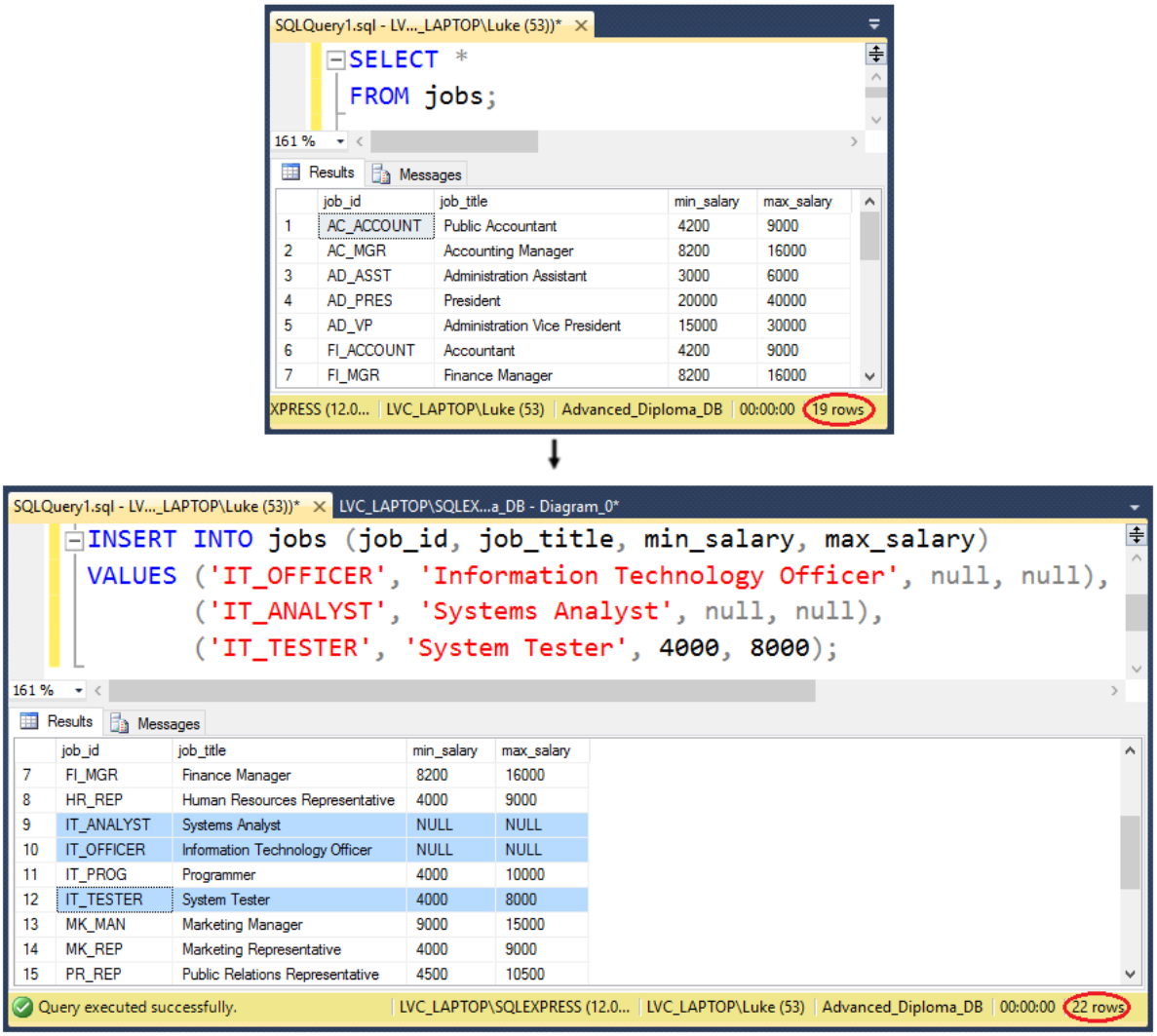


Figure 16 - Result of example 8: three rows added using INSERT VALUES with multiple columns

NOTE:

- In the previous screenshot, explicit method is used in the case of null values (the case with the first two rows inserted).
- Given that all the rows have all the four values, the targeted columns could have been left out.

While using INSERT VALUES one must keep in mind the following common mistakes:

- Columns which are mandatory (NOT NULL) are left without a value
- Columns which have a unique constraint have attempts of duplicate data
- Values which attempt to violate a FOREIGN KEY constraint
- Values which attempt to violate a CHECK constraint
- Values which have an incorrect data type (Data type mismatch)
- Values which are too wide for the column they are intended for.

INSERT SELECT statement

The INSERT SELECT statement is commonly considered as the statement that can be used to copy data from one existing table to another existing table. If the target table does not exist this command will not function.

```
INSERT INTO table [(column [, column . . . ])]
        SELECT {column|expression [,column|expression . .]}
        FROM table
        [WHERE search_condition]
```

It is very important to keep in mind that the number of columns returned by the SELECT statement cannot exceed the number of columns in the target table. Also all the mandatory columns (columns who have NOT NULL and PRIMARY KEY constraints) in the target table should have an entry in the SELECT statement as these columns should always have a value.

Example 9: Write the required statements such that you create a new entity named *employees_names* with three columns – emp_id, name, surname. You are then to copy the name and surname of all the employees to the newly created table.

The below statement will create the employees_names table with the three columns stated above

```
CREATE TABLE employees_names
(
    emp_id INTEGER PRIMARY KEY ,
    name VARCHAR(50) CONSTRAINT en_name_nn NOT NULL,
    surname VARCHAR(50)  CONSTRAINT en_surname_nn NOT NULL
);
```

The below statement copies the name and surname from the employees table to the newly created table employees_names

```
INSERT INTO employees_names
    SELECT employee_id,first_name, last_name
    FROM employees;
```

The below result screenshot, displays all the columns and rows in the newly created table.

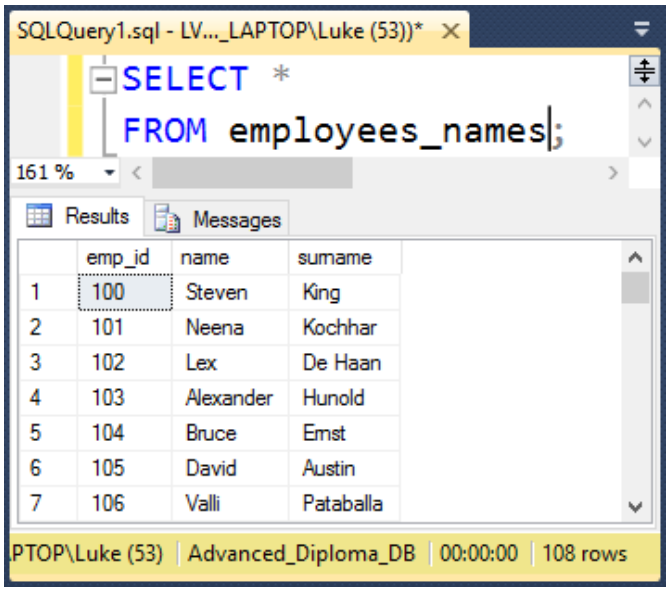


Figure 17 -Result of example 9: rows from employees table successfully inserted in employees_names

Example 10: Write the necessary statements such that a new table (employees_restricted) is created. This table should have 3 columns – *eid* (PRIMARY KEY), *name* (mandatory) and *salary*. You are to include a statement that copies the employee_id, full name and salary from the employees table to the employees_restricted table.

The below code is used to create the target table which will contain the data that is to be copied

```
CREATE TABLE employees_restricted
(
    eid INTEGER PRIMARY KEY ,
    name VARCHAR(50) CONSTRAINT en_name_nn NOT NULL,
    salary NUMERIC(8,2)
);
```

The below INSERT SELECT will take the employee_id, the concatenated name and surname, and the salary of all the employees who earn more than 16999 and place a copy in the employees_restricted entity.

```
INSERT INTO employees_restricted (eid, name, salary)
    SELECT employee_id,first_name +' '+ last_name, salary
    FROM employees
    WHERE salary >= 17000;
```

Figure 18 below shows the three employees which satisfy the condition and which have been copied to the employees_restricted entity

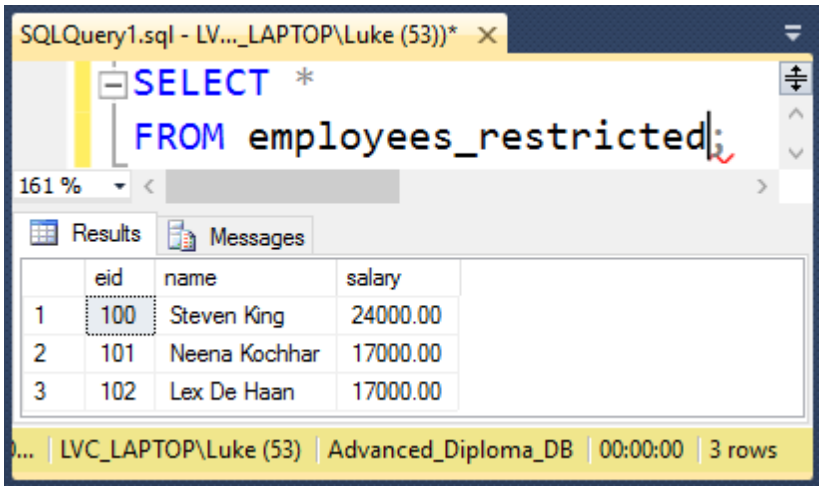


Figure 18 - result of example 10 - employees which were copied from the original table to employees_restricted

SELECT INTO statement

This statement is a bit different from the previous one as it does not require the target table to be created before the actual copying of data takes place. This statement involves a query (found in the SELECT part) and a target entity (created by the statement itself). The syntax for the SELECT INTO statement is the below:

```
SELECT {column [,column . .]}
INTO table
FROM table
[WHERE search_condition]
```

The SELECT INTO statement automatically creates the target entity. The generated target entity will have the same definition of the source database and aspects such as column names, data types, nullability, and IDENTITY property are copied. Although the mentioned aspects are copied, other aspects such as indexes, triggers, permissions and others are not (these need to be specifically applied through scripting from the source entity). Apart from the structure of the entity itself, this statement will also copy the data which satisfies the conditions placed in the SELECT part.

One of the problems of this statement is that you will have no control on the definition of the target entity. Although some things such as IDENTITY and NOT NULL can be modified from within the statement itself, not all the copied aspects can be modified from within the statement. For the scope of this subject we will not attempt to modify the target table from within the SELECT INTO statement itself.

Example 11: Write a query that will copy all the data found in the employee number, name and salary columns within the employees table, into a new table which does not even exist in the database. The new table in which the extracted data is to be placed should be named employees_restricted2.

Figure 19 shows the code that is used to copy all the 108 rows extracted from the employees table to the newly generated table employees_restricted2 table. Notice that after executing the statement and refreshing the tables folder in the Object Explorer, the new table (employees_restricted2) is available and the 3 columns selected in the SELECT part are included.

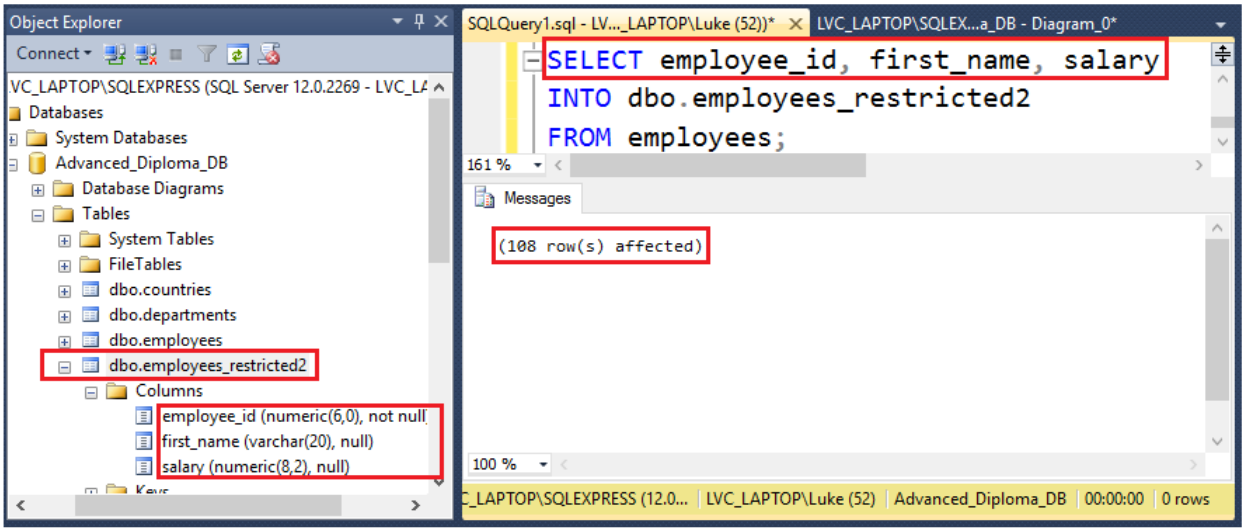
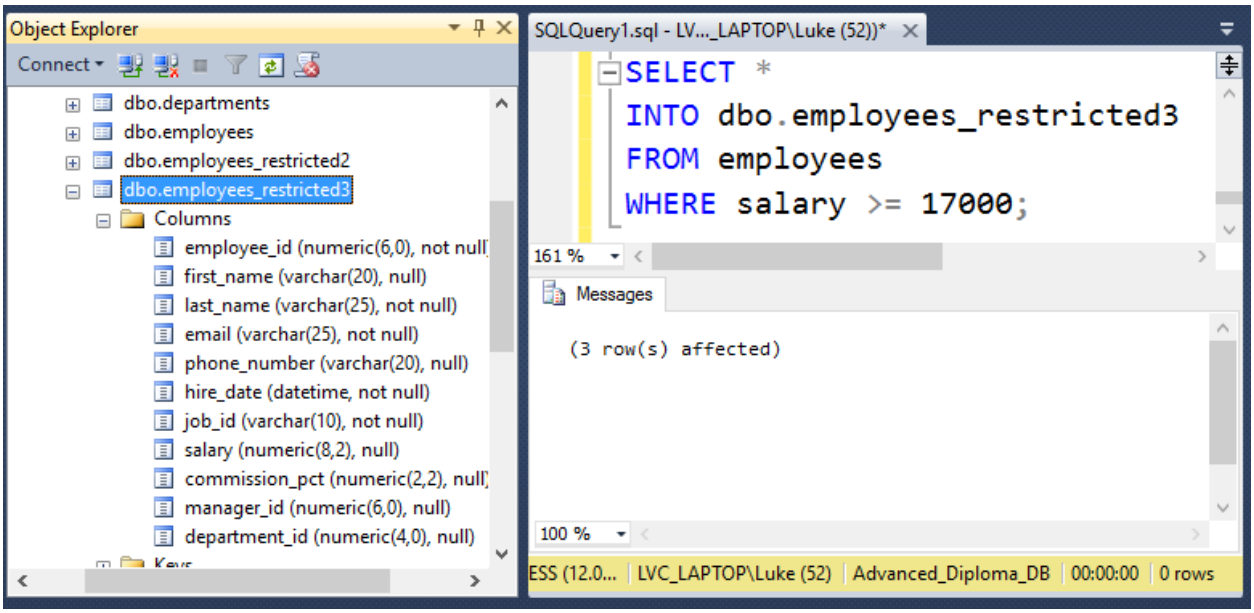


Figure 19 - Result of example 11: extracting 3 rows with the data and placing them in a new table

Example 12: Write a query that will copy all columns and rows of data for those employees who earn 17000 or more. One single statement should be used to extract the desired information and place it in a new table named employees_restricted3. Note that this time round the new table will only have 3 rows of data inserted in it as we included a WHERE condition.



Modifying Rows of data

Whenever data in a table is to be modified the UPDATE statement is to be used. The standard UPDATE statement will be considered in this subject but T-SQL offers a number of extensions which can be used. These extensions require the use of more complex material (some of which will be covered in the coming weeks).

Considering the table below we can be in a situation where we will need to change any of the values in the table below. A common change in such a table is that of the price or the supplier. If we consider a shop which sells the below products, there could be a situation where the price for a particular product changes due to certain circumstances.

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10
3	Banana	0.60	20

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10
3	Banana	1.10	20

In order to be able to perform such changes, you are to make use of the following UPDATE syntax:

```
UPDATE table
SET column = value [, column =value, . . .]
[WHERE condition];
```

The above syntax is made up of the UPDATE keyword which should be followed by the name of the table which is to have the modifications. Soon after the table is declared, the SET keyword is to be used, together with all the columns which are to be effected and their new values. Multiple columns can have their values updated with a single UPDATE statement but it is important that between each column-value combination a comma is included. The WHERE clause is an optional part which needs to be handled very carefully. If the WHERE clause is not included then all the rows will be updated. If you want to restrict the changes to particular row/s, then it is very important to specify the correct WHERE condition.

Example 13: Write a query that will change the name of employee 207 from Charles to Chalie.

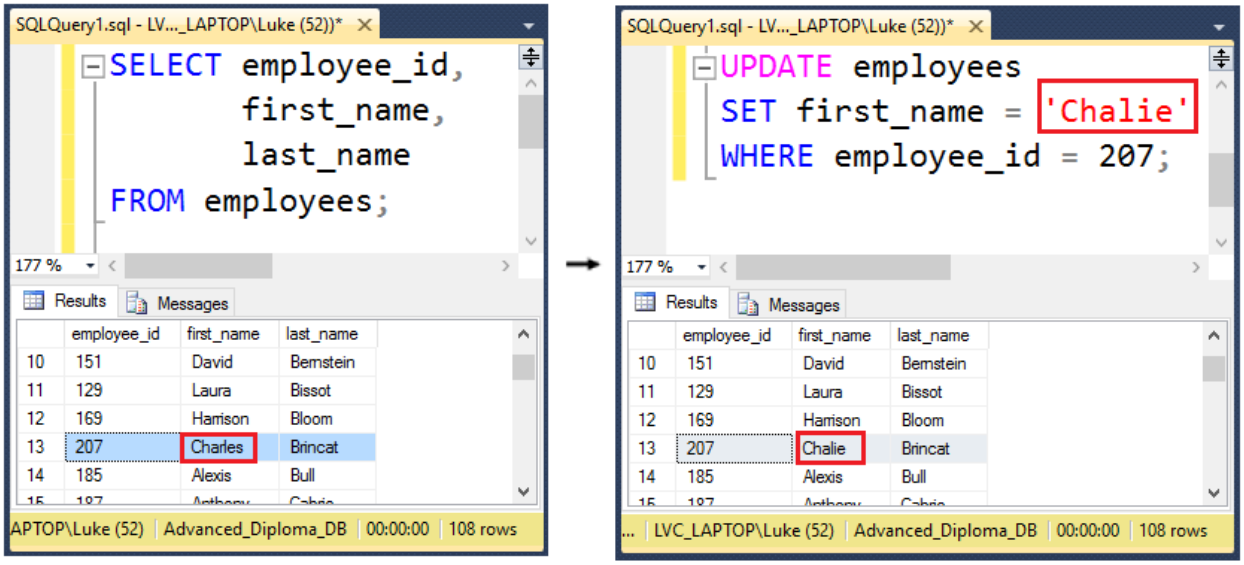


Figure 20 - Result of example 13: Update statement with a single column update

Example 14: Write a query that will change three columns (first_name, hire_date, salary) for the same employee. Employees 207 should be updated such that his first_name is changed to Chalie, hire_date to the last day of the year 2000 and the salary should be an increase of 500 to the current salary.

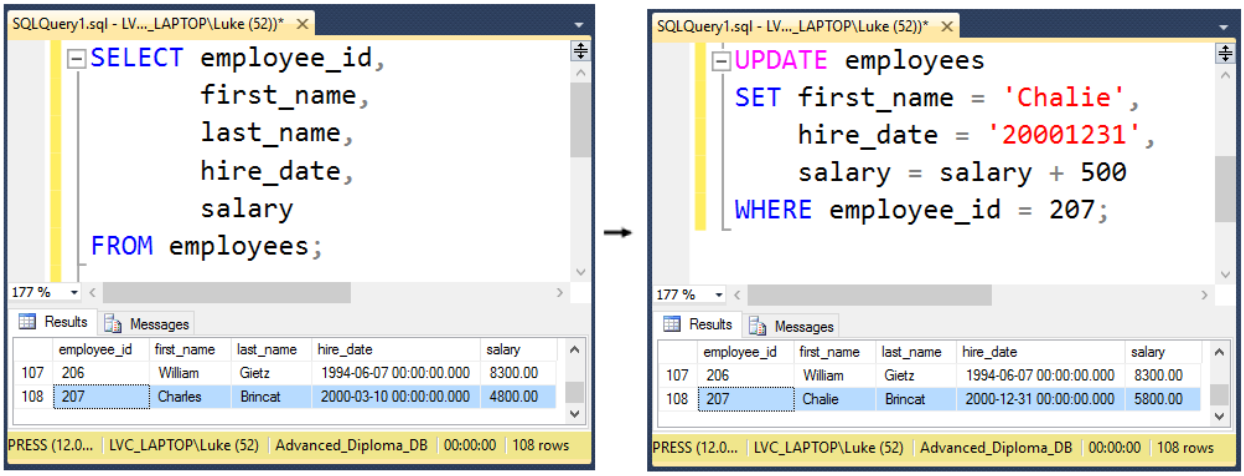


Figure 21 - Result of example 14: the update statement used to change multiple columns

Note: In the above screenshot, it evident that columns storing character and date should be enclosed in single quotes, but numerical columns should not.

Example 15: Write a query that will increase the employees' salary by 500.

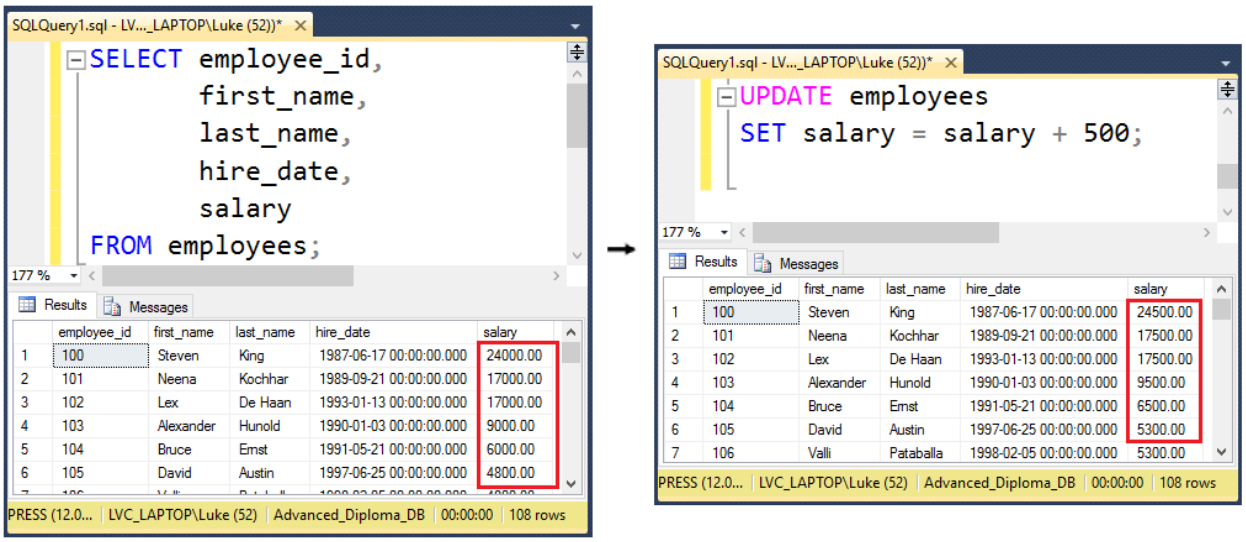


Figure 22 - Results of example 15: all the salaries in the employees table added by 500

Note: given that all the rows have been updated, the WHERE clause was not used.

Example 16: Write a query that will change the job_id and salary of employee 114, such that the values become equal to the one's of employee 205

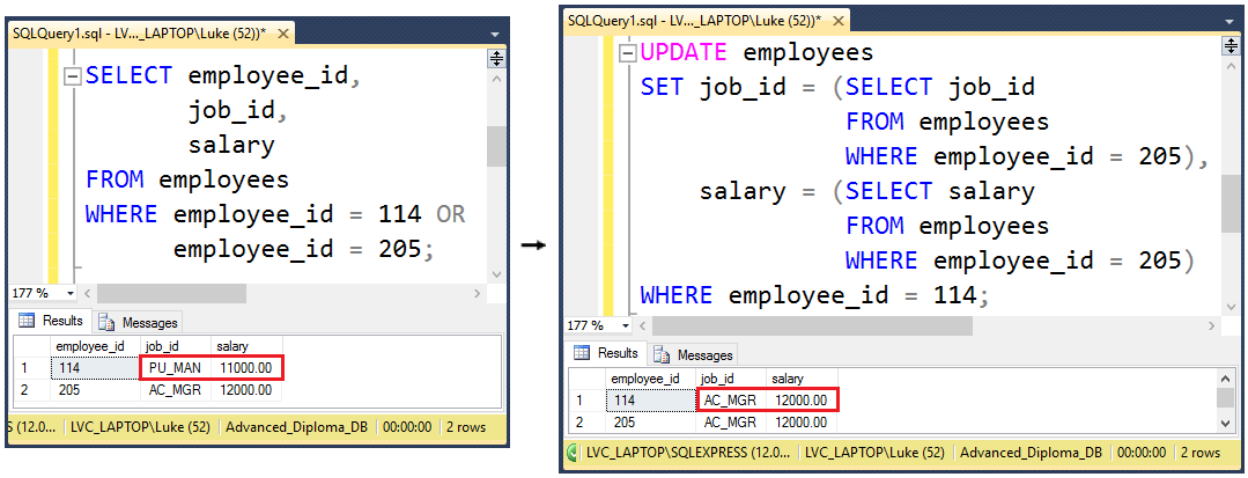


Figure 23 - Result of example 16: change the values using sub queries

Removing Rows of data

Sometimes it is necessary to be able to remove rows of data from a particular table. The reasons behind the use of such an operation might be many but the idea is to use a condition which will select the row/s which are to be permanently removed from the database. In T-SQL there are two main statements that help the user to perform this operation - DELETE and TRUNCATE.

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10
3	Banana	1.10	20

Pid	Product Name	Product Price	SupplierId
1	Orange	0.50	10
2	Apple	0.40	10

The above is what happens when a row is deleted from a particular table. As shown in the above, the third product (Pid=3) has been removed and the current table is now left with only two rows of data.

DELETE statement

The delete statement is used to remove rows of data from the database. The good thing with regards to this statement is that you can select which rows you want to delete by specifying a condition in the WHERE clause

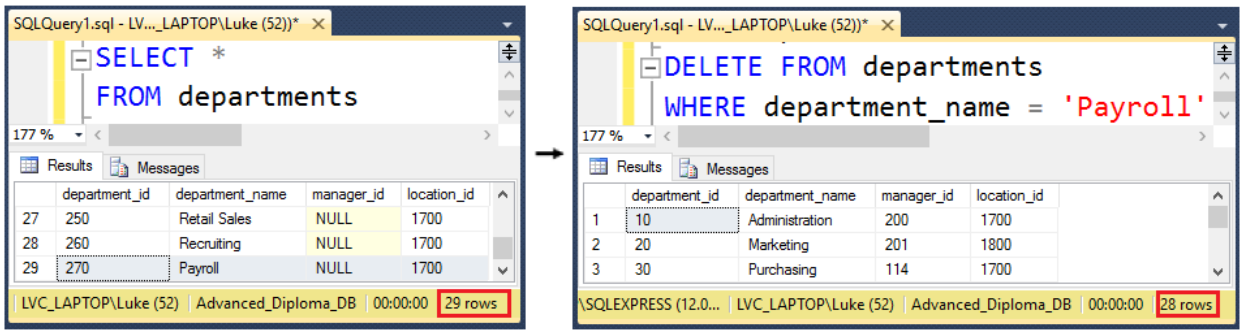
```
DELETE FROM table
[WHERE condition];
```

As with the case of the UPDATE statement it is very important to note that the use of the WHERE clause is paramount. Although this clause is optional, failing to include it will result in the loss of all the rows in the particular table.

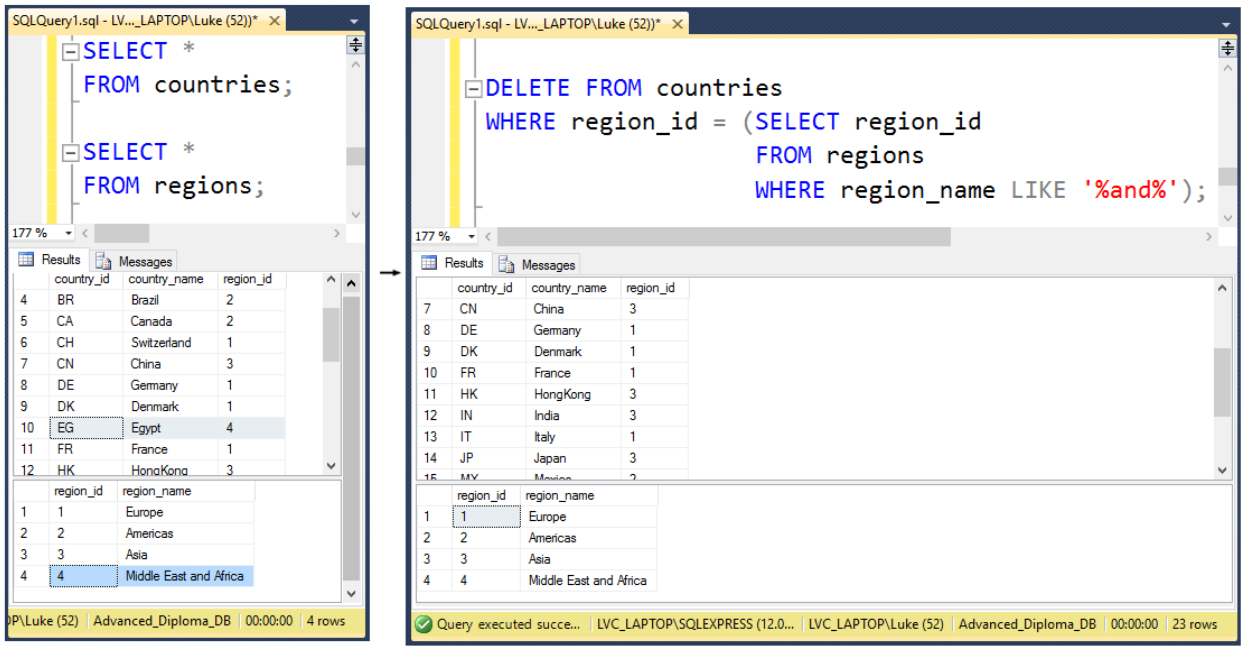
Also note, that if a row is being used in a FK relationship you will not be allowed to remove it.

Another important thing to note, while using this statement is that whenever the database increases in size, delete operations might introduce time delays. Given the fact that delete operations will be logged in the transaction file, this may result in a huge number of entries in this file – hence a lot of storage will be required to store the file. Also locking mechanisms might be applied on the rows or even on the whole table, restricting access to the user.

Example 17: Write a query that will remove the payroll department from the departments table



Example 18: Write a query that will delete all the countries in the countries table which include an 'and' in the region name. The Delete query is using a sub-query which is a query within another query.



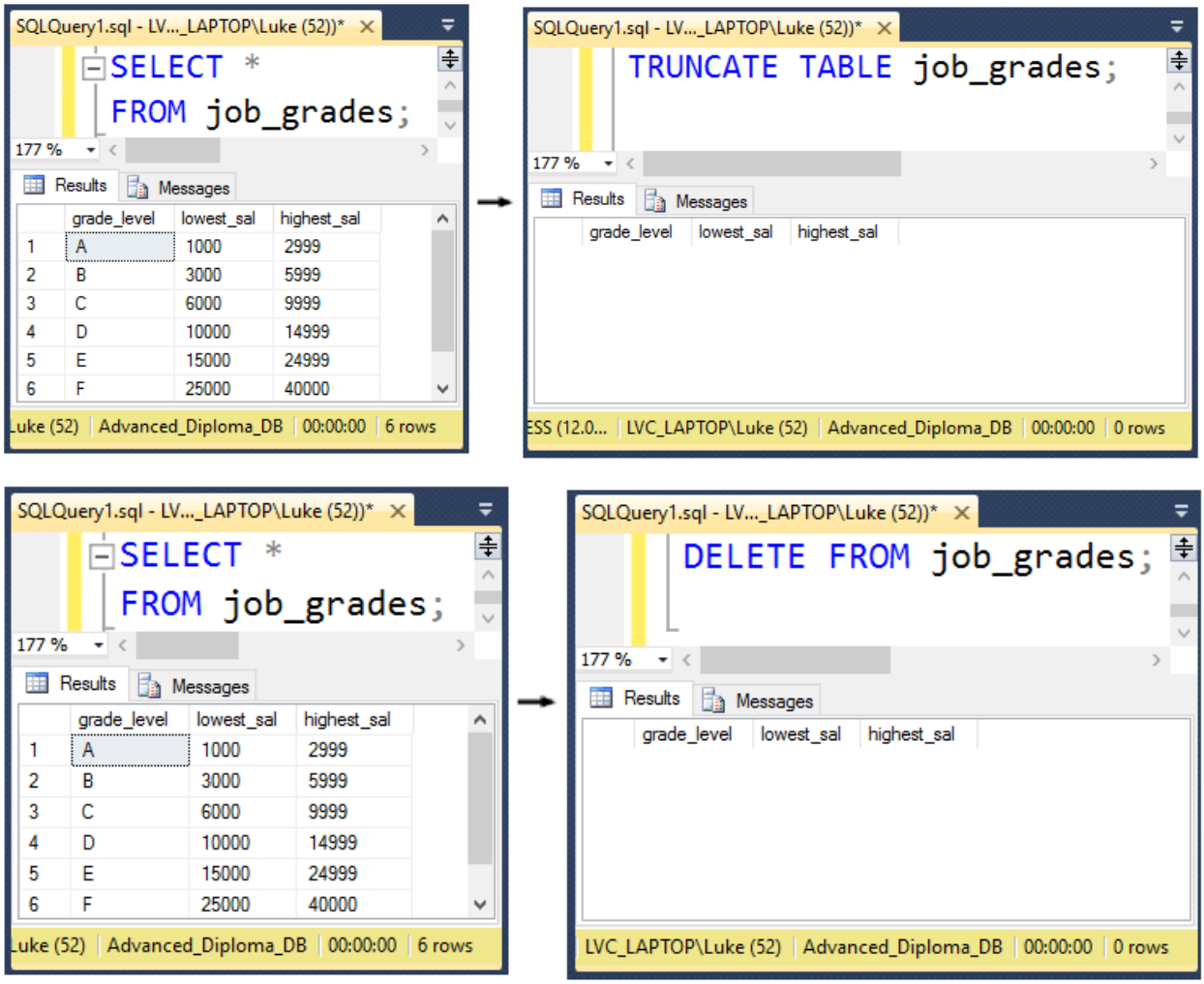
Note that 6 rows will be deleted and in the screen shot above one of the countries which was deleted (Egypt) is being displayed.

TRUNCATE statement

TRUNCATE is the alternative of the DELETE statement when a user intends to remove rows of data from a table. Unlike the DELETE option, this statement does not allow the use of the WHERE clause, therefore you either delete all the rows or nothing. Once that you use the TRANCATE statement the target table will become empty. The syntax for the TRUNCATE statement is listed below:

TRUNCATE TABLE table

Example 19: Write a query that will delete all the rows in the job_grades entity



Note that the two queries above perform the exact same thing