Title: ITMD526_Assignment_03

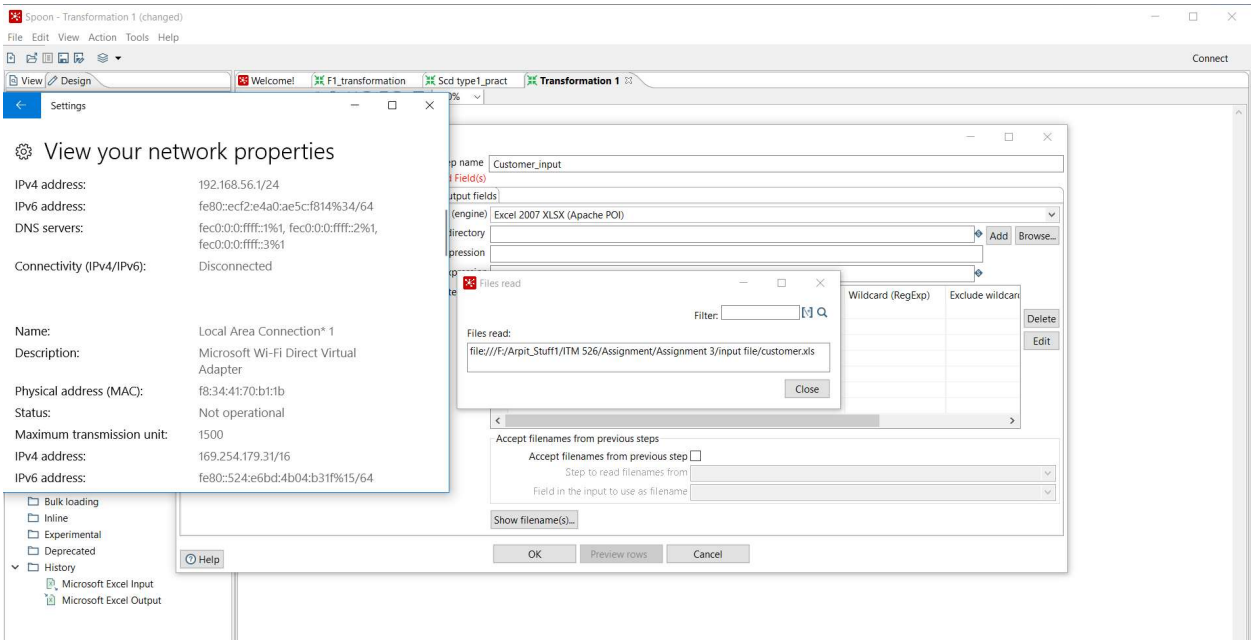
First Name	Last Name	CWID
Arpit	Khandekar	A20409171

Table of Contents

1. Slowly Changing Dimension Type |(Update) Transformation

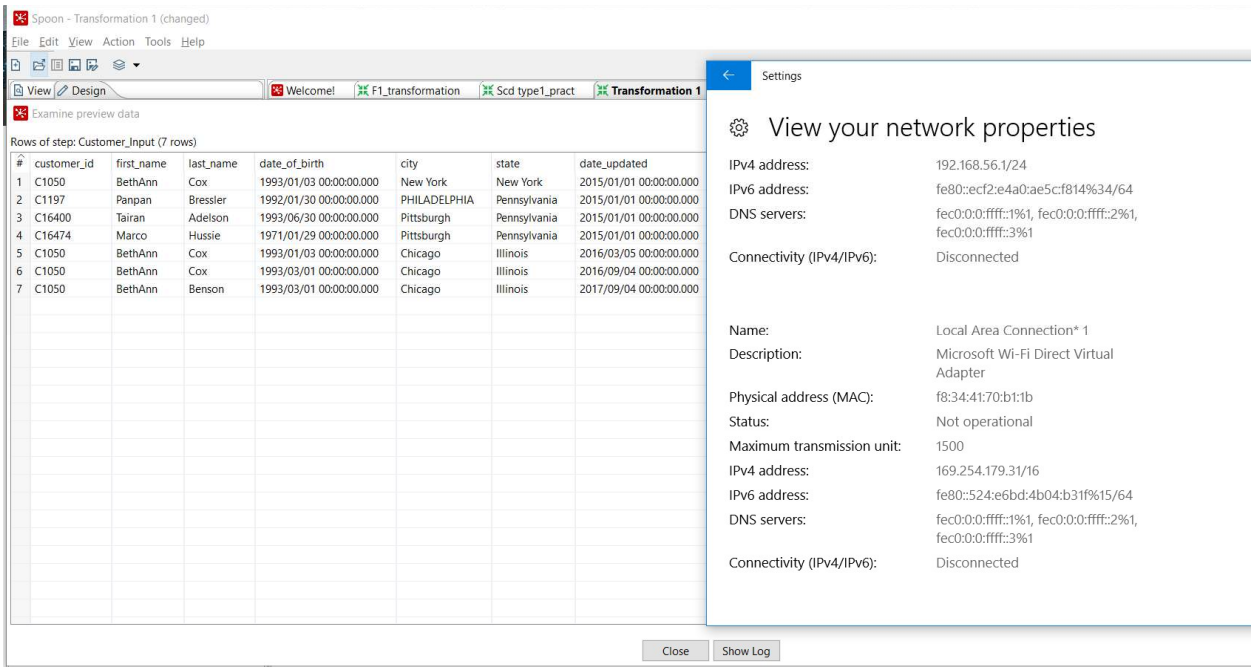
1. First, we need to take Microsoft Excel input as an input in our transformation.

2. Then in Pentaho tool I have drag and drop Microsoft excel as an input named as **‘Customer_input’**



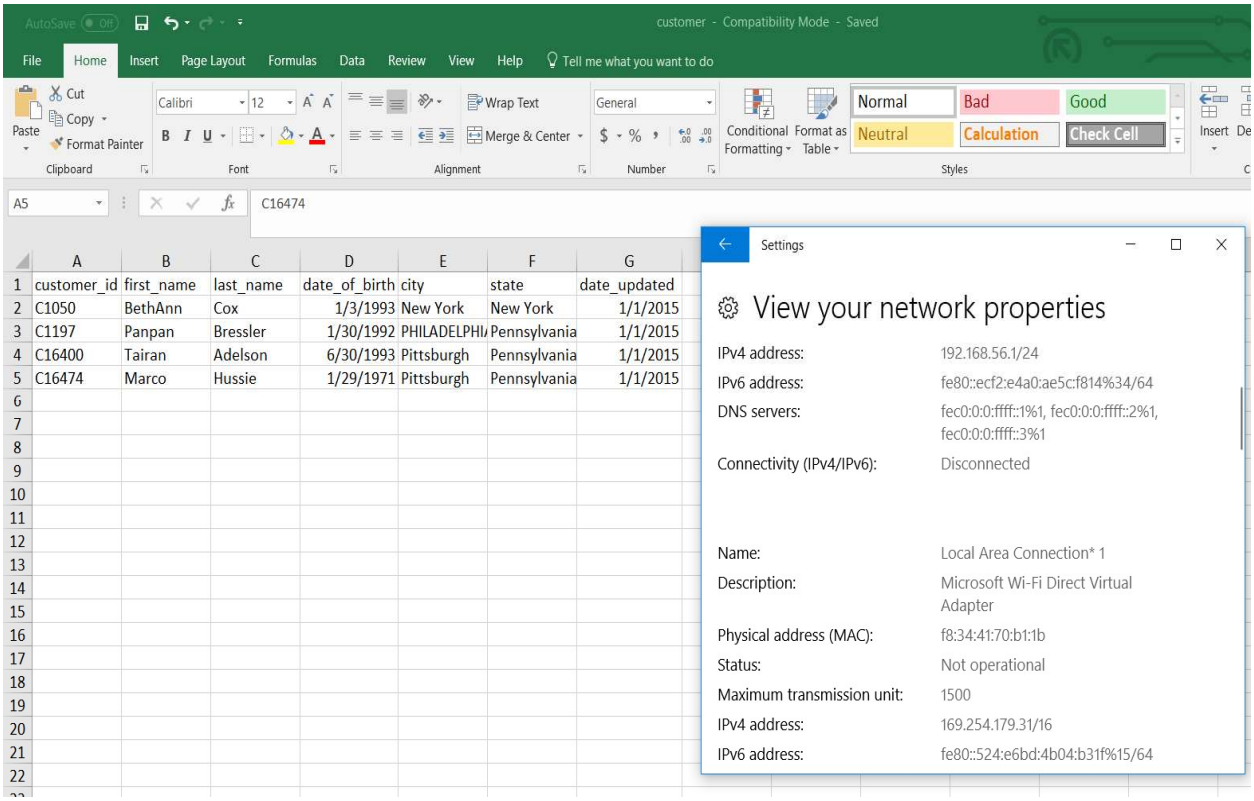
3. I have browse a xlsx file from the directory and uploaded in Excel input.

4. Preview of 1000 rows from the excel.



5. Selected Spread Sheet type(engine) as Excel 2007 XLSX (Apache POI), so that pentaho should know the type of input file.

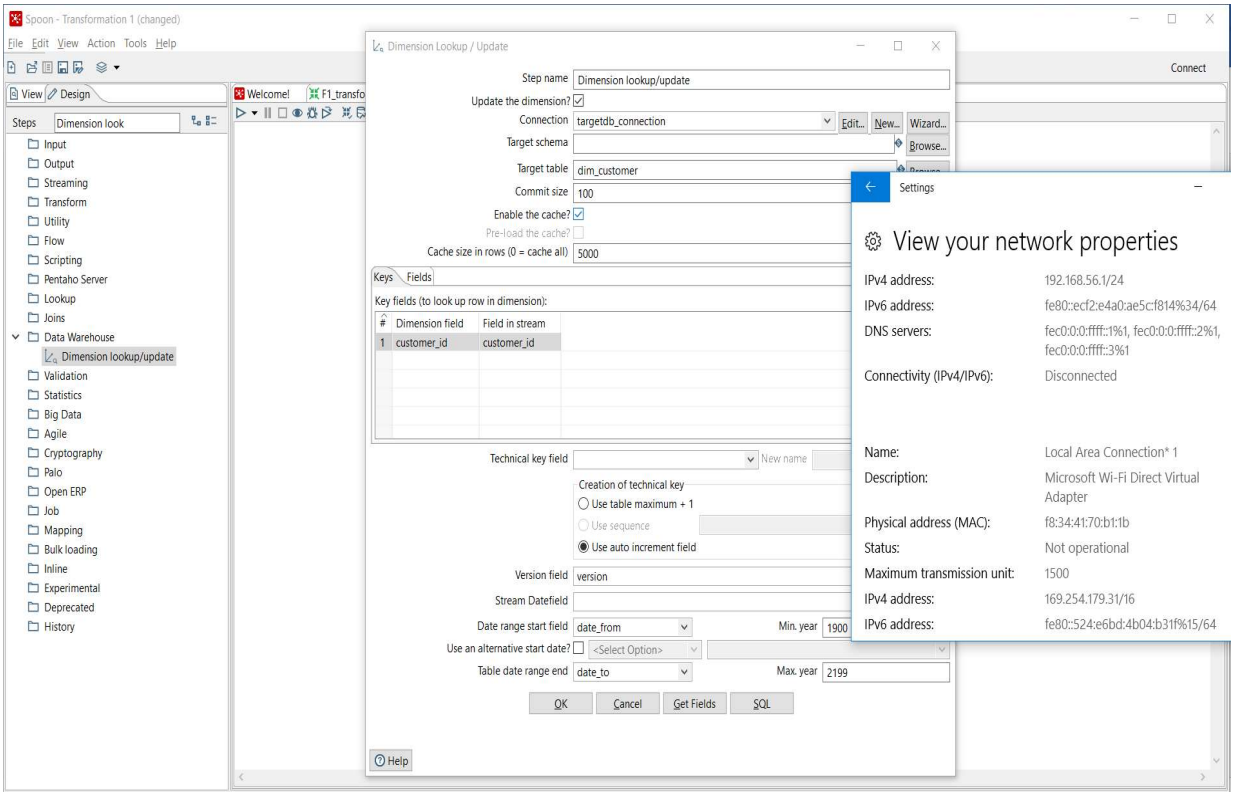
6. In the field tab we can get the header row of xlsx input file.



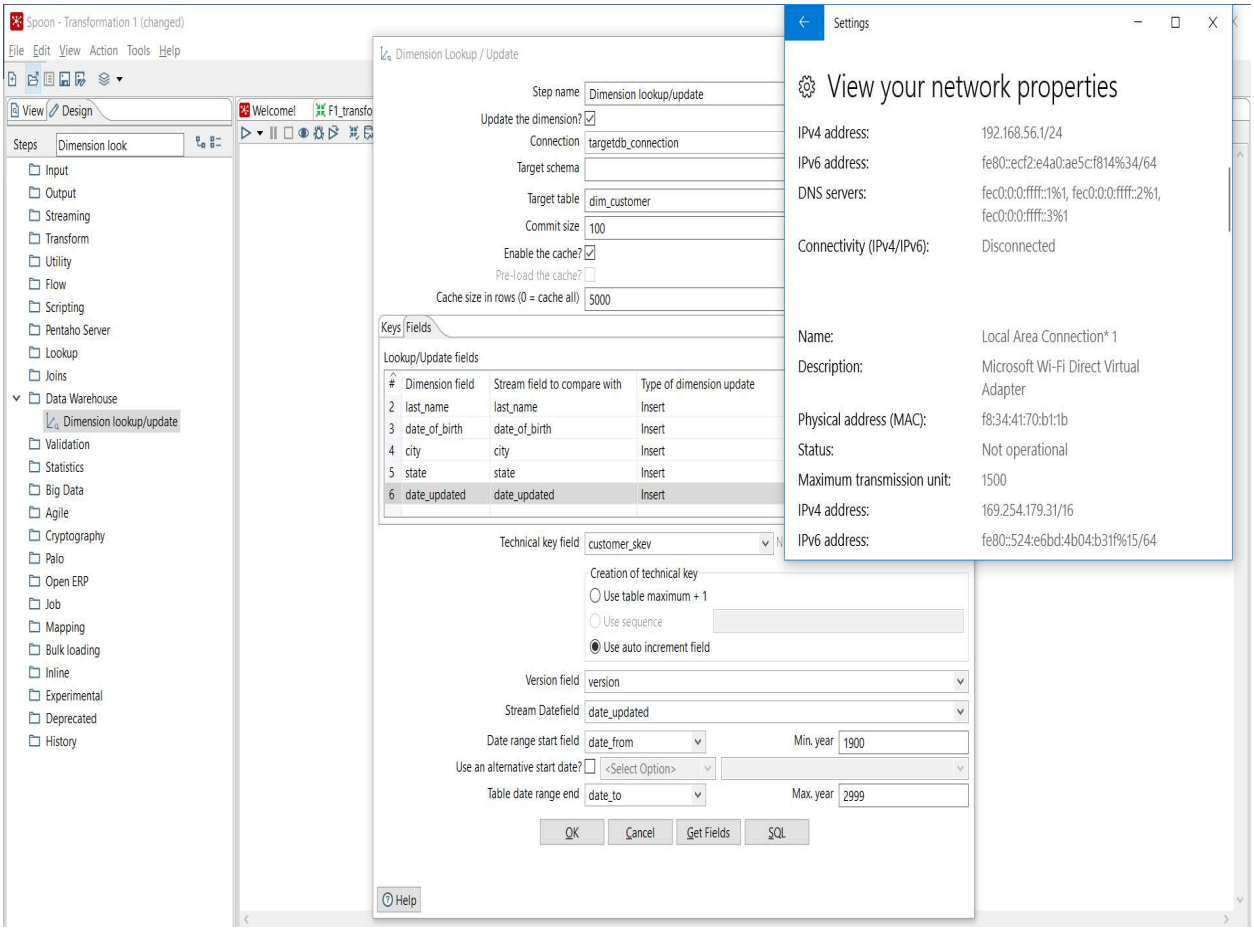
7. Now I have taken Dimension lookup/update output in the transformation.

8. It is connected with my DB “target_db” and target_table as “dim_customer”.

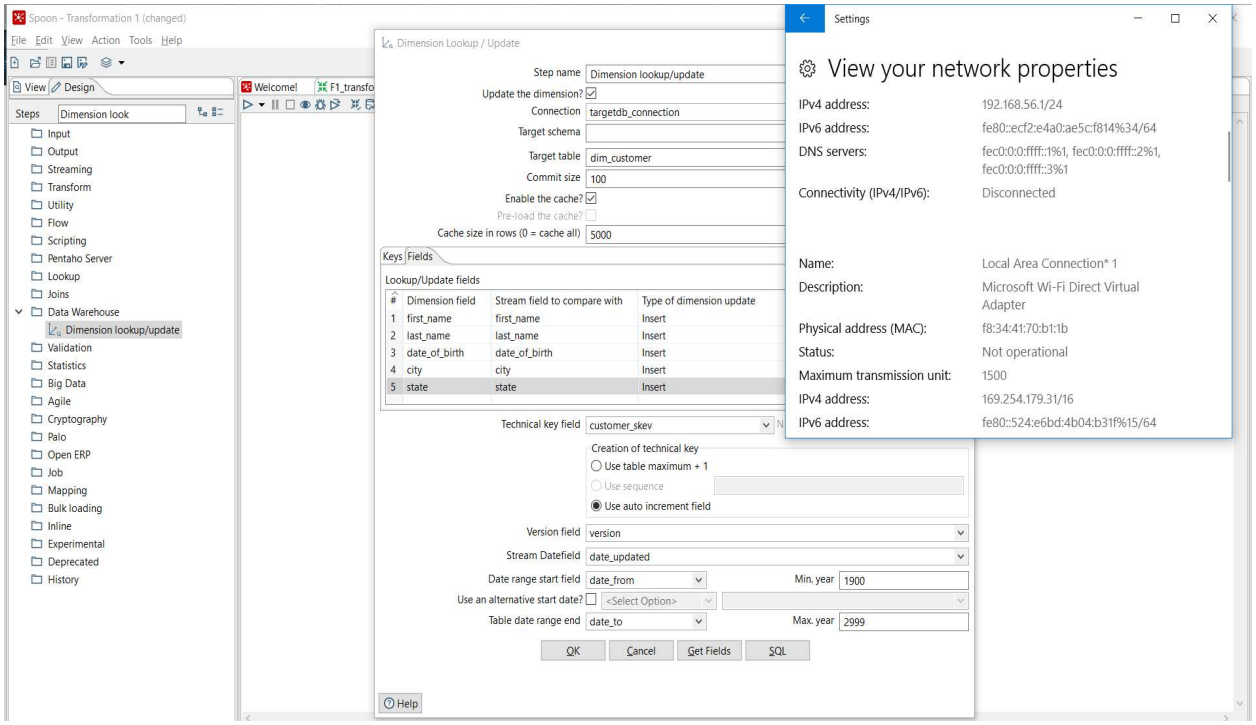
9. Customer_di is identified as a natural key therefore I have put this field in keys section of Pentaho.



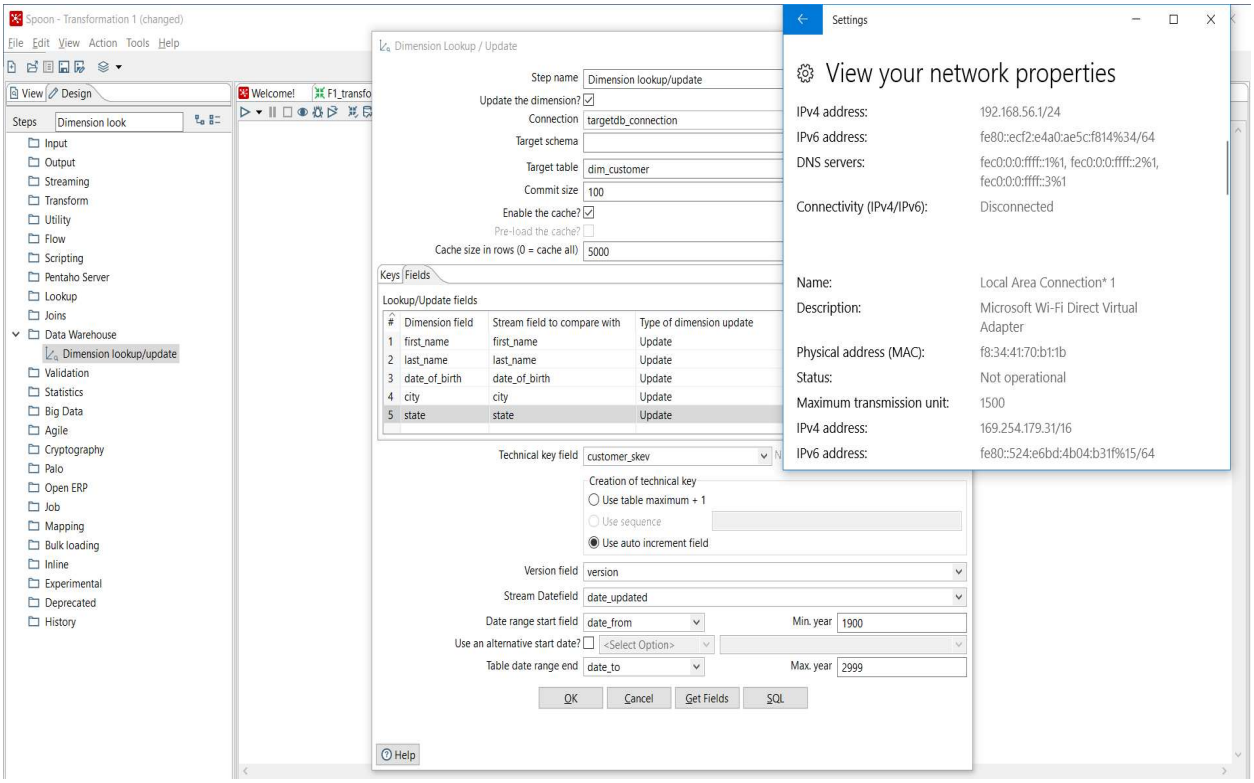
10. To define surrogate key we have defined surrogate key in Technical key field “**customer_skev**” and used auto increment field.
11. We have changed date range start field from year **1900** and table date range end date to **2999,** furthermore I have set Stream Datefield as **date_updated.**
12. We will select get fields to get all the fields in lookup/update fields.



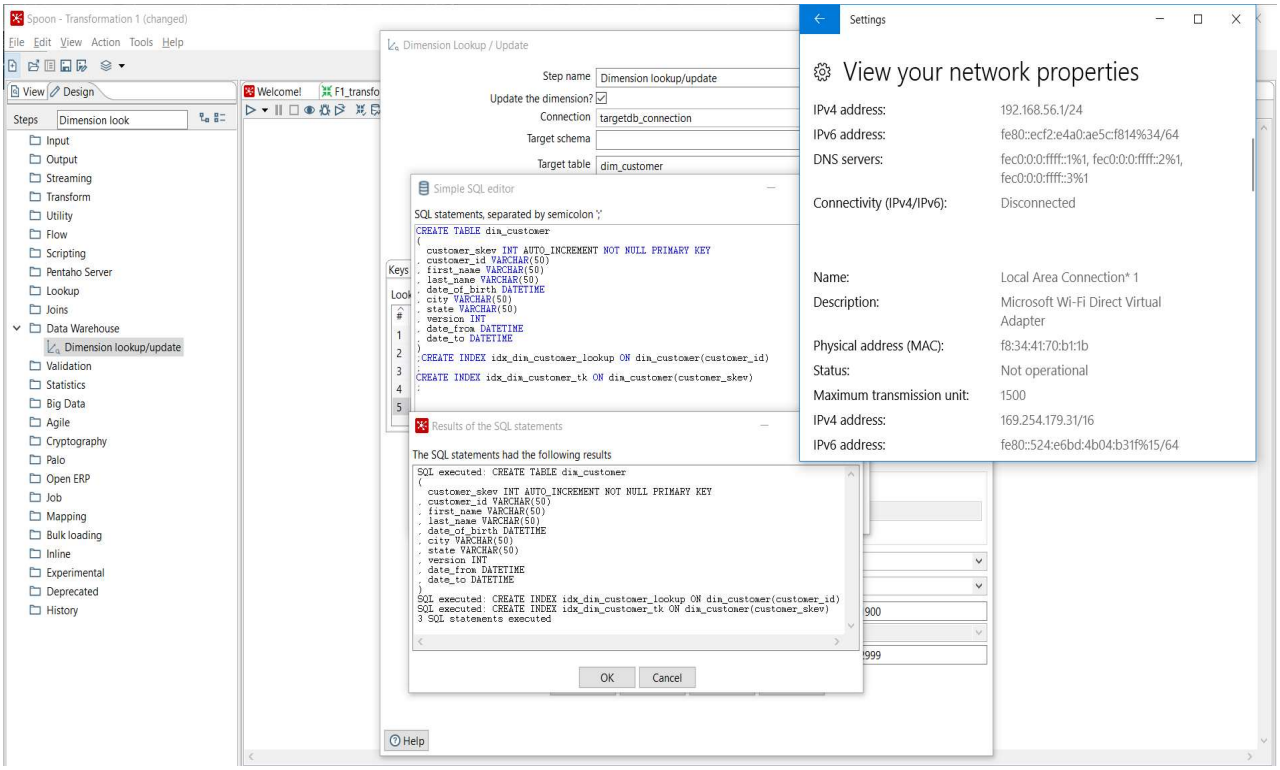
13. Now, I have **deleted** the last column **date_updated** from the lookup/update fields, as I have already defined that as Stream **Datefield** (unique field).



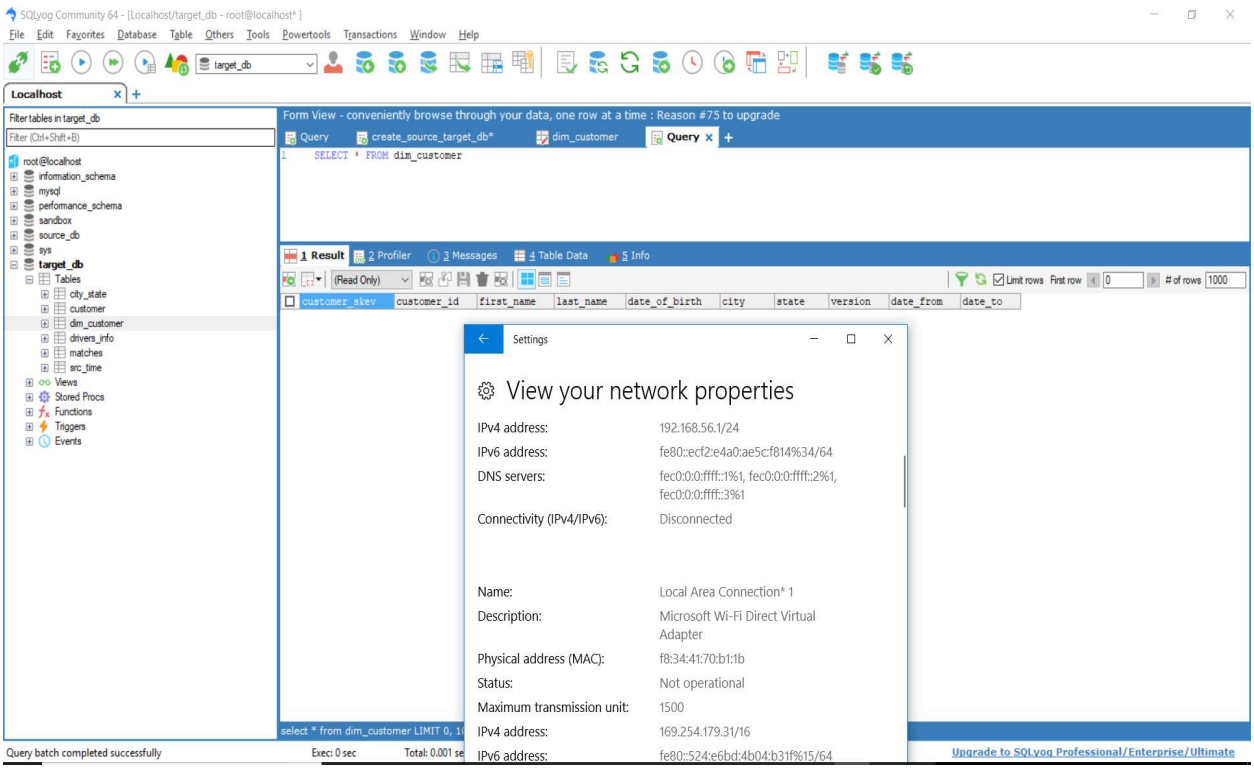
14. I have defined all the columns as type |(Update).



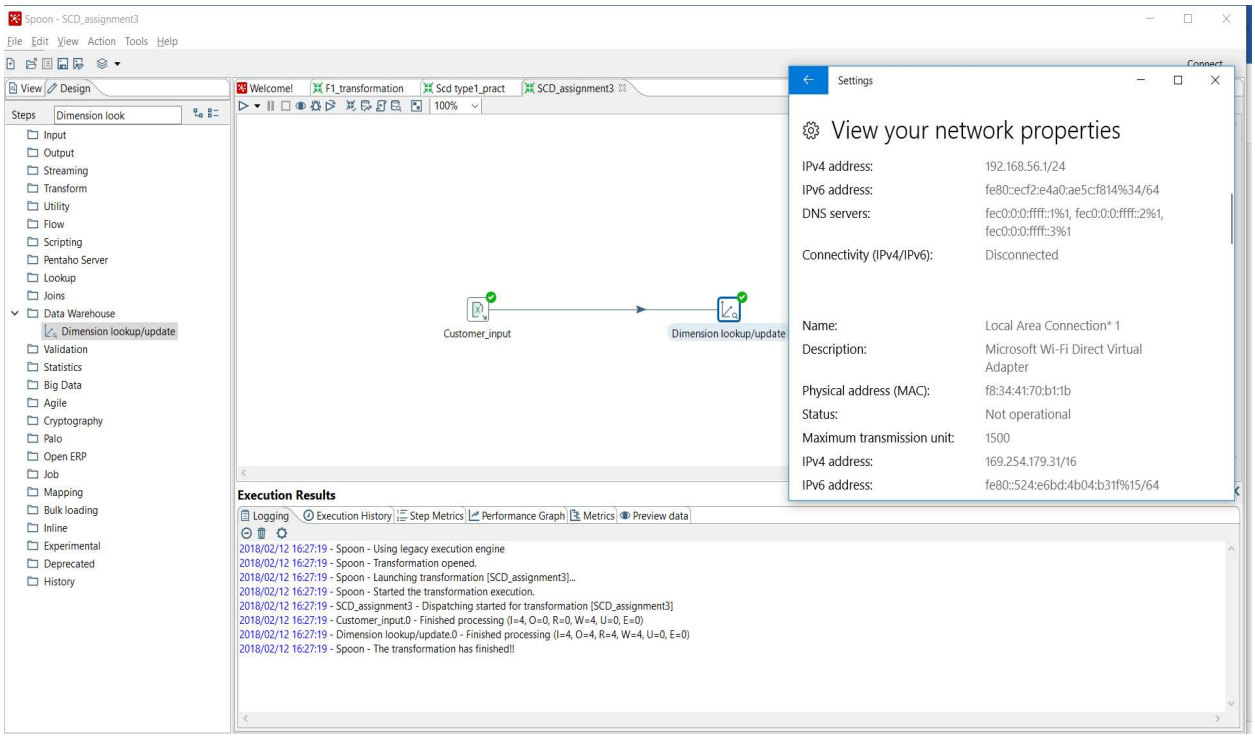
15. Now I have created DDL creation of dimensional table by clicking SQL button and automatically we have DDL scripts.



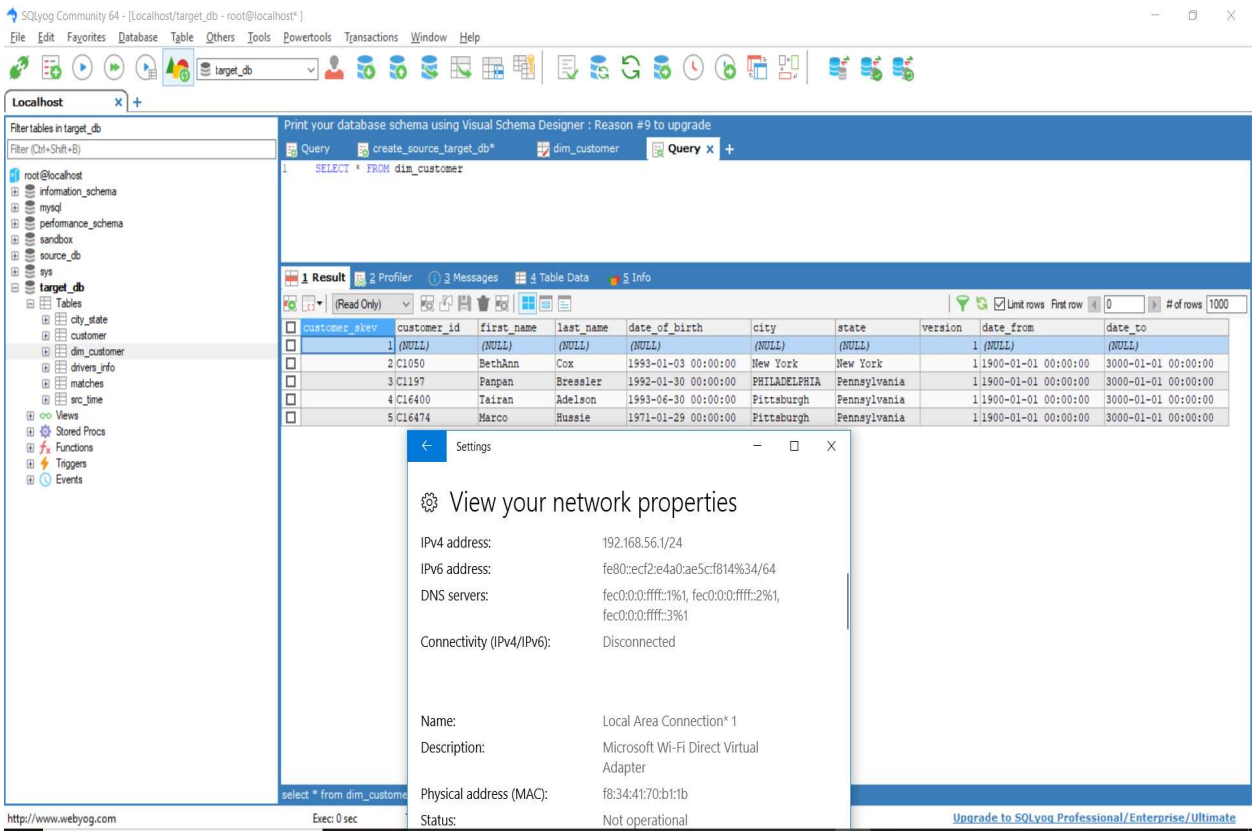
16. I checked the table created in our database, hence I can check this through SQLYOG.



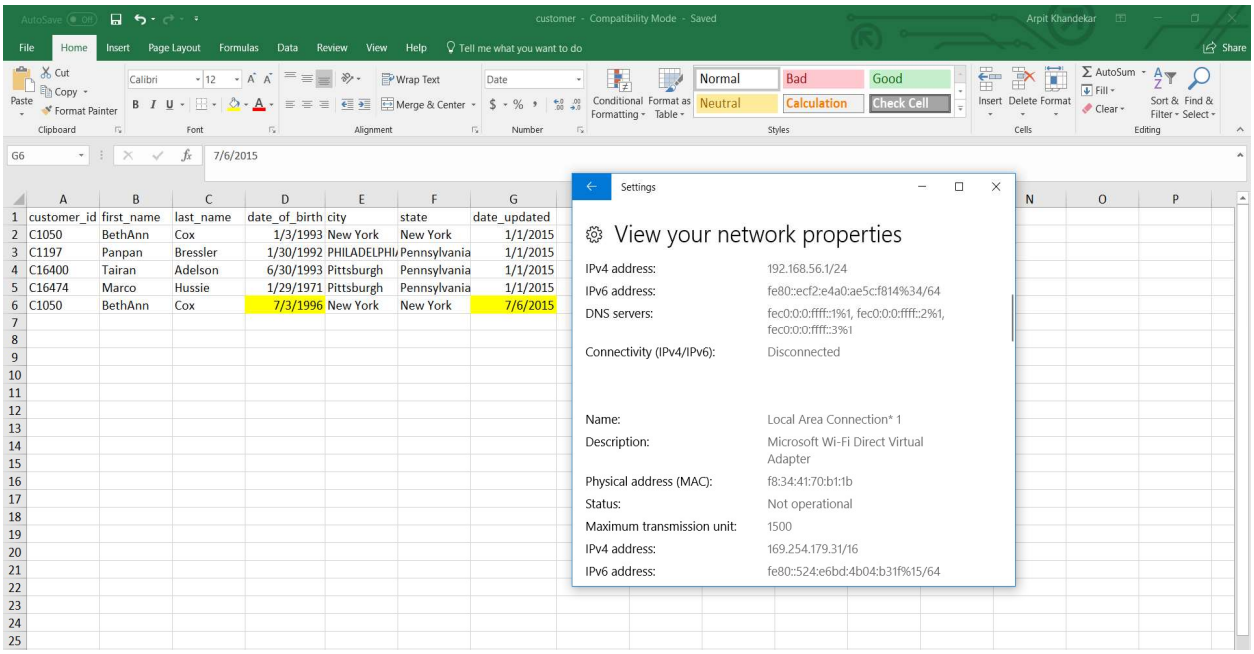
17. Now, I have run the transformation, it has run successfully.



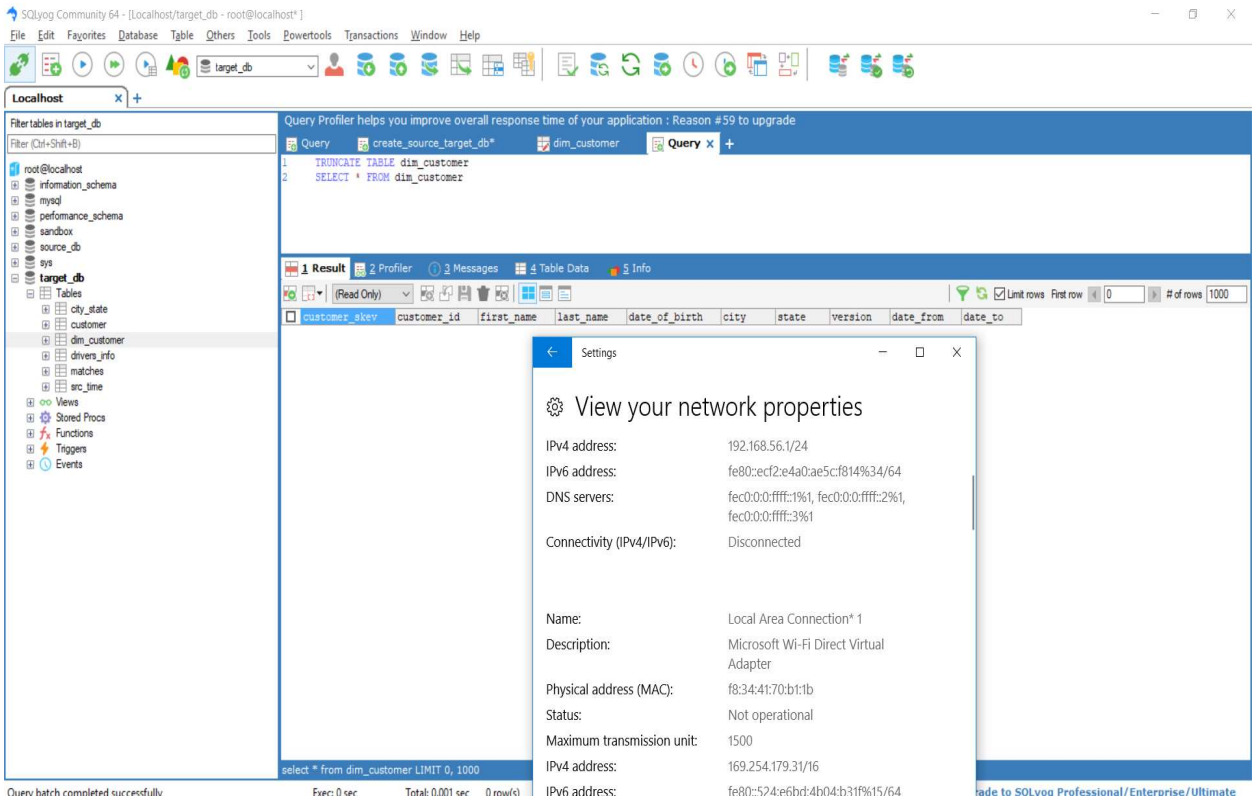
18. We can check in MYSQL DB by selecting all rows of the table, hence we found all rows are fetched by the below screenshot, hence the transformation is done successfully.



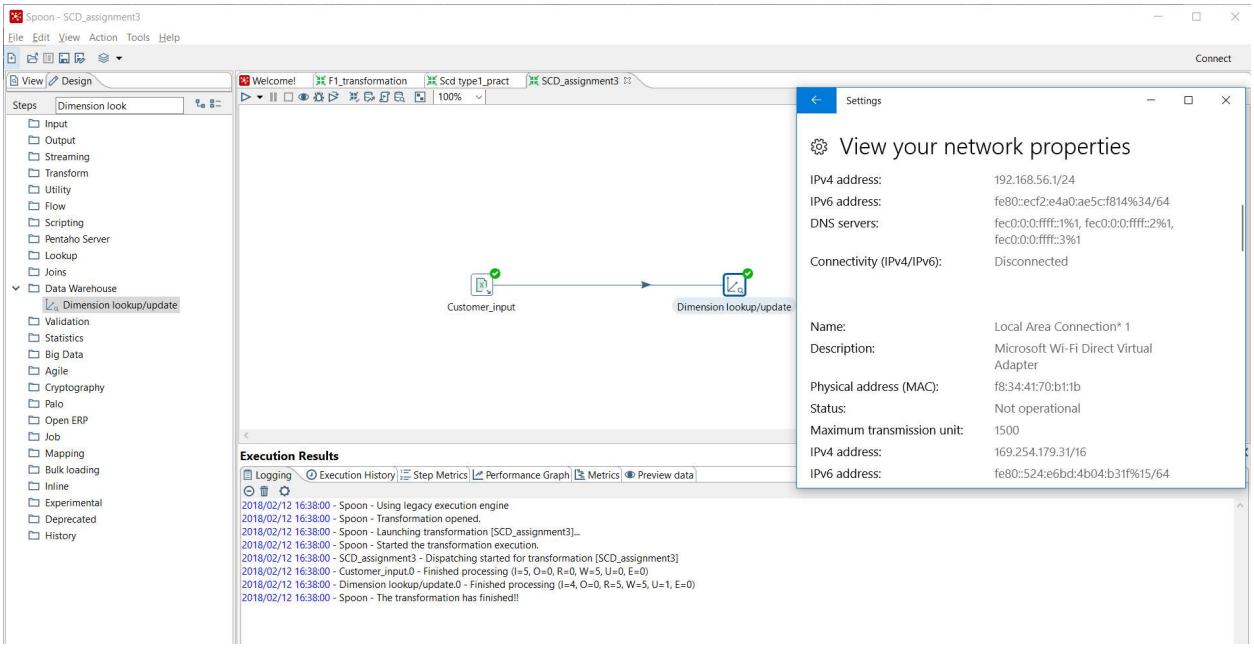
19. Now, we have updated records from the excel sheet, for eg we have updated date_of_birth and date_updated.



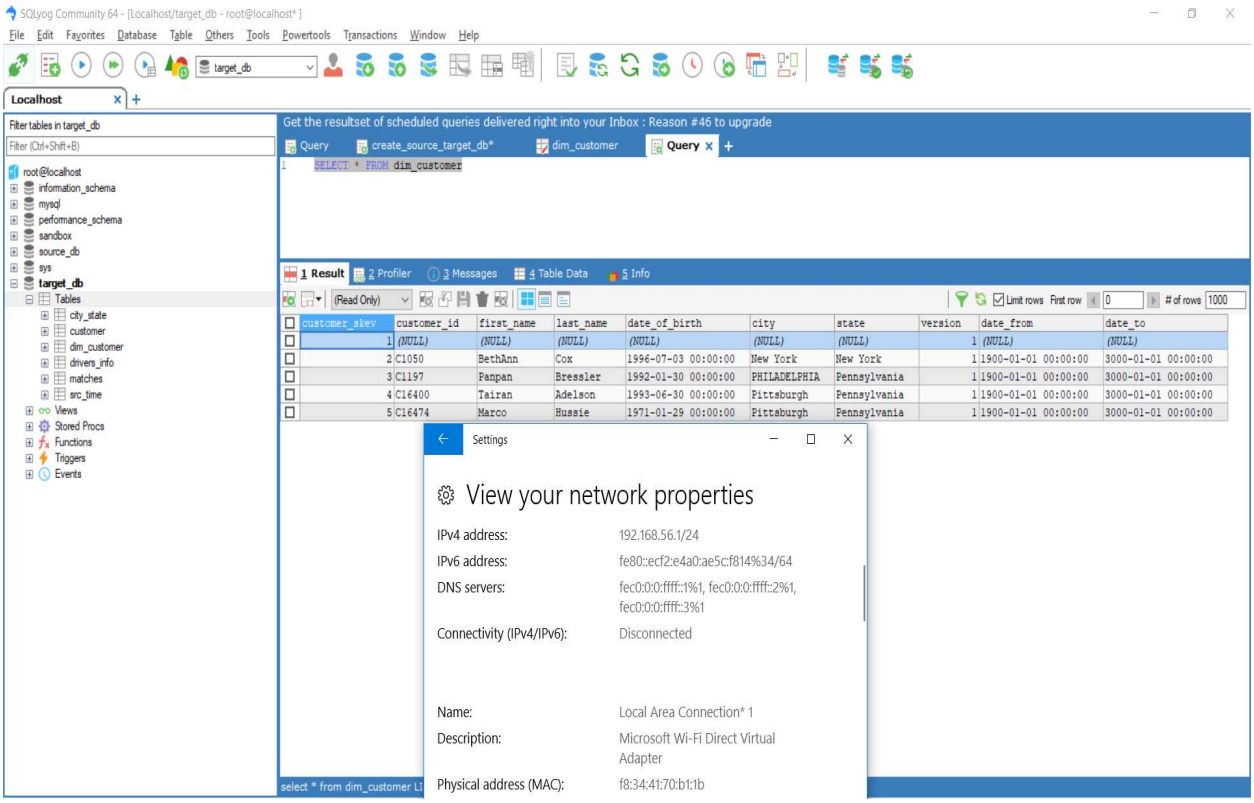
20. Before running transformation, we have truncated the table.



21. Now again I will run the transformation.

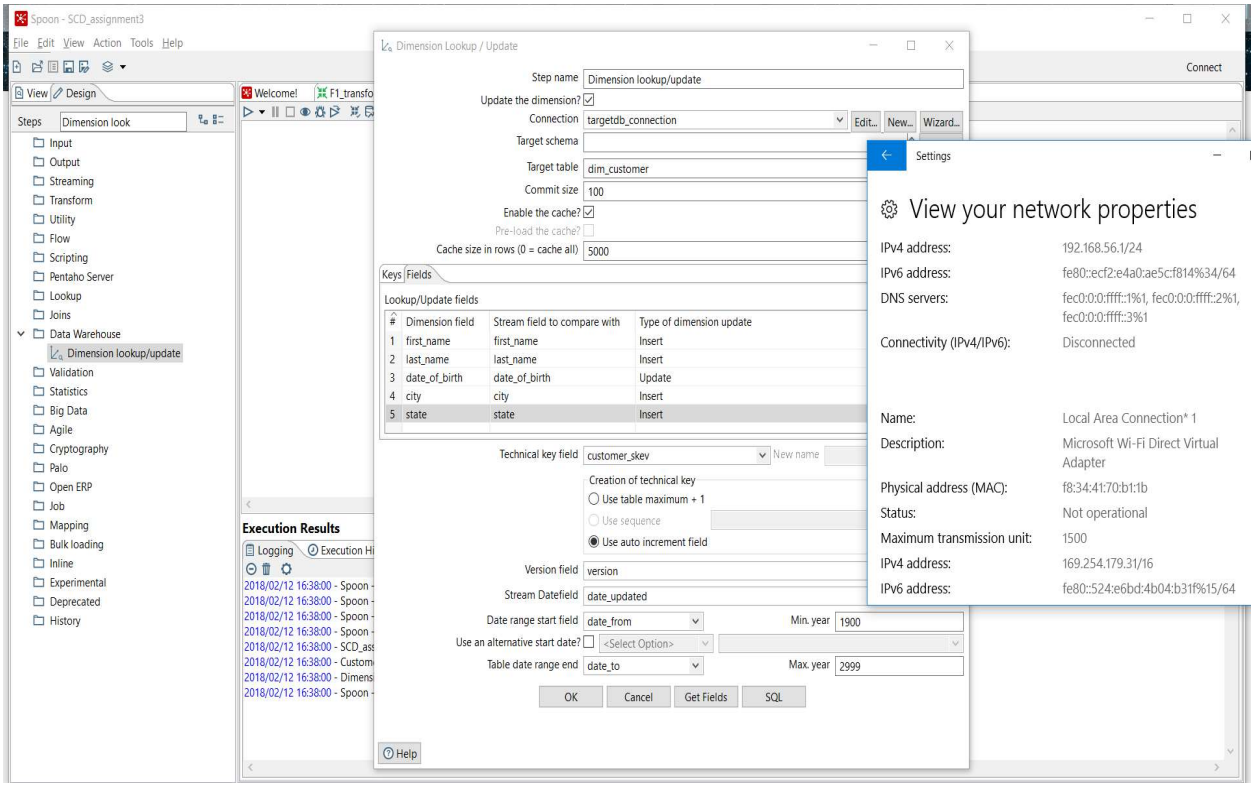


22. Transformation ran successfully, and now I check the entries in the database, I found that the birthdate column has been updated.

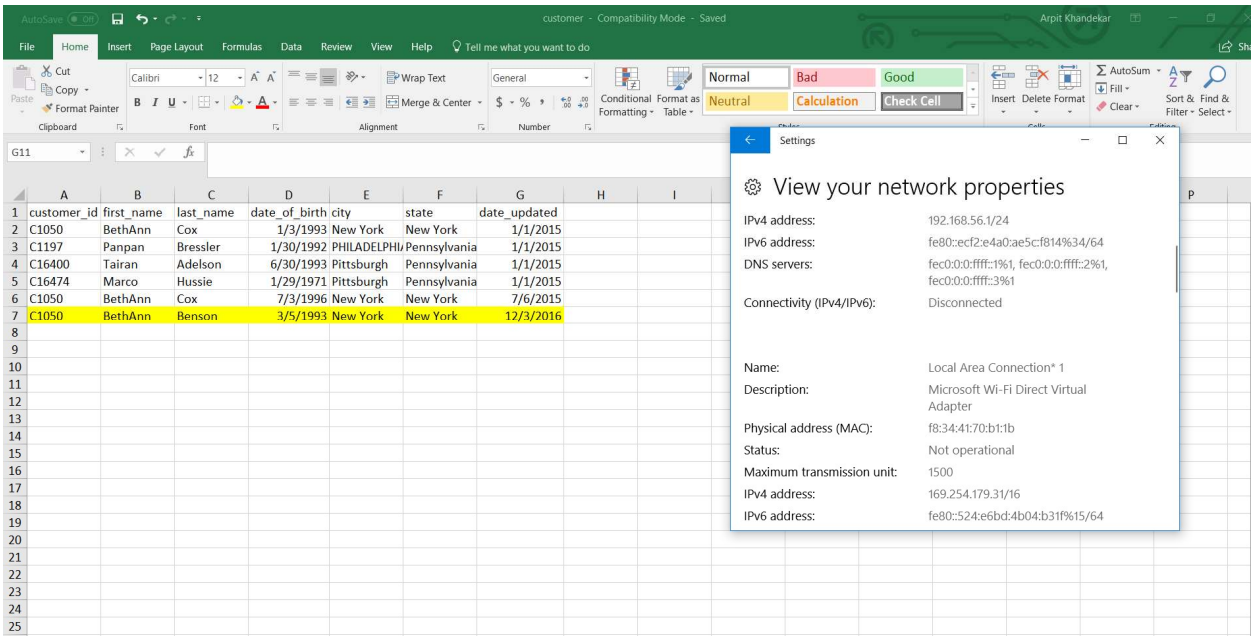


2. Slowly Changing Dimension Type ||(Insert) Transformation

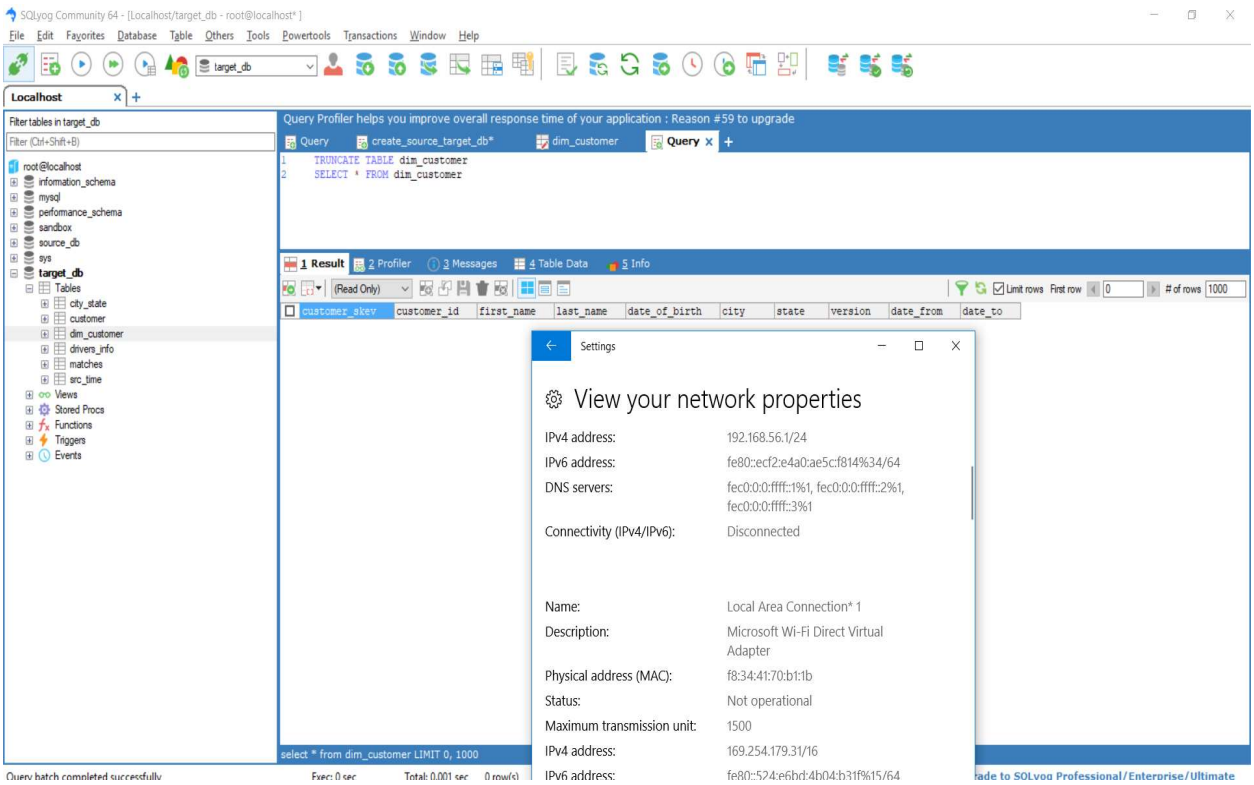
1. For Type || Transformation, we will change the type of dimension update as **Update.** So we will change the type of dimension for all the fields except **date_of_birth**.



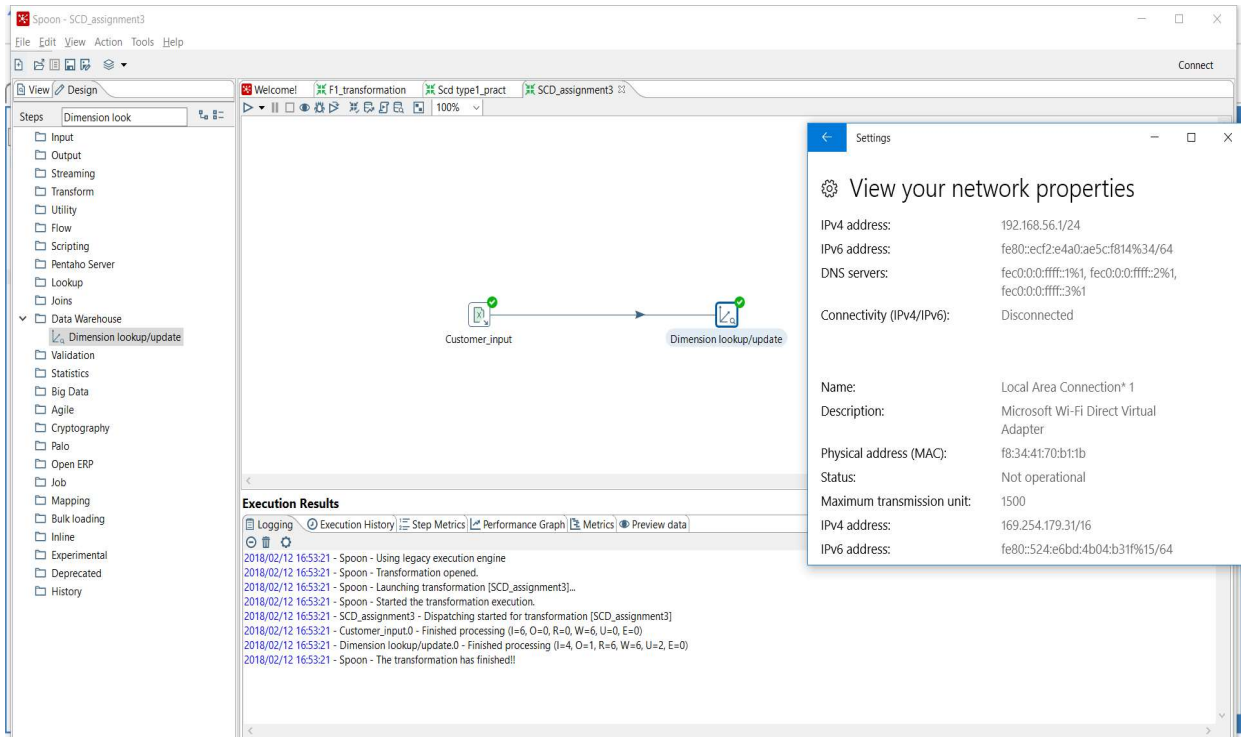
2. In the excel sheet I have insert one row for same customer '**BethAnn'** with same customerID '**C1050'**



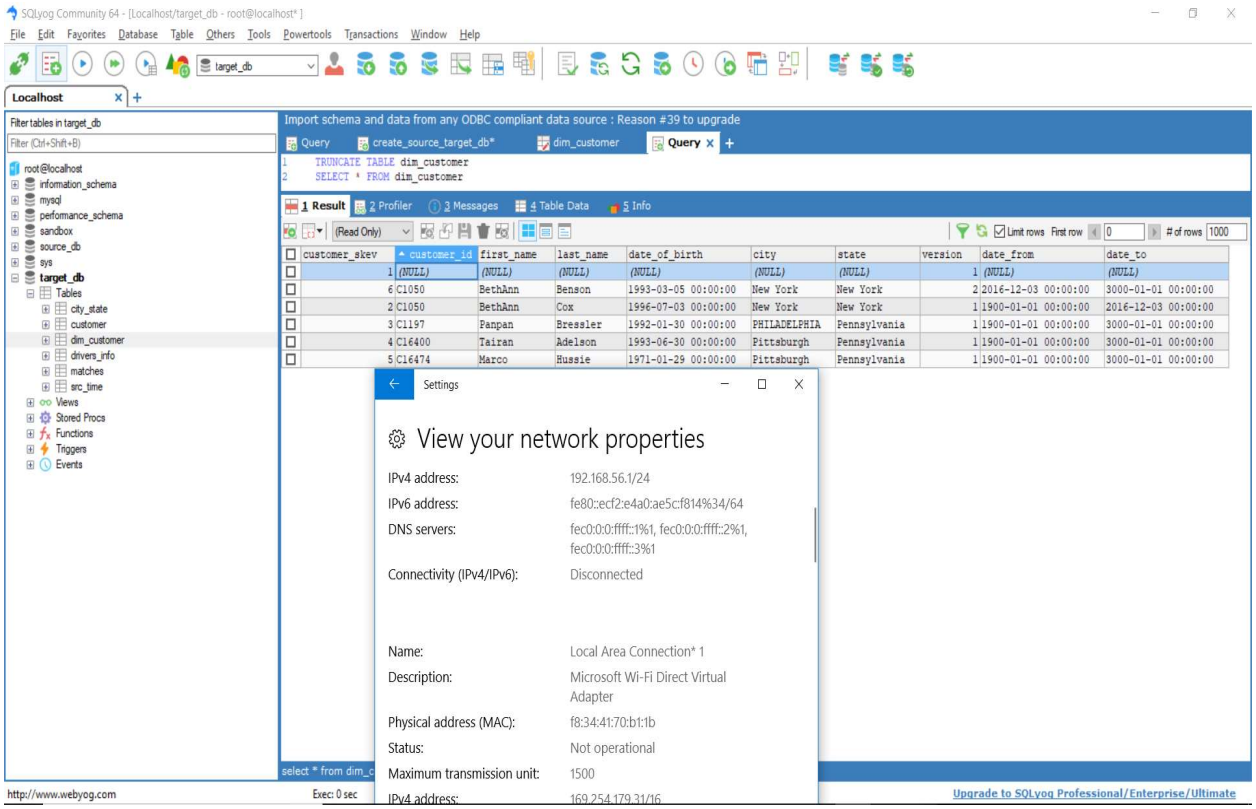
3. Before running transformation, we have truncated the table.



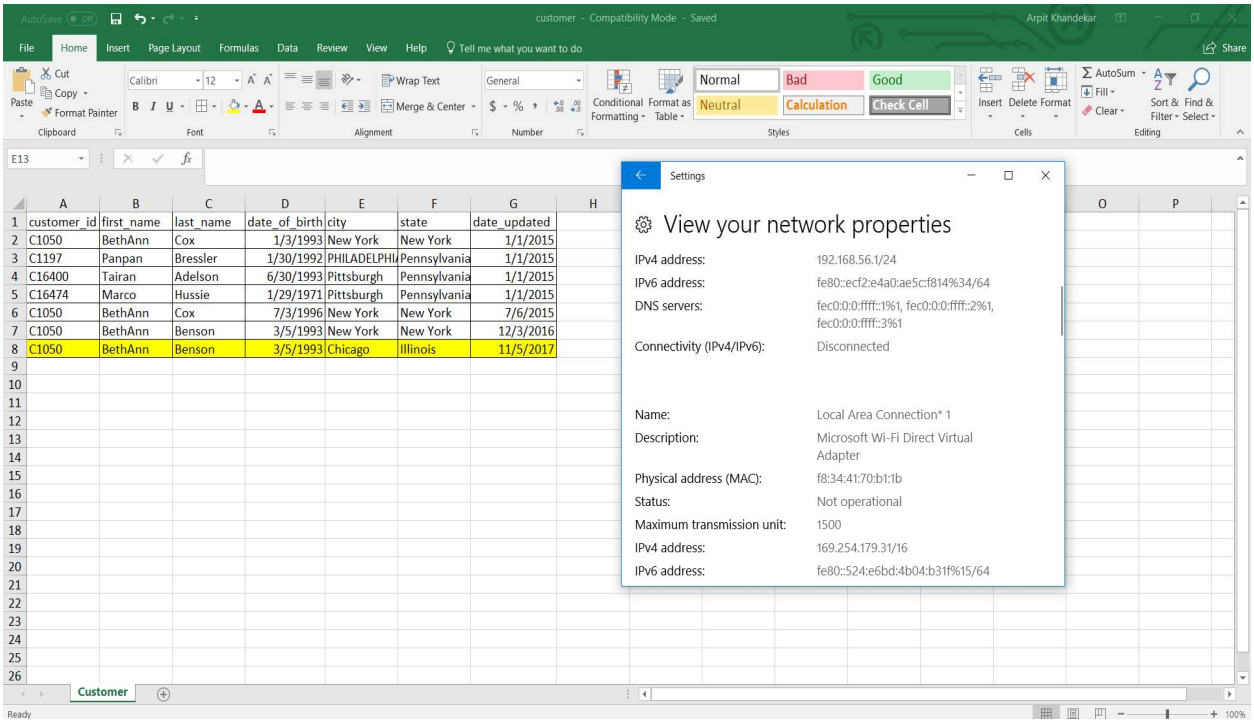
4. We will run the transformation again, we found that it has ran successfully.



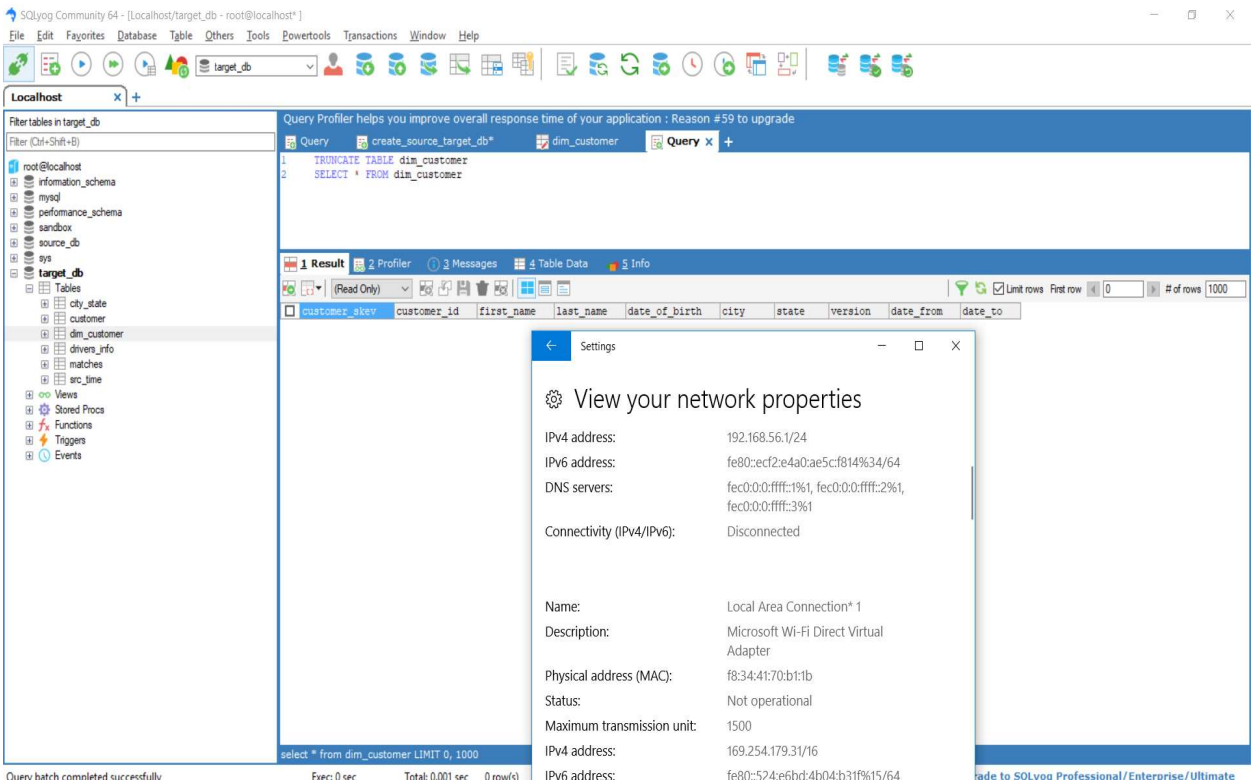
5. We can check the new record inserted in database through SQLYOG, hence we see that record is inserted successfully, also version is incremented by 1 and customer_skev(surrogate key) is incremented by 1. Hence, we can say that BethAnn lastname **Cox** will be upto 2016-12-03 and thereafter it changes to **Benson.**



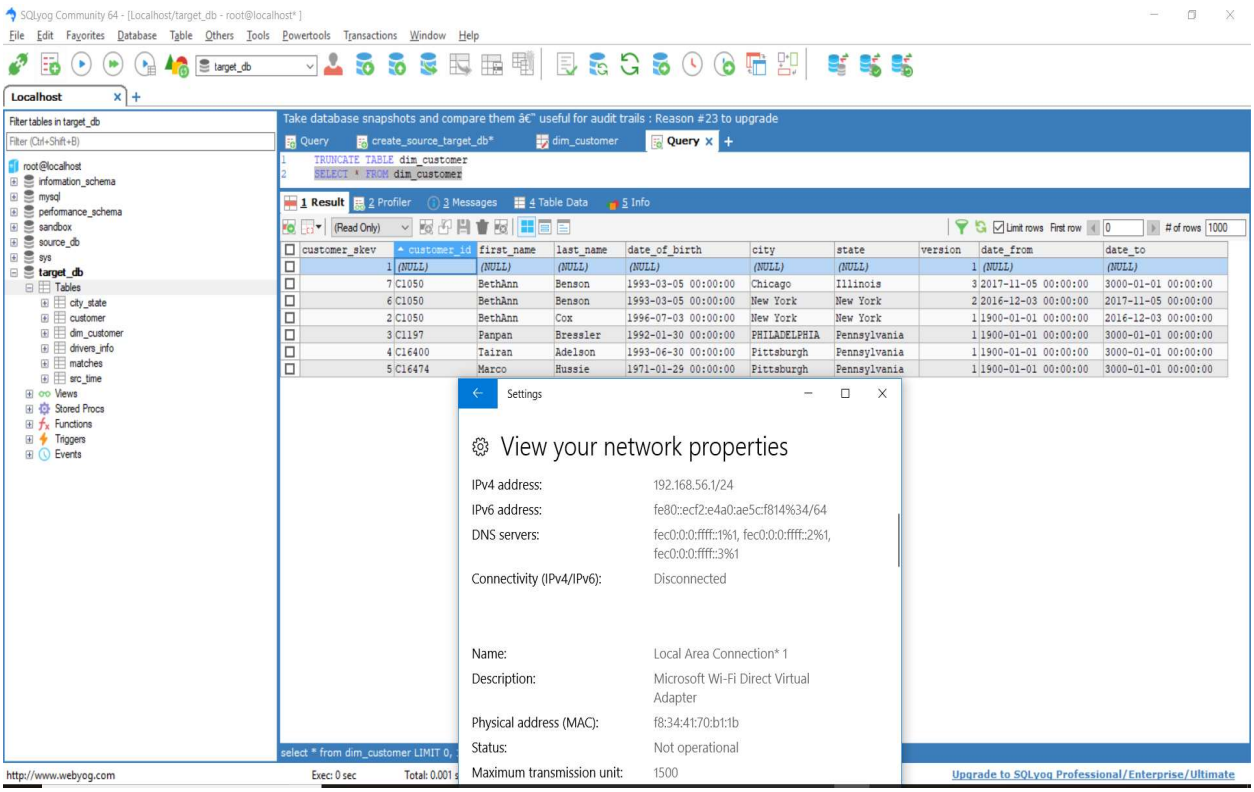
6. Now, we have updated BethAnn address to Chicago Illinois from particular date in excel sheet, hence we will run the transformation and lets check.



7. Before running transformation, we have truncated the table.



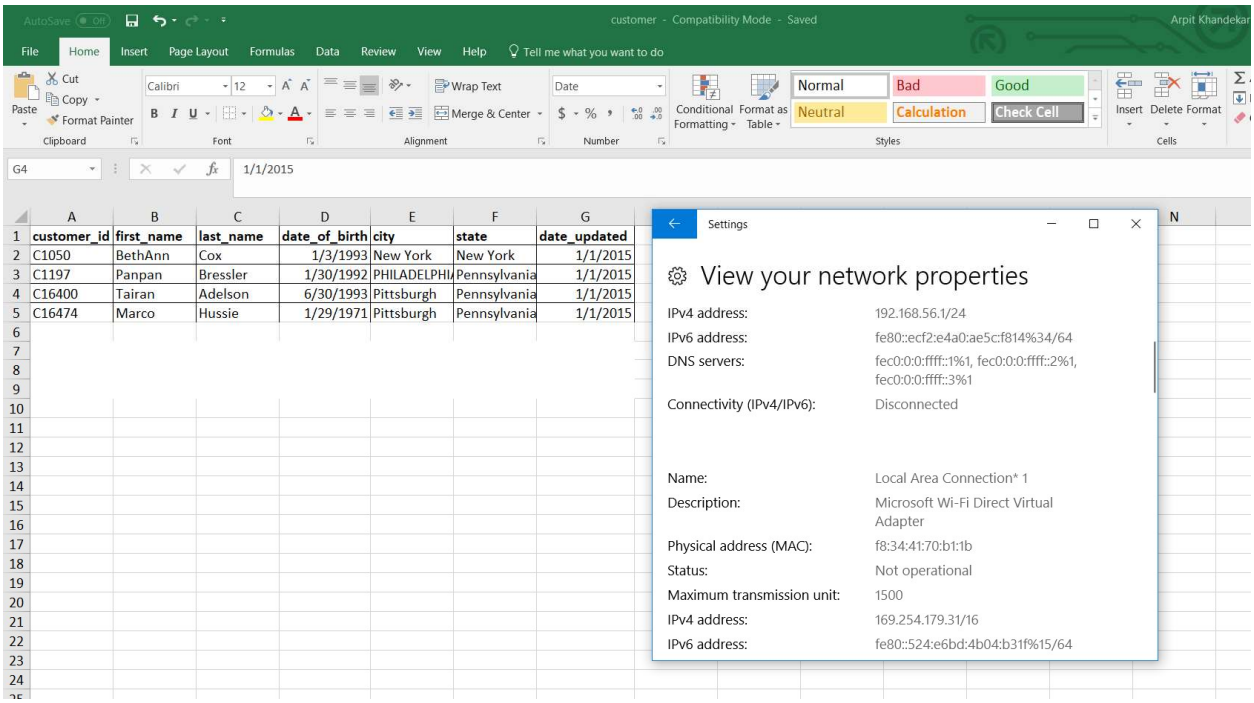
8. Hence, we see that address of customer BethAnn is updated from **‘2015-11-05’**
We can see from the below screenshot.



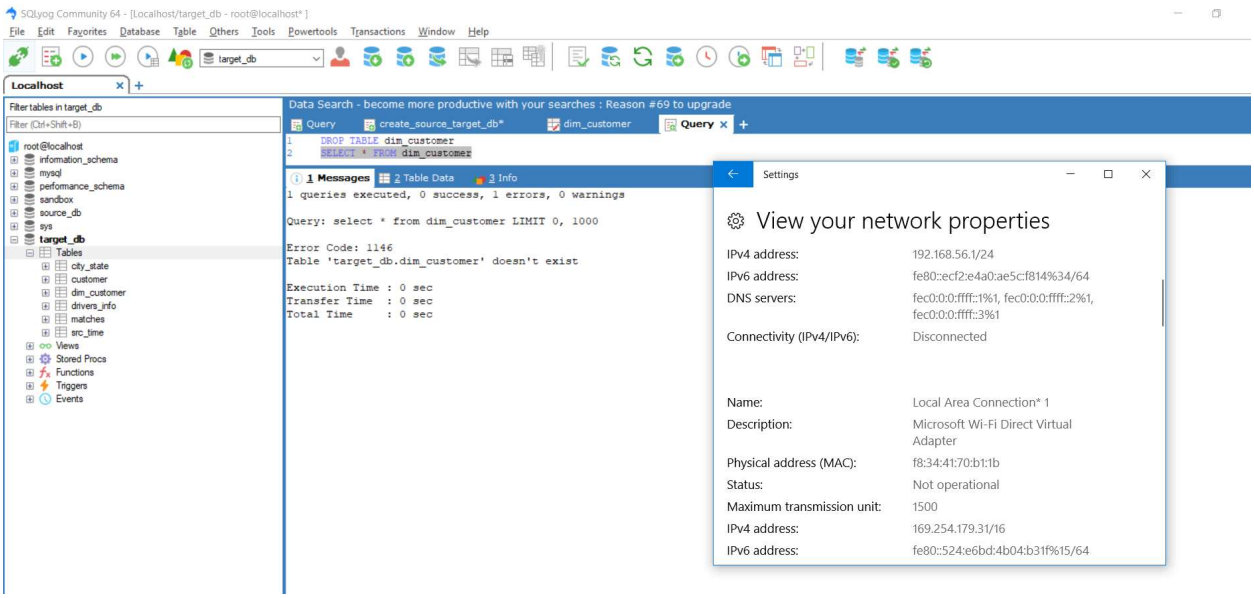
Thus, it successfully shows slowly changing **dimensional model type ||.**

3. Slowly Changing Dimension Type | (Push Through) Transformation

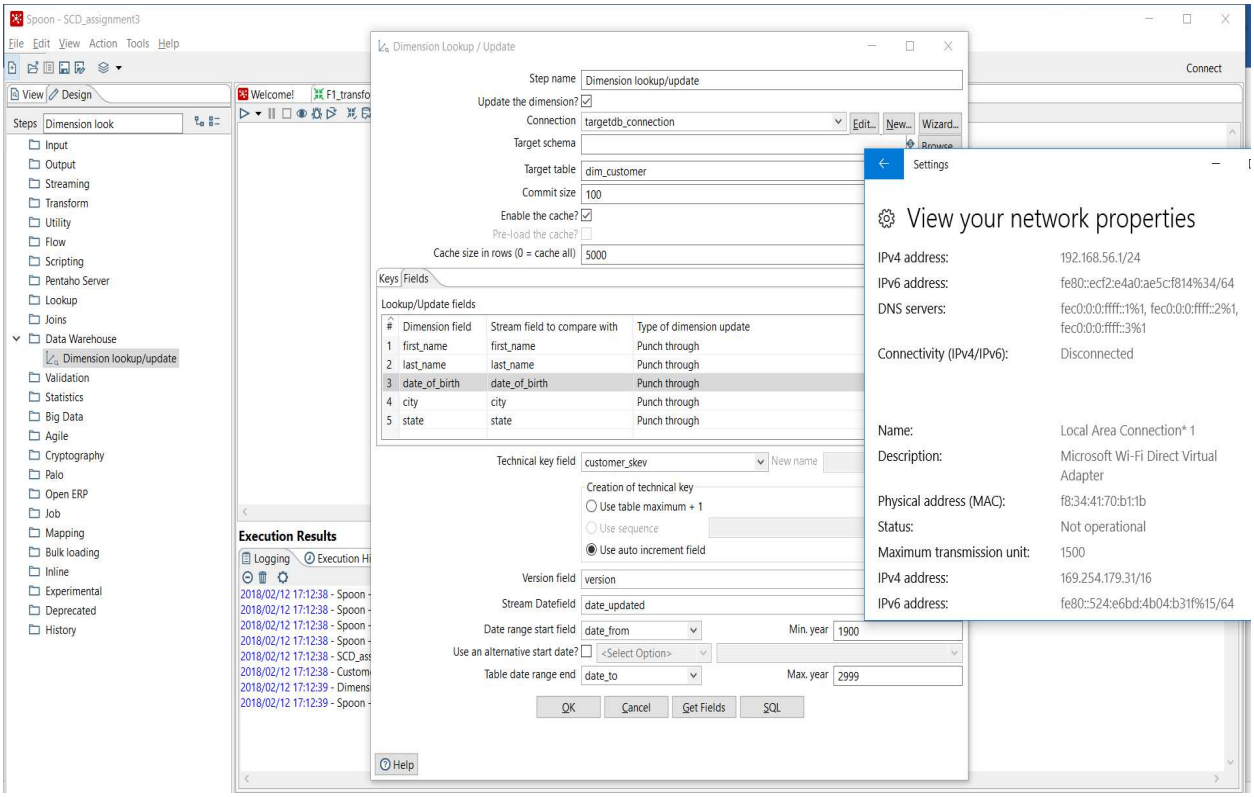
1. We will take excel input and dimension lookup/update.



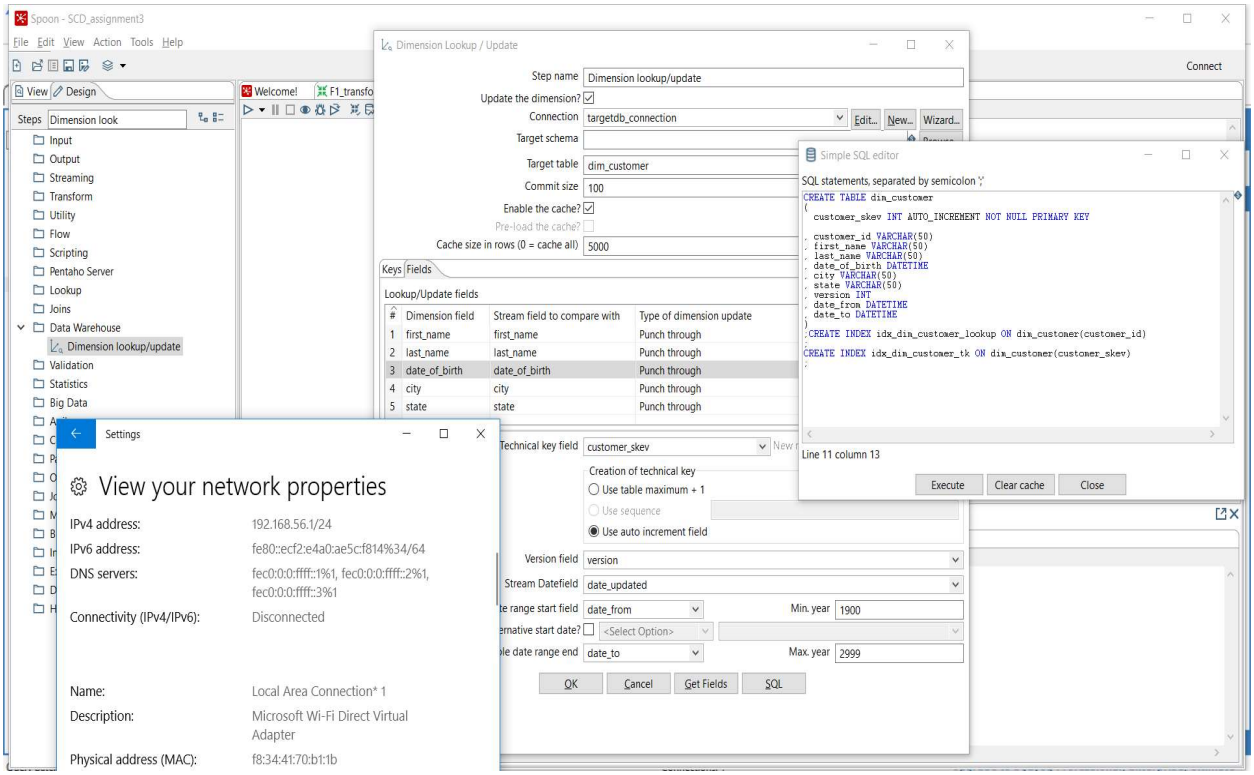
2. Before running transformation, we have dropped the table from the databse.



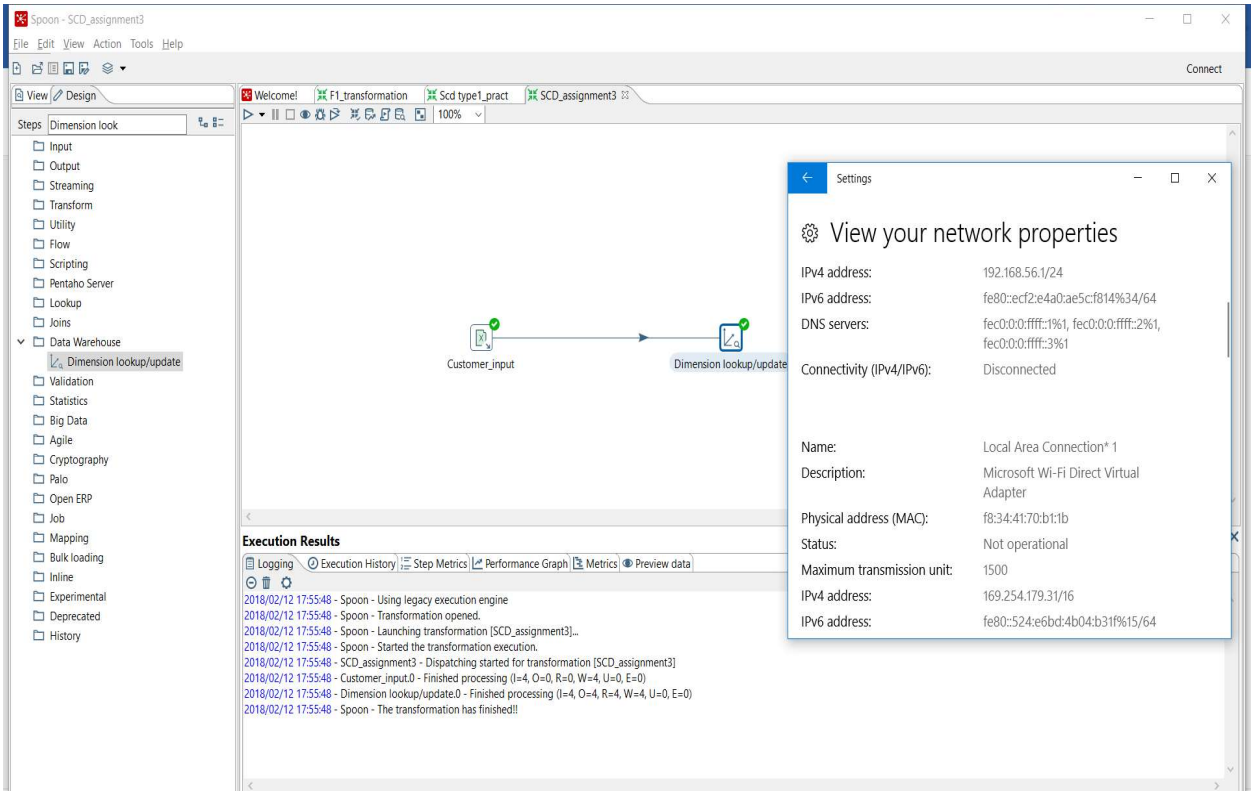
3. Now we change the type of dimension update filed to **‘Punch through’.**



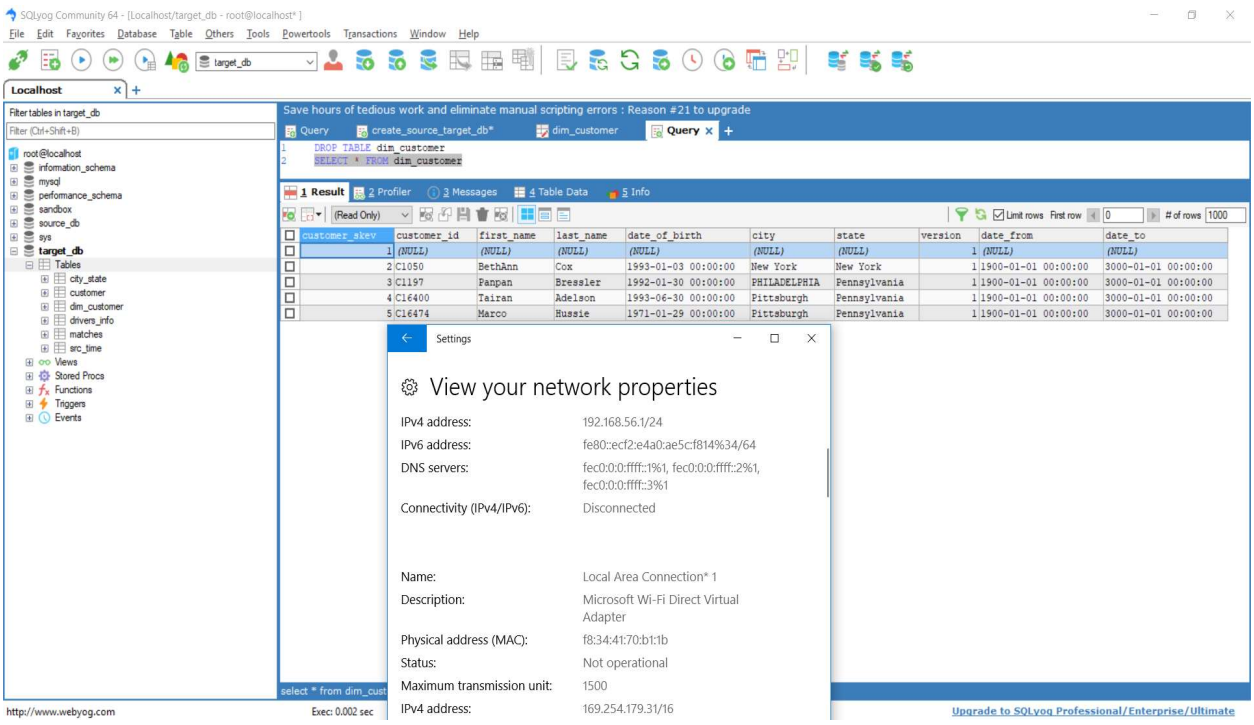
4. Create SQL query to create a table.



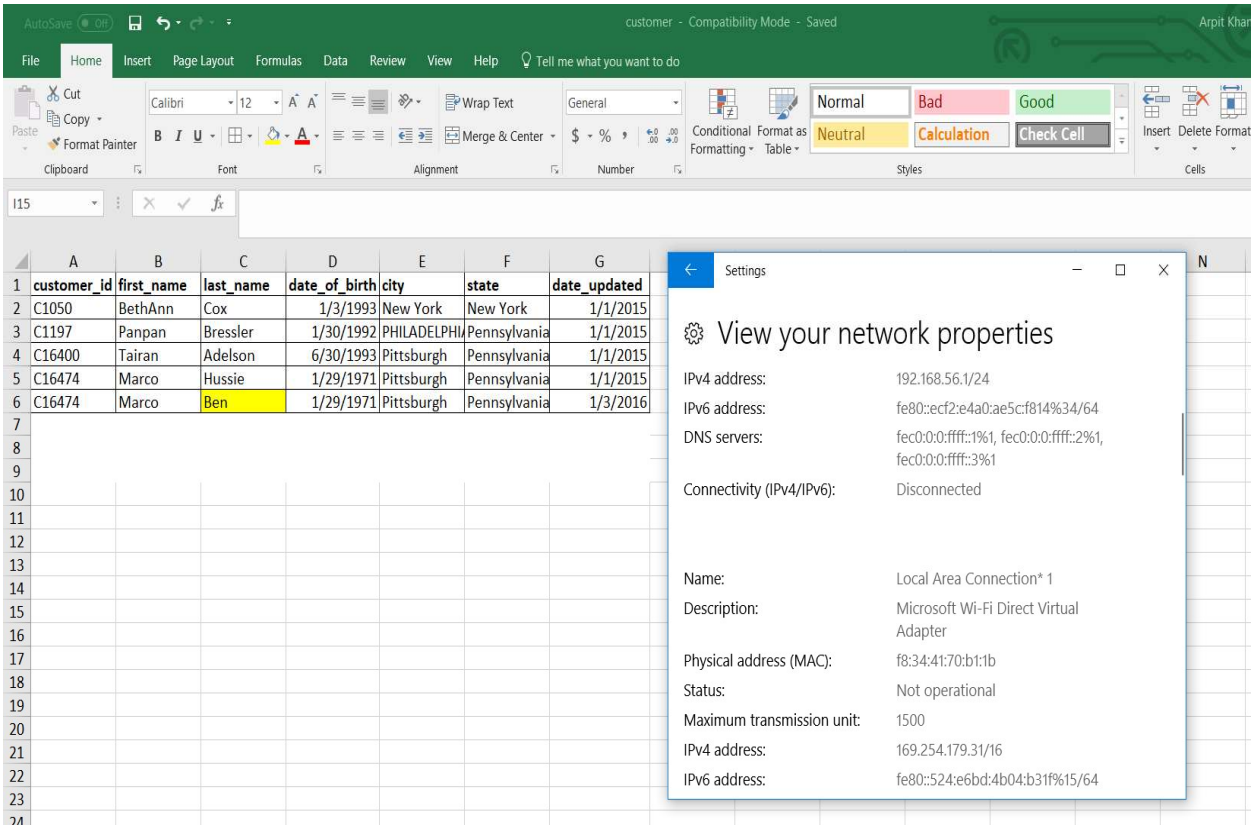
5. Run the transformation, it has run successfully.



6. Now we can see the changes in database mysql, hence from below screenshot we can see that all the records are inserted successfully.

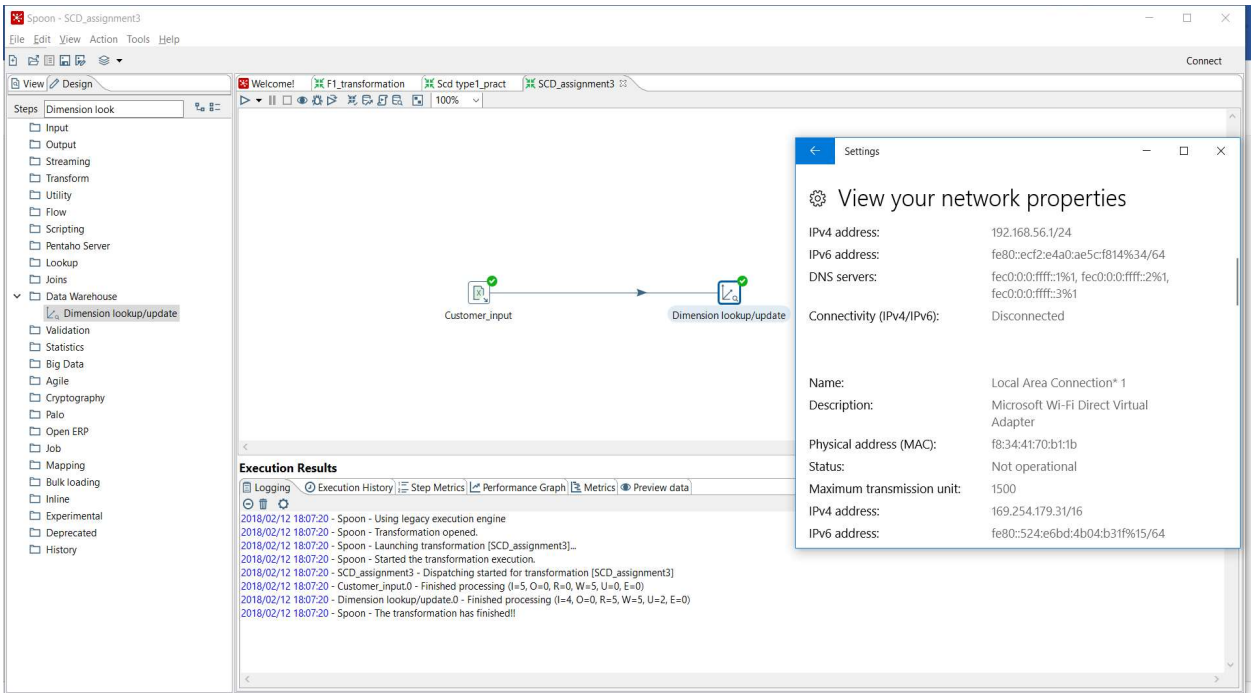


7. Now, if I update any record then it can be updated and can be seen the database. For example if we want to update Customer last_ name for Macro to Ben.

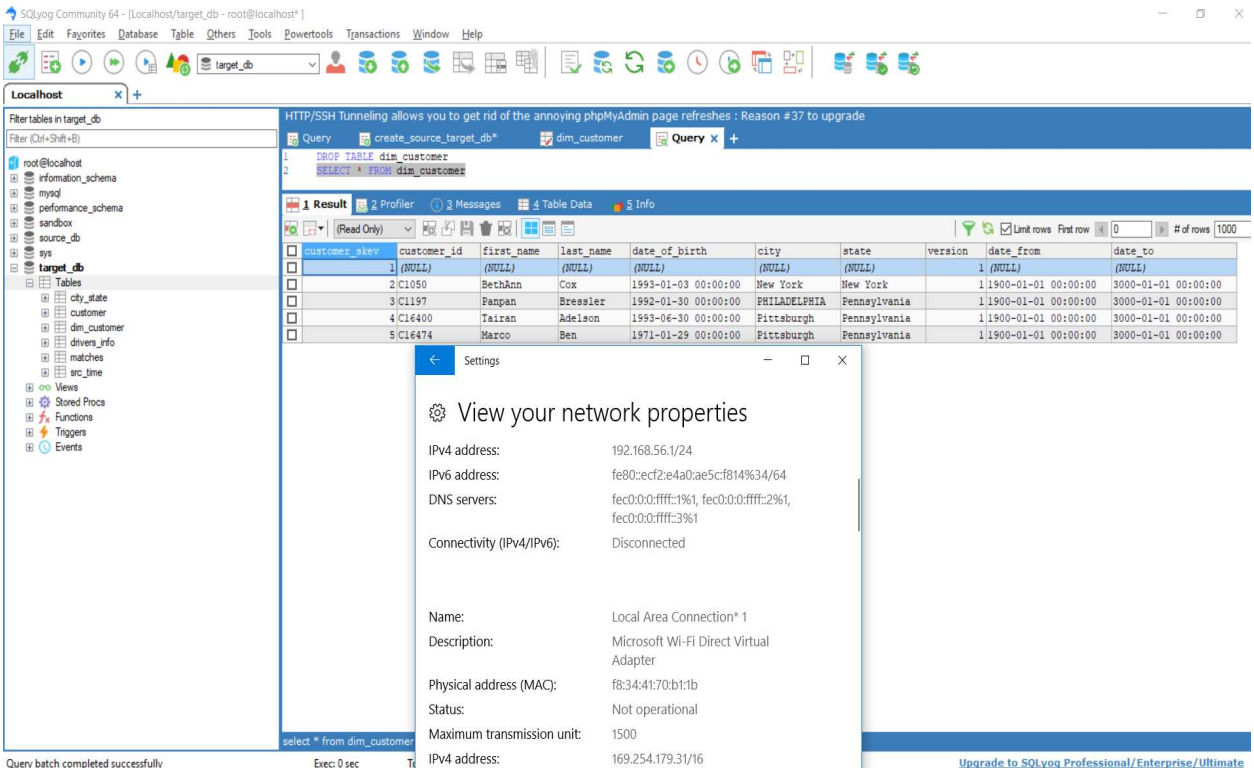


customer_id	first_name	last_name	date_of_birth	city	state	date_updated
C1050	BethAnn	Cox	1/3/1993	New York	New York	1/1/2015
C1197	Panpan	Bressler	1/30/1992	PHILADELPHIA	Pennsylvania	1/1/2015
C16400	Tairan	Adelson	6/30/1993	Pittsburgh	Pennsylvania	1/1/2015
C16474	Marco	Hussie	1/29/1971	Pittsburgh	Pennsylvania	1/1/2015
C16474	Marco	Ben	1/29/1971	Pittsburgh	Pennsylvania	1/3/2016

8 Now we will the run the transformation.



9 We can see that transformation has run successfully, hence changes will be reflected in Database Mysql.



Hence, we saw that Macro last name has changed to **Ben**. Thus, we have perform punch through operations successfully in Pentaho.