**Data Warehouse and**

**Business Intelligence**

INFO 7290

(SECTION-1)

Assignment-05

NAME: Ayush Mangal

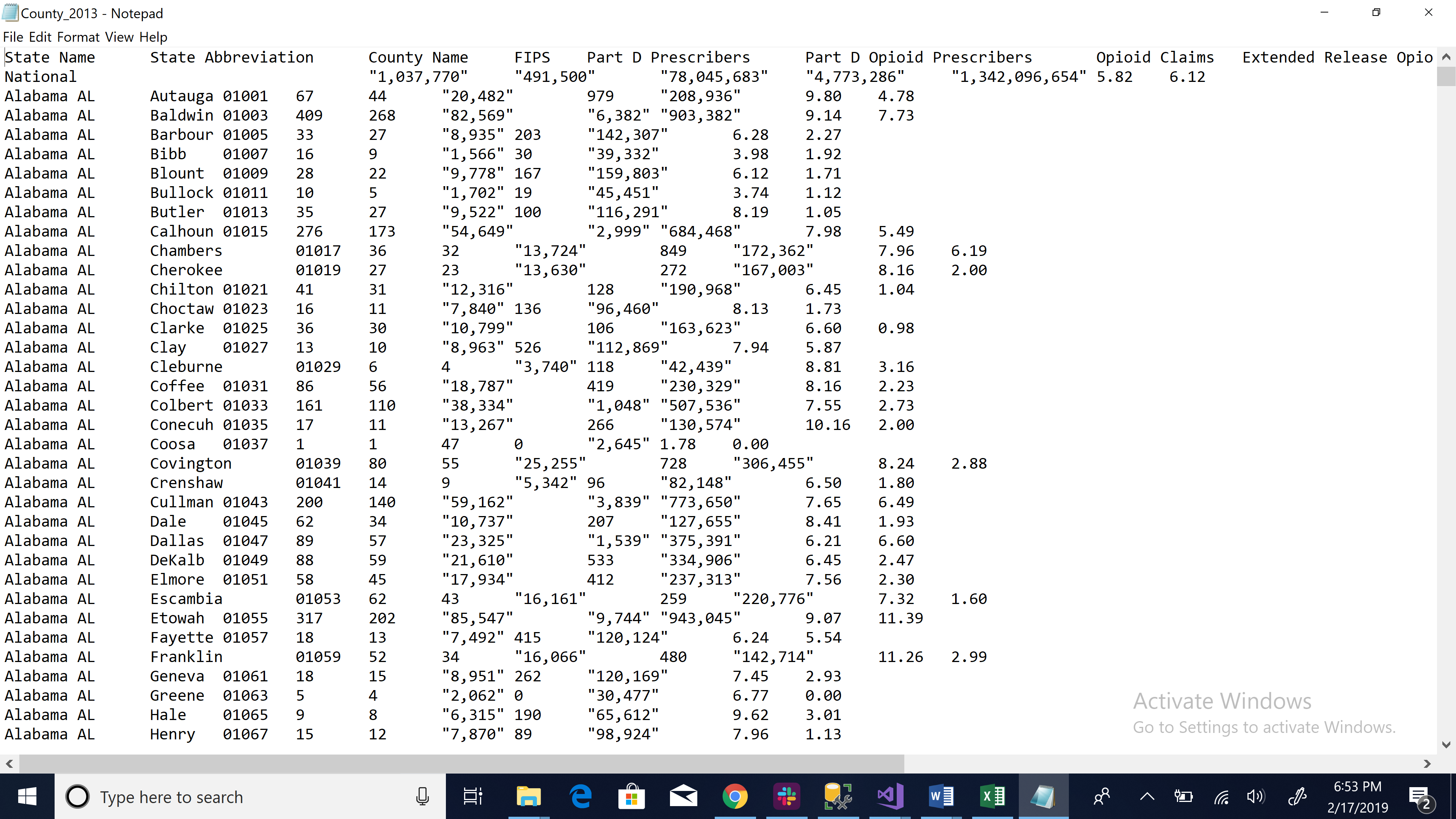
NUID: 001873174

**Solution 1:**

Following steps have been done for this solution:

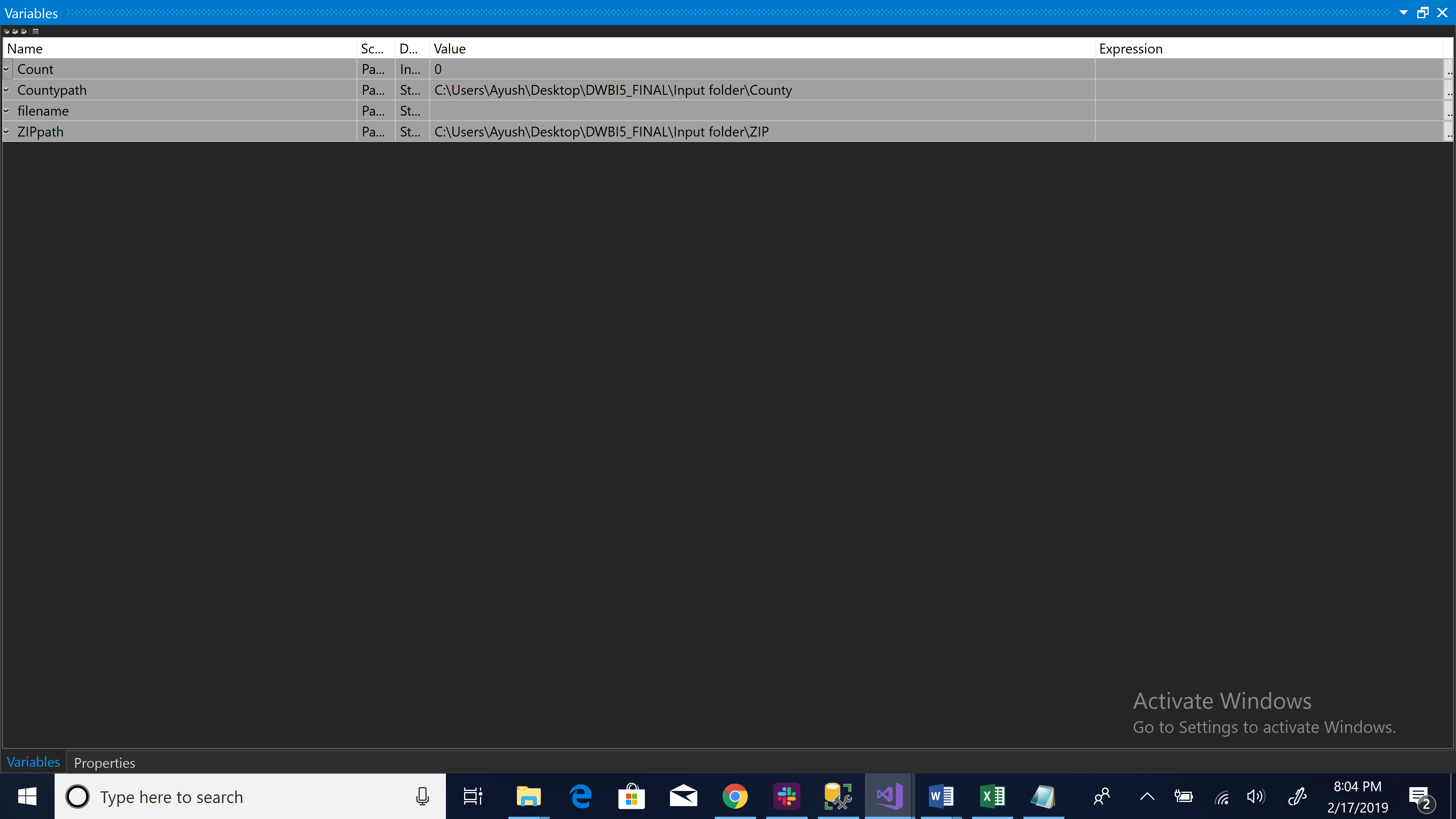
1. Files have been renamed to County\_2013, County\_2014,County\_2015,County\_2016,ZIP\_2013,ZIP\_2014,ZIP\_2015,ZIP\_2016 respectively.
2. Files have been edited to remove the extra items at the top of the tab but the data is formatted as is.
3. Files have been modified from xls to tab delimited.

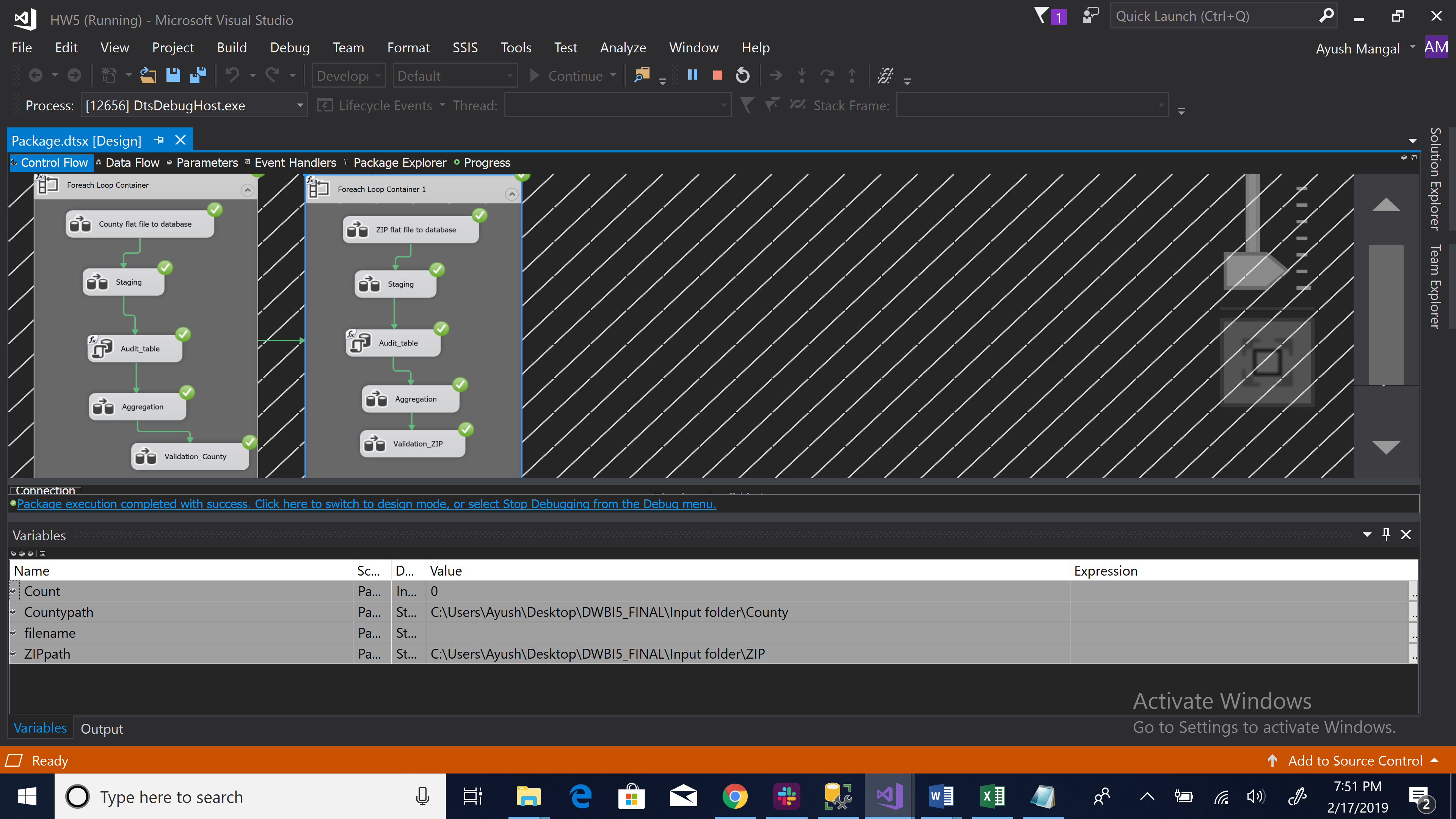




For the purpose of this package, we have to create the following variables in SSIS:

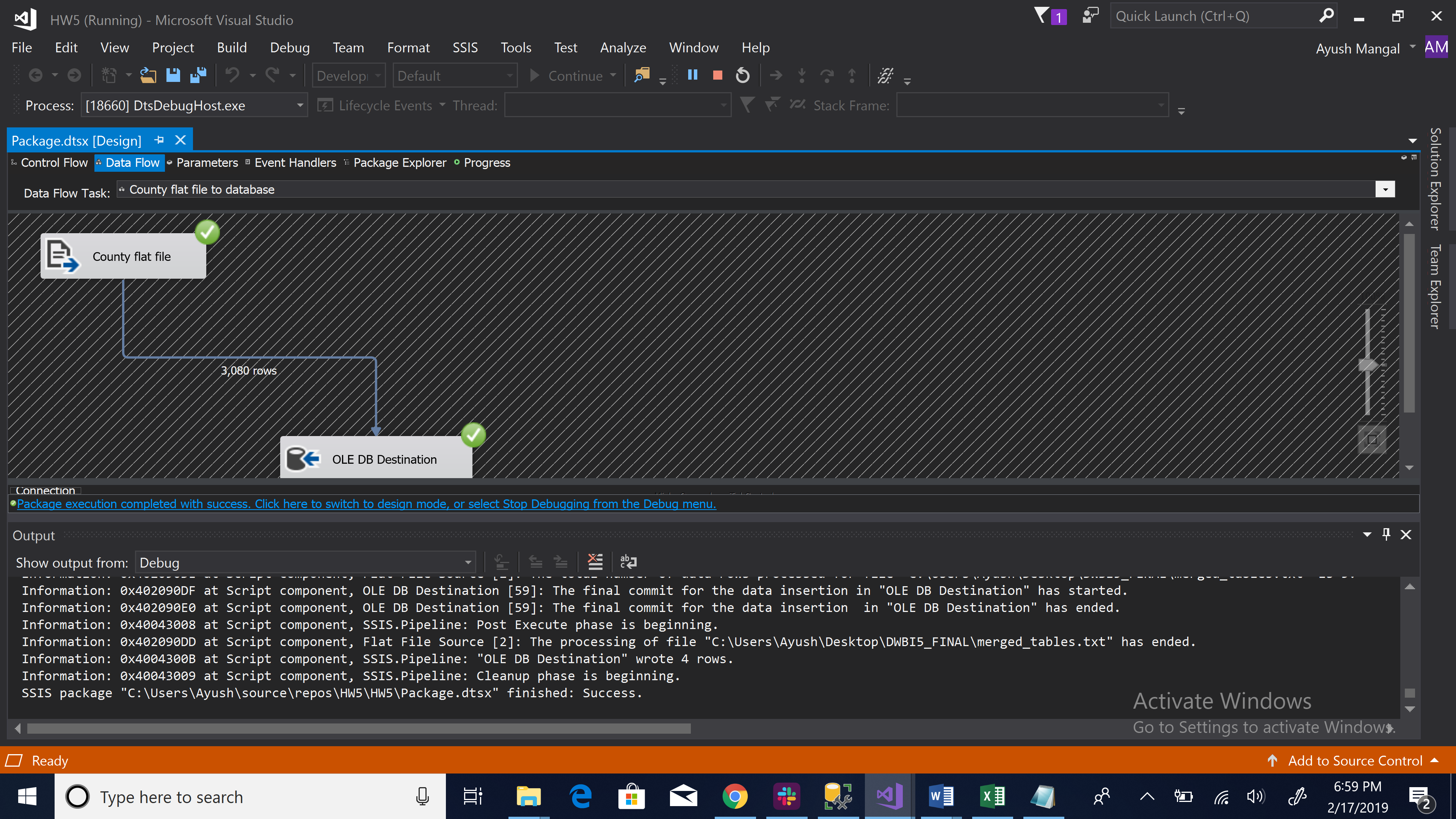
1. Countypath: Variable to get the path of the particular .txt files in County folder so that it can be extracted.
2. ZIPpath: Variable to get the path of the particular .txt file so that it can be extracted.
3. filename: Variable to get the .txt files from the Input folder where all the .txt files have been saved.
4. Count: Variable to count the rows of a particular file in the audit table task.





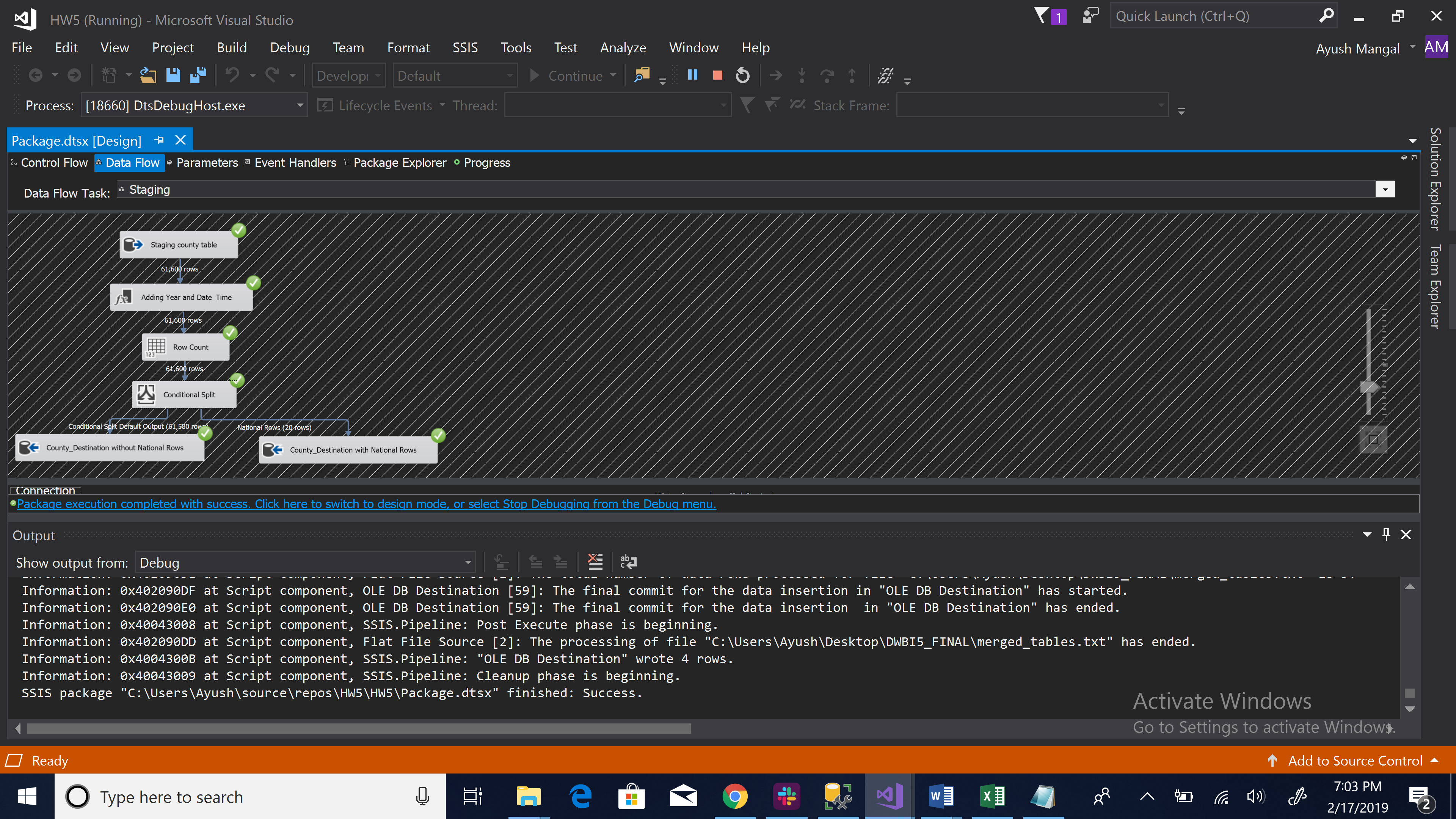
The above package comprises of the following steps:

Step 1: Copying flat file to the SQL server database:



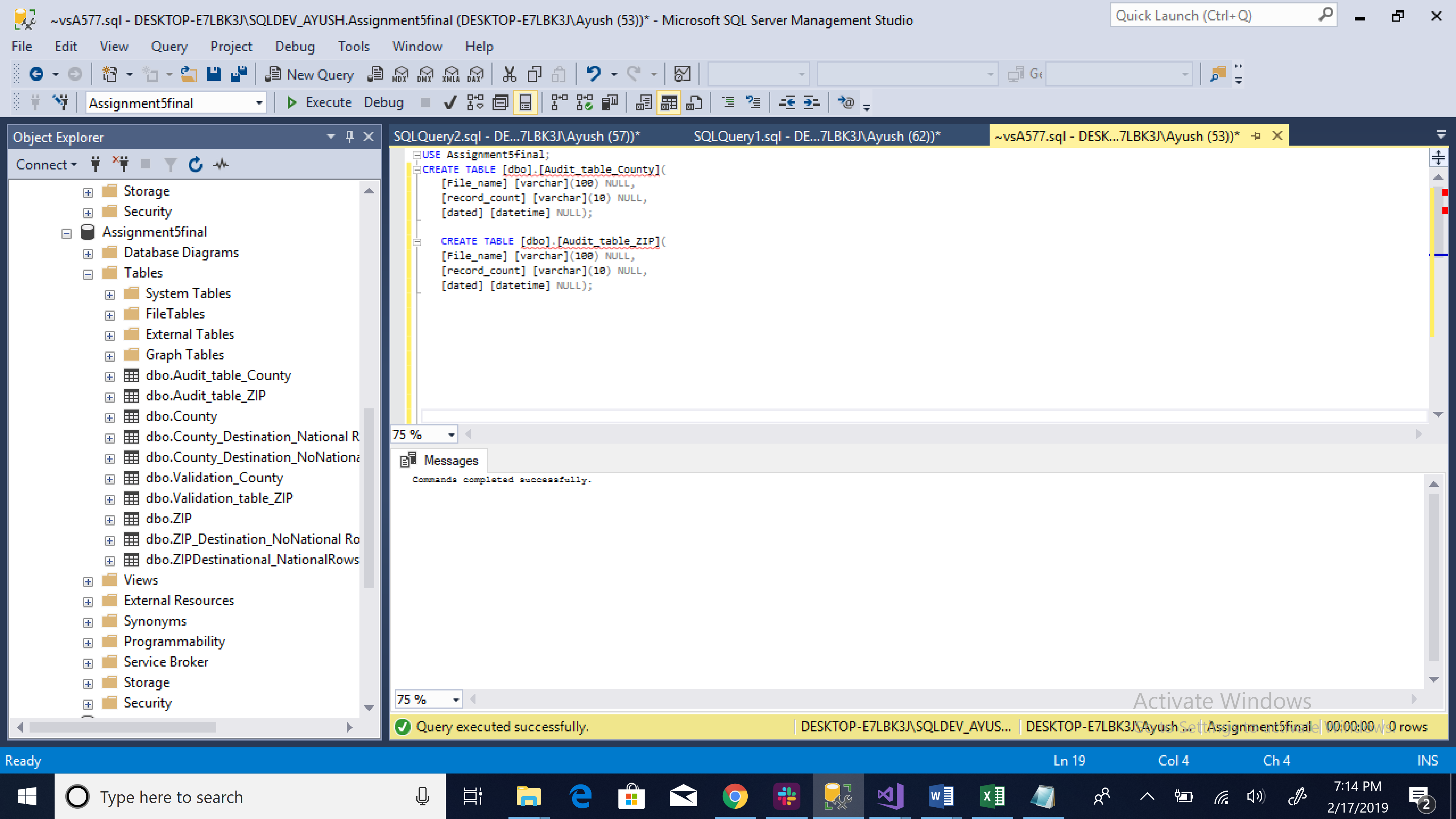
Step 2:

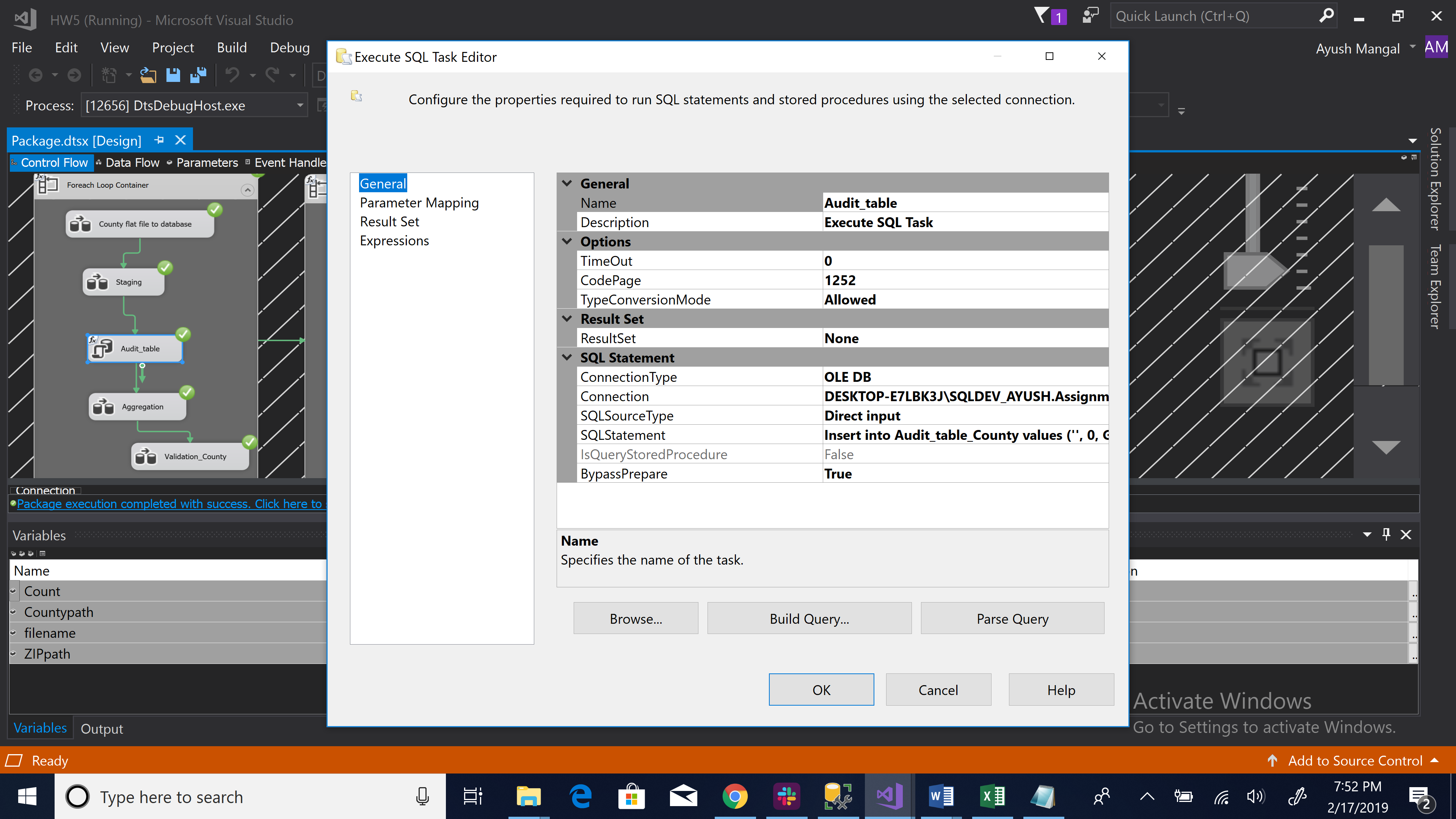
* Addition of Year and Date\_Time column using Derived column task.
* We used the Row count task to count the total number of rows.
* Then using conditional split task, we segregate the National Rows from the table by providing condition (State ABBRIVIATION==””)
* Using two different OLE DB source we send the data with National rows to “County\_Destination with National Rows” table and without national rows in “County\_Destination without National Rows” Table.



Step 3:

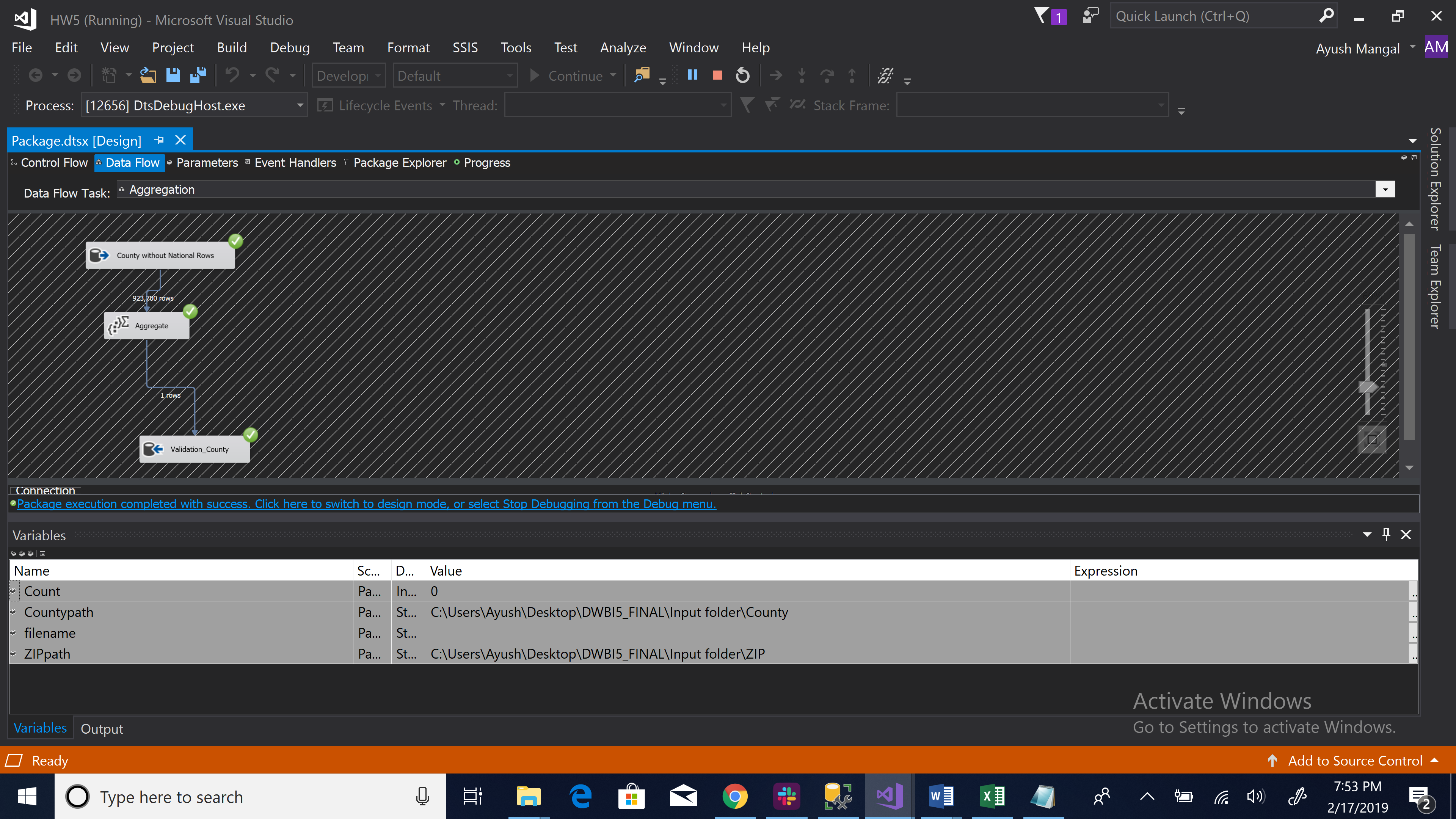
* We create a audit table in SSMS as shown in the fig.
* IN SSIS we execute sql task to fill the data in the audit table we created in SSMS by insert SQL command.





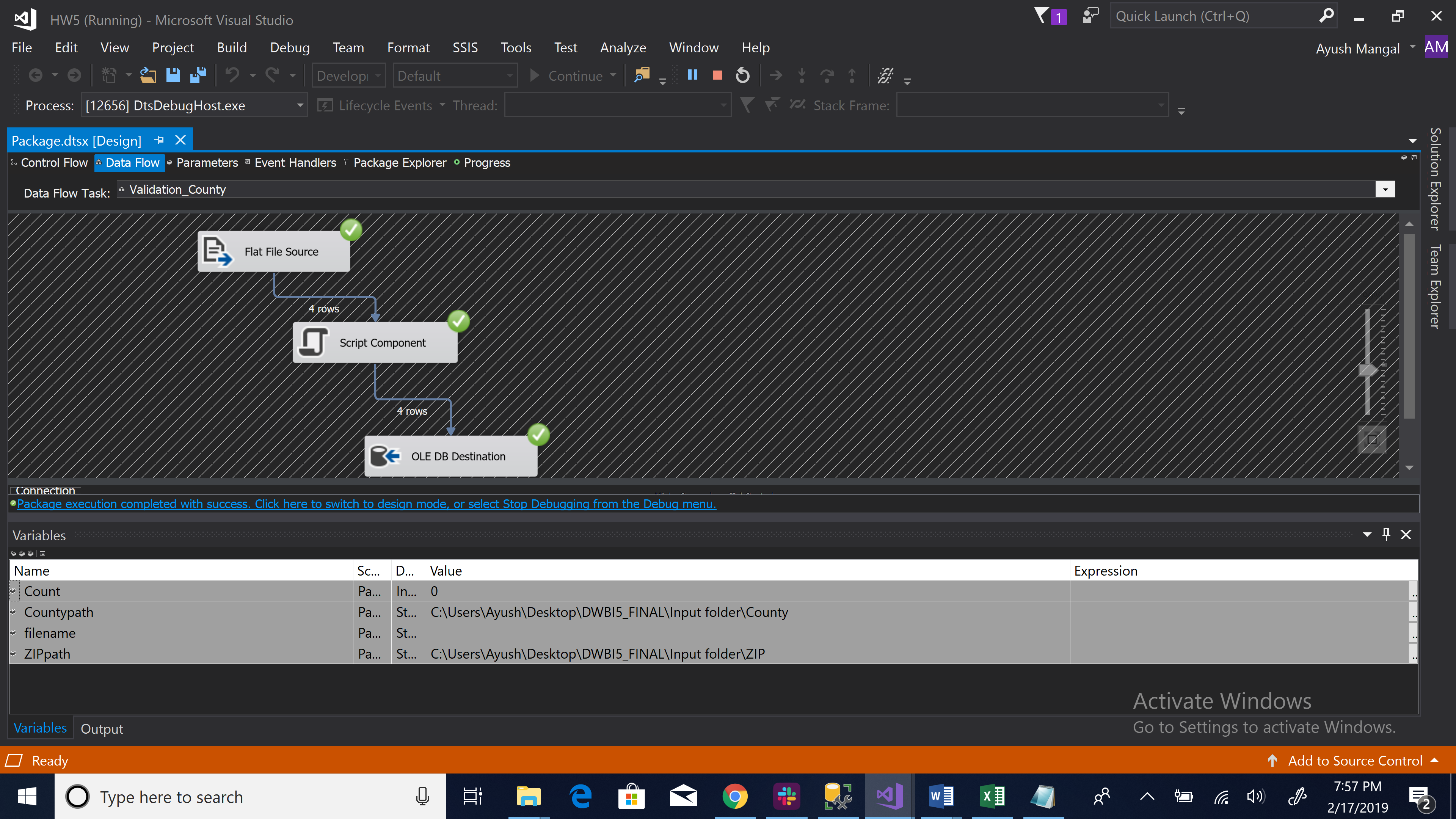
Step 4: Aggregation:

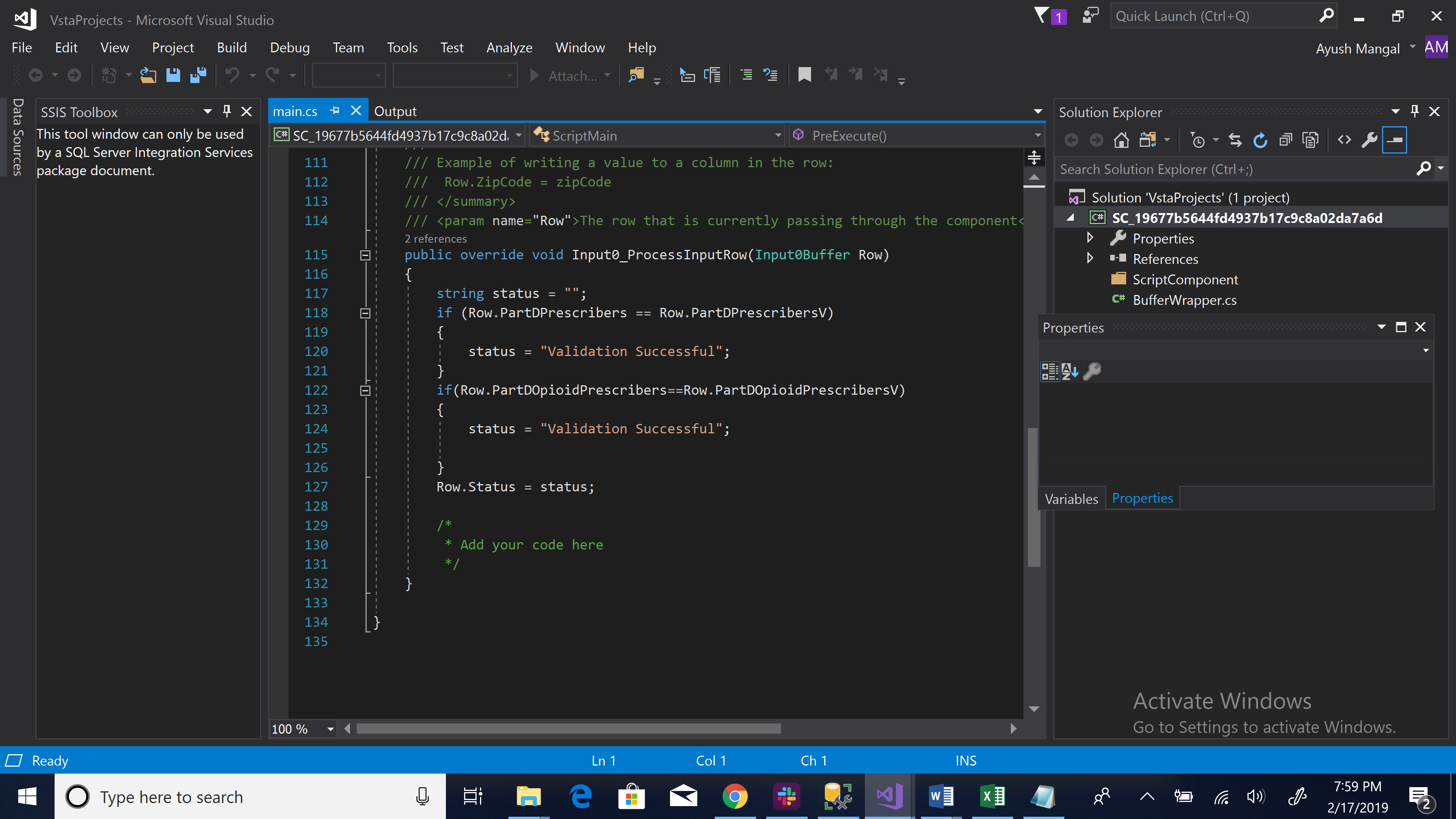
* We take the data flow task and use the County\_Destination without national rows table as our source.
* Using aggregation task in SSIS, we aggregate the selected colums and choose the sum function.
* Finally, we take a destination source as OLE DB destination and make a new table as Validation\_County to see the validated result.



Step 5:

* We used the merged table of validation as flat file source.
* Second, we used c# to validate the two tables total national rows and aggregated table.
* Finally, we took destination source as OLE DB source.





For ZIP foreach loop task we follow the same steps as County.

**Solution 3:**

**DDL Script generated by taod modular.**

