Step 1: Create a table to stage data as is. So, let the data type of all its column be varchar(50)

```
CREATE TABLE StgData

(
CredentialNumber VARCHAR(50),
LastName VARCHAR(50),
FirstName VARCHAR(50),
MiddleName VARCHAR(50),
CredentialType VARCHAR(50),
Status VARCHAR(50),
BirthYear VARCHAR(50),
CEDueDate VARCHAR(50),
FirstIssueDate VARCHAR(50),
LastIssueDate VARCHAR(50),
ExpirationDate VARCHAR(50),
ActionTaken VARCHAR(50))
```

Step 2: Create a Dest table to store the clean data (with correct data types)

```
CREATE TABLE DestData

(
CredentialNumber VARCHAR(10) PRIMARY KEY NOT NULL,
LastName VARCHAR(15) NOT NULL,
FirstName VARCHAR(15) NOT NULL,
MiddleName VARCHAR(15),
CredentialType VARCHAR(45) NOT NULL,
Status VARCHAR(10) NOT NULL,
BirthYear INT NOT NULL,
CEDueDate DATE NULL,
FirstIssueDate DATE NULL,
LastIssueDate DATE NULL,
ExpirationDate DATE NULL,
ActionTaken VARCHAR(3) NOT NULL
)
```

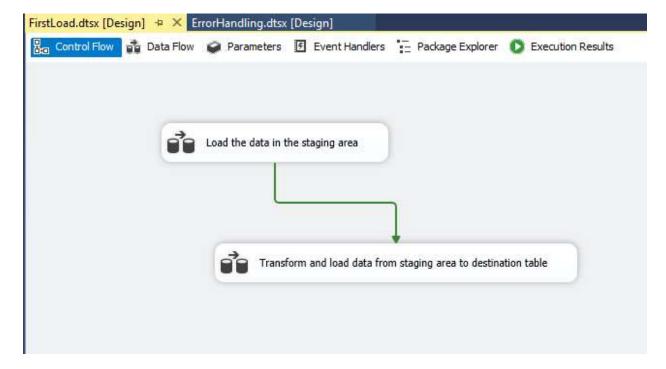
Create ErrorData table to store all the rows that errored out, along with the type of error:

```
CREATE TABLE ErrorData

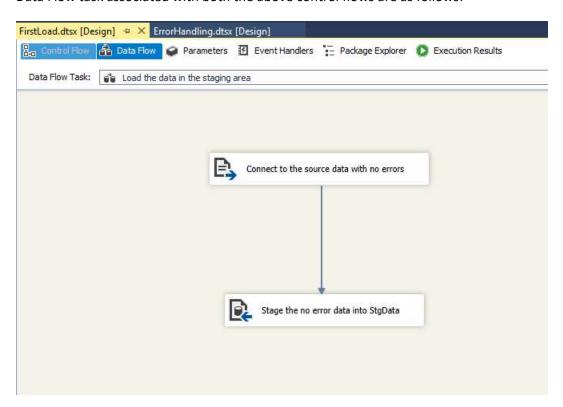
(
CredentialNumber VARCHAR(10) NOT NULL,
LastName VARCHAR(15) NOT NULL,
FirstName VARCHAR(15) NOT NULL,
MiddleName VARCHAR(15) NULL,
CredentialType VARCHAR(45) NOT NULL,
Status VARCHAR(10) NOT NULL,
BirthYear INT NOT NULL,
CEDueDate DATE NULL,
FirstIssueDate DATE NULL,
LastIssueDate DATE NULL,
ExpirationDate DATE NULL,
ActionTaken VARCHAR(3) NULL,
ErrorType NCHAR(50)

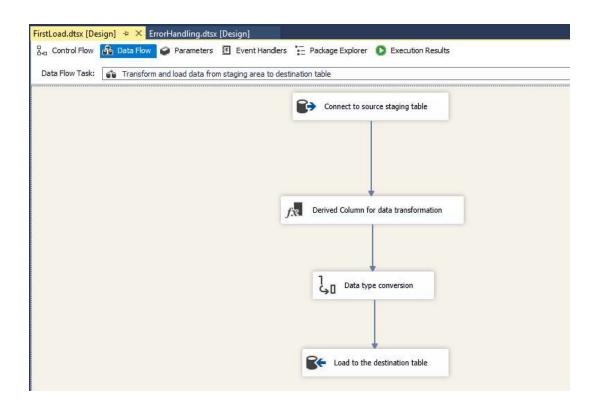
)
```

Step 3: Load the data from the no error flat file source directly to the staging table using SSIS. And then, use a data conversion task to convert the columns to proper data types and load it to the destination. This is done by the package FirstLoad.dtsx. It has 2 control flow components as follows:

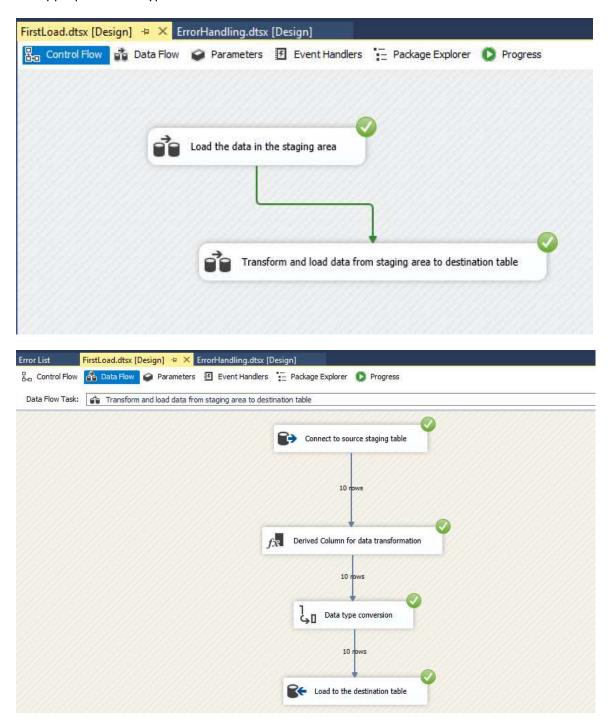


Data Flow task associated with both the above control flows are as follows:





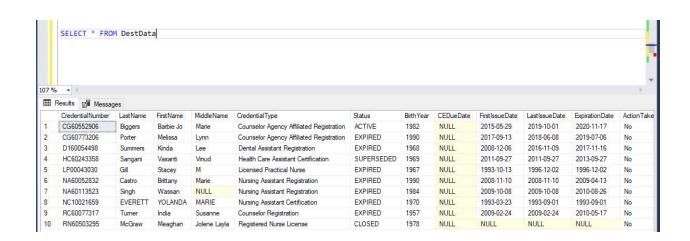
Finally, execute the package, to load the data in the staging table, and then to the destination table with the appropriate data types.



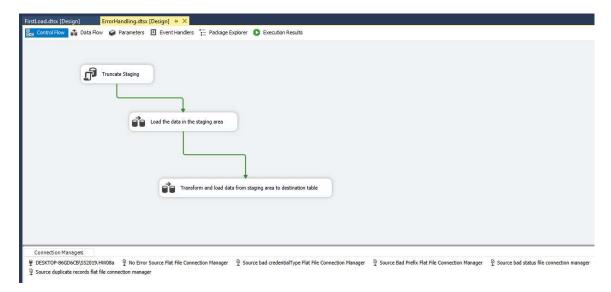
This shows the successful execution of SSIS package, and 10 rows are loaded into the staging table and the destination table.

Let's verify that in SSMS.

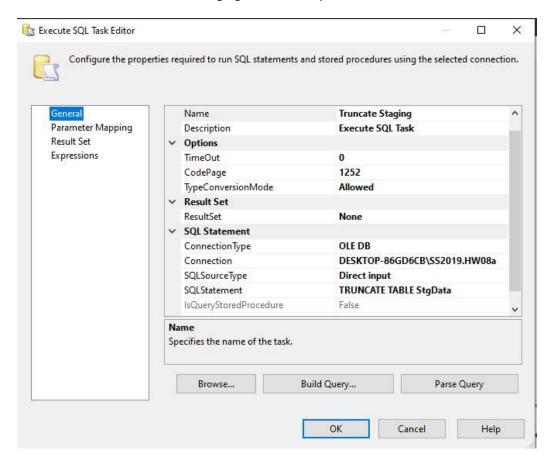




Step 4: Error Handling part: Configure the Error Handling package as follows:



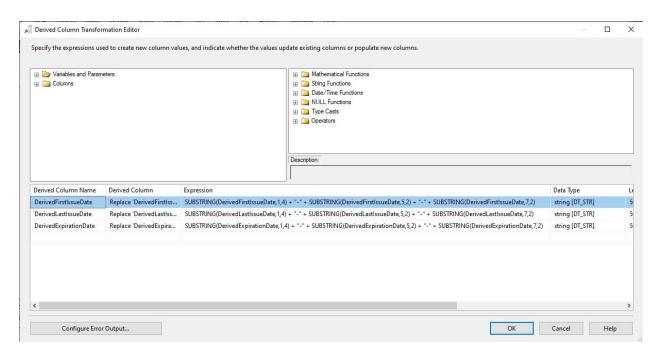
Truncate Staging- First Load populates the staging area and the dest table with the no error data. We do not want that to be loaded again and again, so whenever we pass in a new file, we need the previous data to be truncated from the staging tale, and only the new data is inserted.



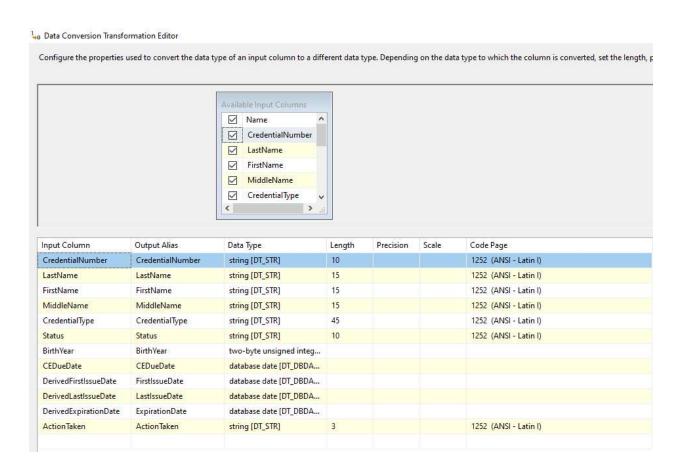
Then load the error data in the staging table, we have 4 different connection managers to connect to 4 different error flat files.

Finally, we put the data in the destination table after doing the data conversion and error handling part as follows:

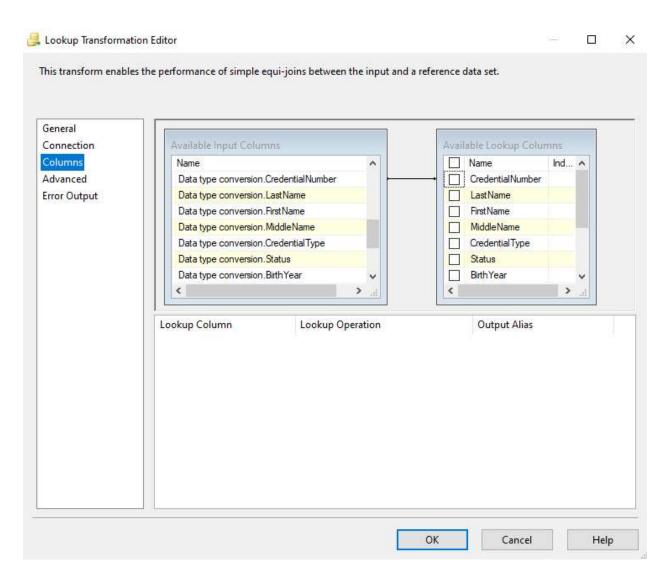
i) Converted the date into proper format using derived column as follows:



ii) Performed data type conversion to change the data types of all the columns as required by the destination table in SSMS as follows:

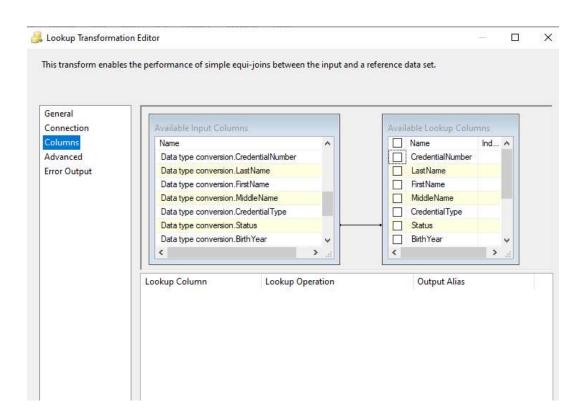


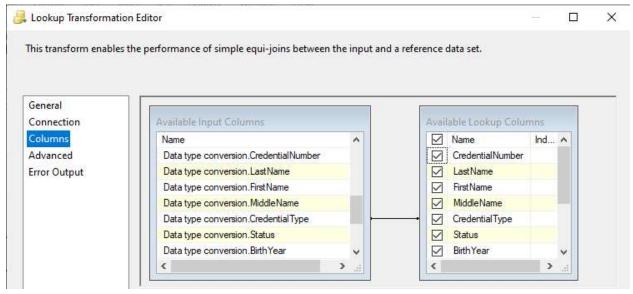
iii) Lookup to handle duplicate records:



This lookup makes sure that no row with the duplicate credential number is inserted into the dest table. If it finds a duplicate record, the record is inserted in the ErrorData table.

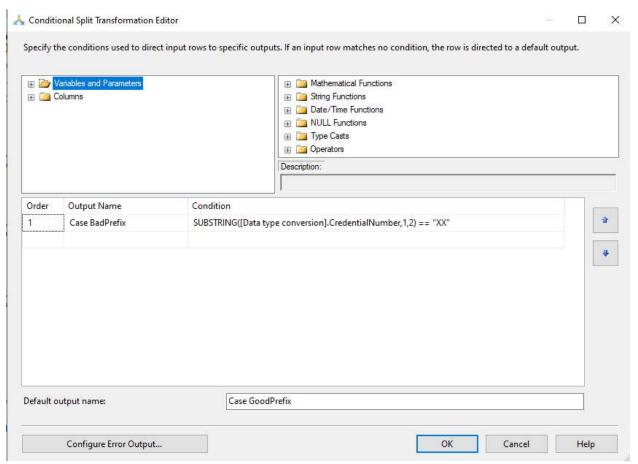
iv) Similarly, lookup to handle bad status type and bad credential type as follows:





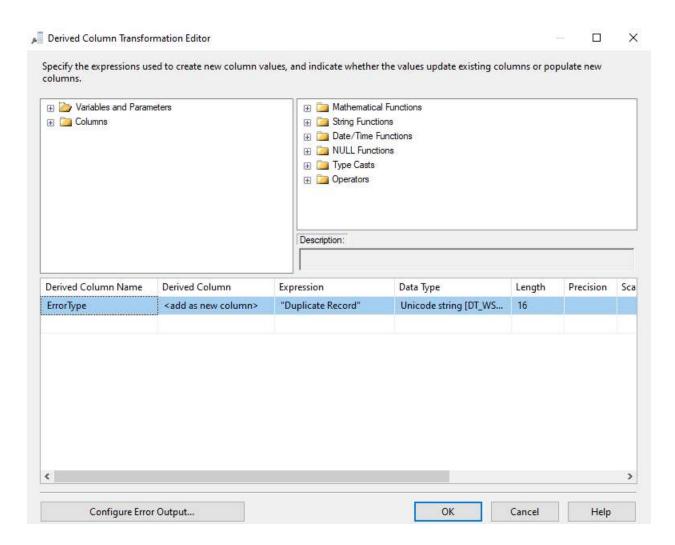
These two lookups make sure that the status type and the credential type is amongst the one that is already present in the destination table. If not, the entire row associated with it is inserted in the ErrorData table, with their corresponding error types.

v) Conditional split to handle Bad Prefix for Credential Number:



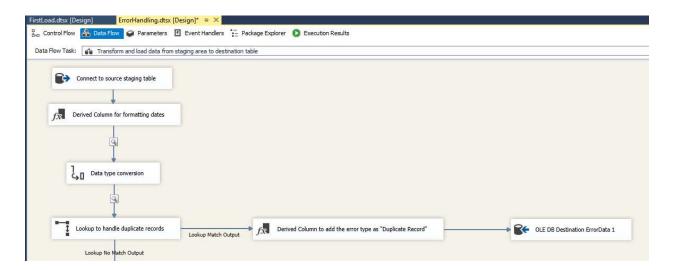
Here, I have created 2 cases, case BadPrefix if the credential number starts with "XX", otherwise it is a good prefix. If this encounters a bad prefix, the entire row is redirected to be inserted into the ErrorData table with error type "Bad Prefix"

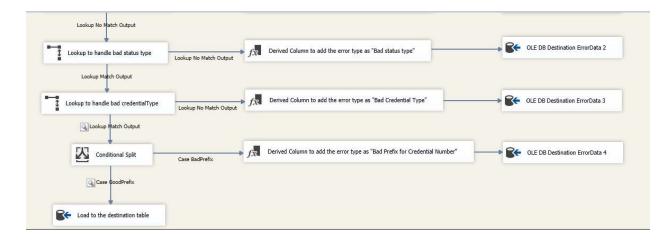
vi) Configuration of the ErrorType value (example of "Duplicate records" is shown below):



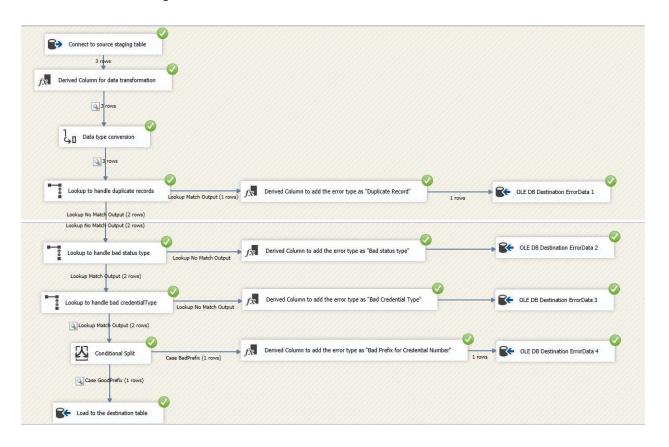
All the other error types are configured in the similar fashion.

vii) Entire data flow:



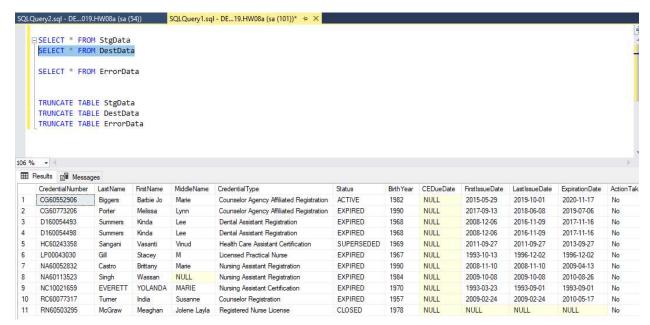


- viii) Finally, we execute the package by changing the source files.
- ix) Bad Prefix Error handling:



Here, we have only 1 row inserted into the destination table, as the other 2 had errors, one associated with duplicate record and the other with the bad prefix for credential number.

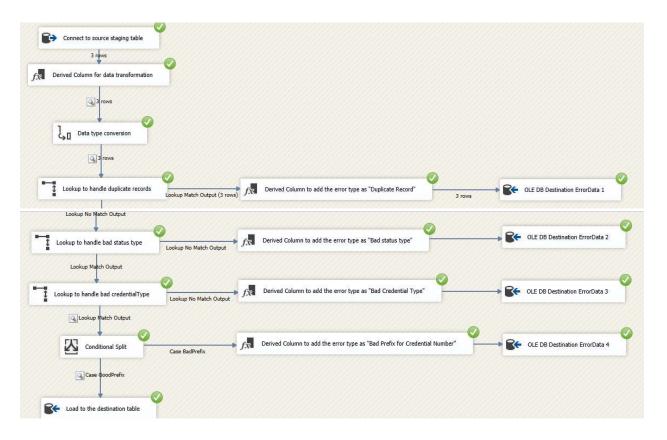
Verifying the tables in SSMS:



We have 11 rows now, one inserted from the source file, and 2 errored out and inserted in the ErrorData table with error type as "Duplicate Record" and "Bad Prefix for Credential Number" as below:



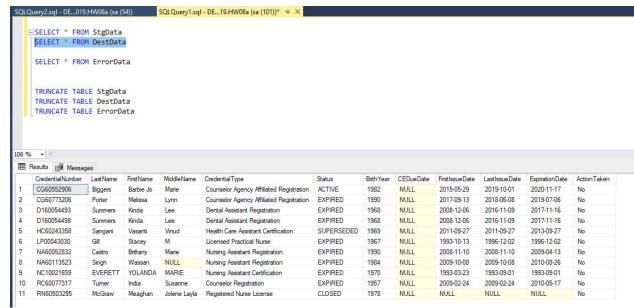
x) Check for duplicate records:



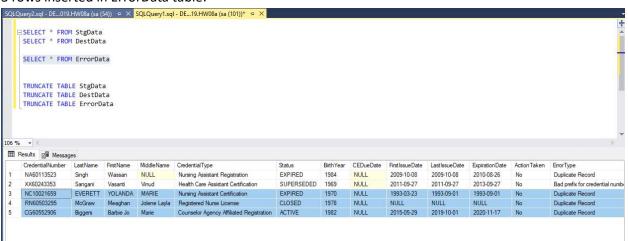
So, none of the rows were inserted into the destination table, and all the 3 errored out with error "Duplicate records"

Verifying in SSMS:

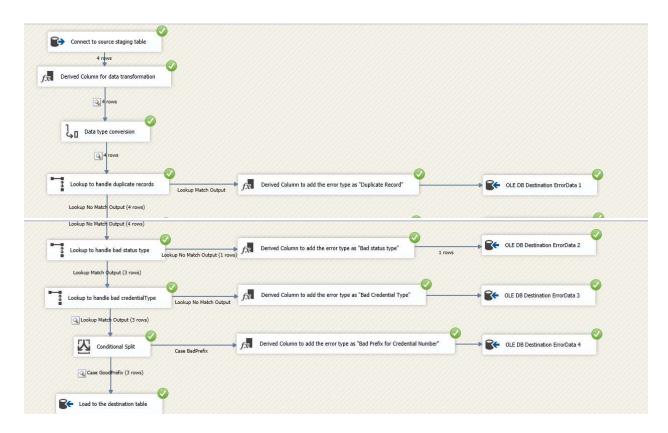
0 rows inserted in DestData:



3 rows inserted in ErrorData table:

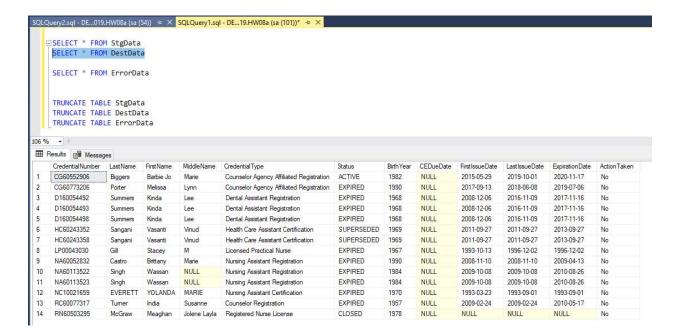


xi) Check for Bad Status type:

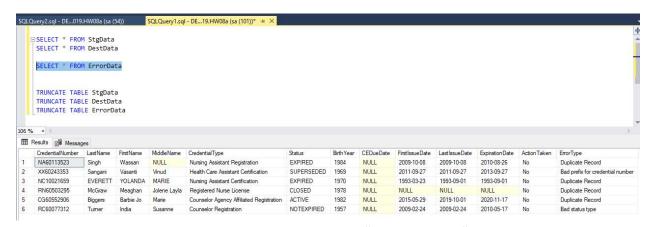


Here, we have 3 rows with good data and inserted into the destination table, and 1 row errored out with bad status type and inserted into ErrorData table.

Verifying in SSMS:

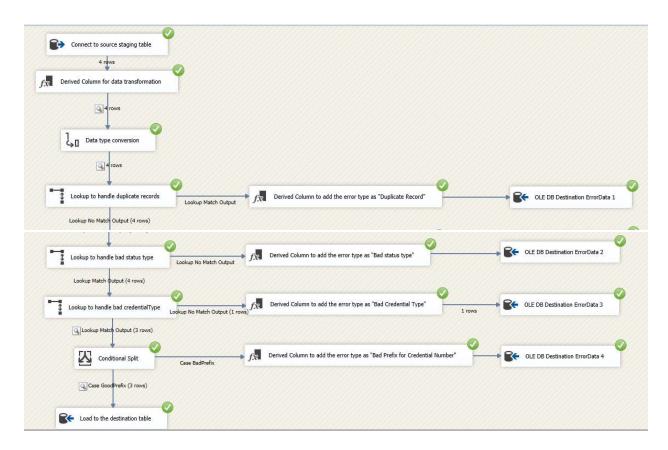


3 new rows inserted (row no. 12, 13 and 14)



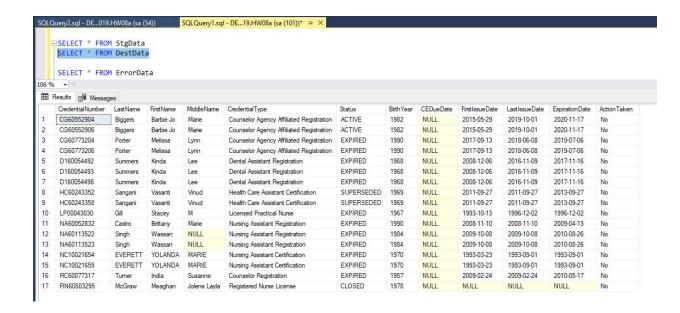
1 row inserted into the ErrorData table, with error type as "Bad status type"

xii) Checking for bad credential type:



3 rows inserted into the destination table, and one row with bad credential type inserted into the ErrorData table.

Verifying in SSMS:



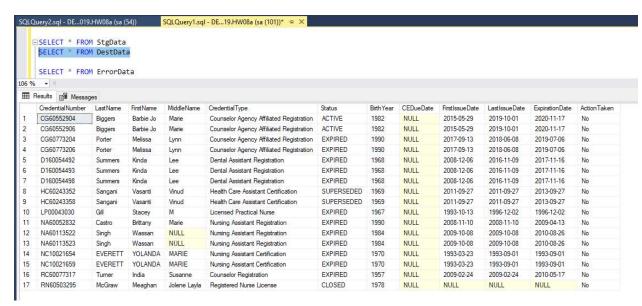
Last 3 rows inserted into the DestTable



Last row inserted into the ErrorData table with error type "Bad credential type"

Final output:

17 rows in the DestData table:



7 rows errored out, with the errors mentioned under the column "Error type"

