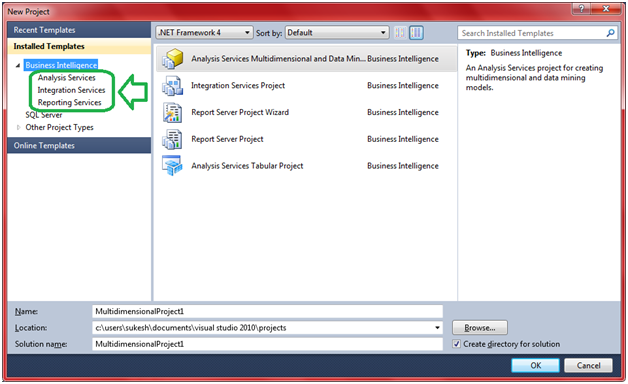
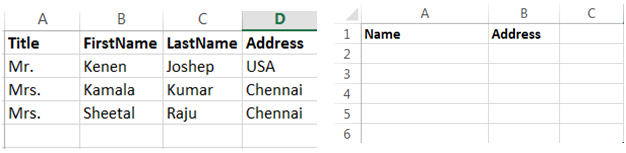
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Hope Foundation’s,**  **Finolex Academy of Management and Technology, Ratnagiri** | | | | | | | | | |
| **Department of Information Technology** | | | | | | | | | |
| Subject name: OLAP LAB | | | | | | | | Subject Code: ITL503 | | | |
| Class | | TE IT | | Semester – V (CBCGS) | | | | Academic year: 2018-19 | | | |
| Name of Student | | **Kazi Jawwad A Rahim** | | | | | **QUIZ Score : 04** | | | | |
| Roll No | | **32** | | | Experiment No. | | | | | 09 | |
| Title**: To implement ETL process** | | | | | | | | | | | |
|  | | | | | | | | | | | |
| 1. **Course objectives applicable:**   **LOB1**- To create awareness of how enterprise can organize data from various sources. | | | | | | | | | | | |
| 1. **Course outcomes applicable:**   **LO1**- Apply the knowledge of ETL processes of data warehouse. | | | | | | | | | | | |
| 1. **Learning Objectives:**  * To be able to design considerations for implement an ETL solution. * To understand various operations of data warehouse. | | | | | | | | | | | |
| 1. **Practical applications of the assignment/experiment:**  * Business/ Enterprise application | | | | | | | | | | | |
| **5. Prerequisites**: Knowledge of Data warehouse | | | | | | | | | | | |
| **6. Hardware Requirements**:   1. PC with 4GB RAM, 500GB HDD,   **7. Software Requirements:**  1. SQL Server 2012 2. SSIS -Sql server Integration services | | | | | | | | | | | |
|  | | | | | | | | | | | |
| **8. Quiz Questions (if any): (Online Exam will be taken separately batch wise, attach the certificate/ Marks obtained)**  https://goo.gl/7xBdp2 | | | | | | | | | | | |
|  | | | | | | | | | | | |
| **9. Experiment/Assignment Evaluation:** | | | | | | | | | | | |
| **Sr. No.** | **Parameters** | | | | | | | | **Marks obtained** | | **Out of** |
| **1** | Technical Understanding (Assessment may be done based on Q & A **or** any other relevant method.) Teacher should mention the other method used - | | | | | | | |  | | 6 |
| **2** | Neatness/presentation | | | | | | | |  | | 2 |
| **3** | Punctuality | | | | | | | |  | | 2 |
| **Date of performance (DOP)** | | |  | | | **Total marks obtained** | | |  | | **10** |
| **Date of checking (DOC)** | | |  | | | **Signature of teacher** | | | | | |

**Result :-**

**Step 1**. Open Sql Server Data Tools

**Step 2.** Click on File >> New >> Project. A dialog box will popup similar to like this

**We have Two excel files as follows :**

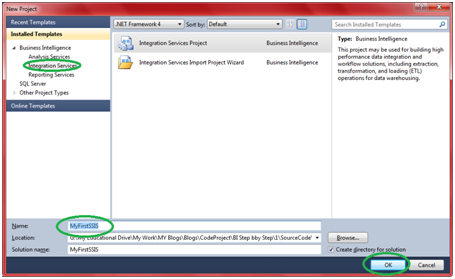


* We will collect data from first file (Datafile.xlsx).
* Convert that data so that it match to second excel file format (merge Title, FirstName and LastName and call them as Name).
* Dump final result to second file (Result.xlsx).

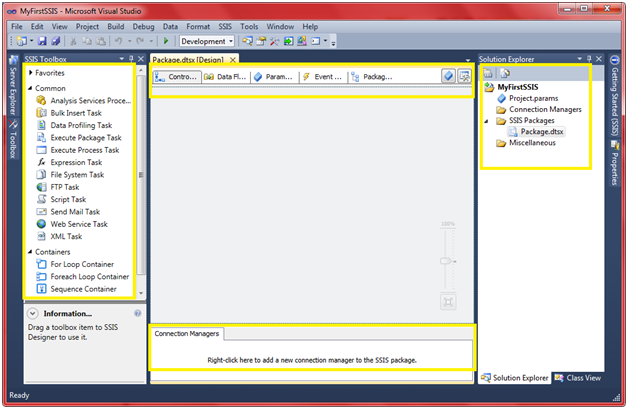
### Simple ETL process using SSIS :

**Step 1. Create New Project**

Click File >> New >> Project. Select Integration Services from the group. Specify some nice name. Say Ok.



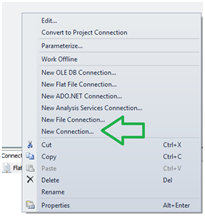
It will open up SSIS designer which you will use for creating and maintaining Integration service packages. It looks like follows,



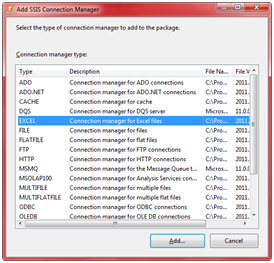
In the solution explorer under “SSIS packages” folder we will see one default package created with name “Package,dtsx”. If we want you can simply “rename it”or “remove it and add new one” (right click the folder and say “New SSIS Package”).

**Step 2. Create Connection Manager for Excel File**

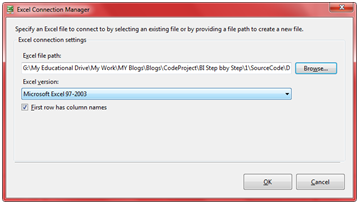
2.1 Right click Connection Manager and Say New Connection.



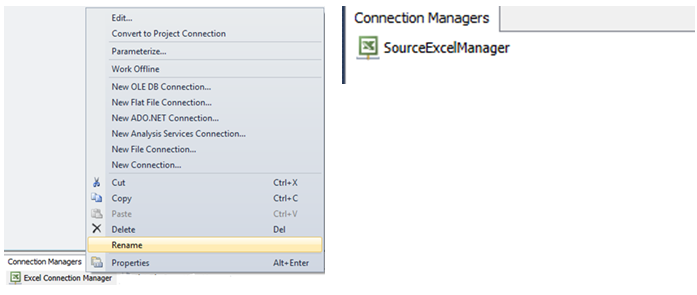
2.2 Select Excel from the popup and click on Add.



2.3 Click the browse button and select the excel file and click on OK.



**Step 3. Rename Connection Managers**

3.1 Right click the connection manager just added and rename it to SourceExcelManager.

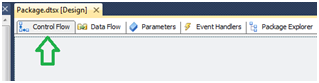
**Step 4. Create Destination connection manager**

4.1 Follow the Step no 3 and create one more connection manager pointing to Result.xlsx file.

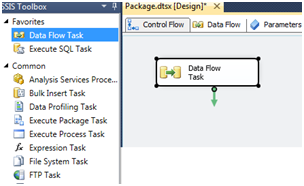
4.2 Rename connection manager to ExcelDestinationManager.

**Step 5. Create control flow – Pass data from Source Excel to Destination Excel.**

5.1 Make sure control flow tab is selected in SSIS designer.



5.2 Select data flow task from the toolbox and drag it into designer.



5.3 Rename Data Flow Task to “Source excel to Destination excel transfer task”

**Control Flow**

Control flow will be used to define the workflow. As the name implies it control the flow of execution.

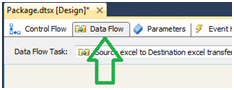
**Step 6. Create Data Flow.**

**Data Flow**

Data Flow defines the flow of data between source and destination.

6.1 Double click the control flow created in last step.

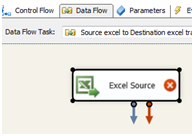
6.2 It will take you to second tab that is Data Flow tab.



**Step 7. Create Excel Source**

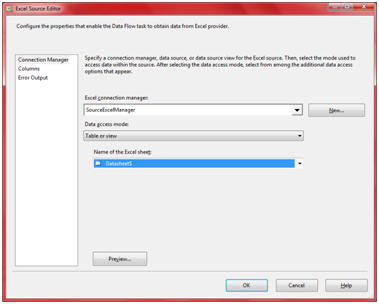
7.1 Now you will see a different SSIS toolbox all together. In toolbox you will see couple of groups defined like Sources, Destinations and Transformations.

Take excel source from Source group and place it in designer.



**Step 8. Configure Excel Source**

8.1 Red Cross mark on excel source indicates that, it’s not configured yet. Double click the excel source. It will show up a dialog box something like this.



8.2 Select Data Source as "SourceExcelManager", Data Access Mode as "Table or View" and Name of the sheet as "DataSheet1" (Name of the sheet in the excel file).

**Step 9. Create Derived Column**

9.1 From the SSIS toolbox from Transformation group drag Derived column to SSIS designer.

**Step 10. Connect Source to Derived Column**

10.1 Click the Excel source added in prior step.

You can see a small blue arrow attached to the source. We call it “Data Flow Path”.

https://www.codeproject.com/KB/database/751447/image_18.png

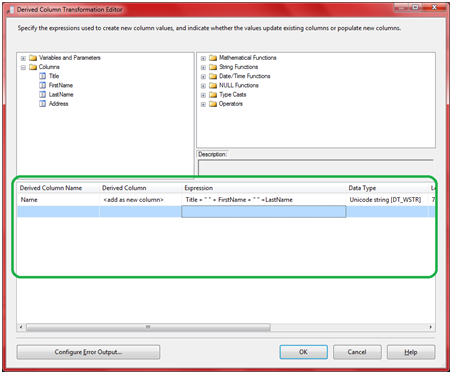
**Data Flow Path:** It lets you define how data will flow.

Click on the blue arrow and connect it to Derived Column.



**Step 11. Configure the derived column**

11.1 Double click the derived column. Popup looks like follows.



11.2 Put down Derived Column Name as **Name**, Select Derived Column as <add>and expression as **Title** **+ “ “ + FirstName + “ “ + LastName**  
11.3 Click Ok.

**Step 12. Create Excel Destination**

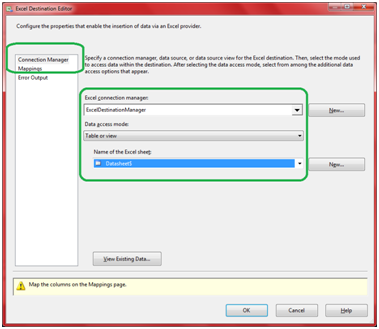
12.1 Add Excel Destination from the Destination group in SSIS toolbox.

**Step 13. Connect Derived Column to Excel Destination**

13.1 Just like step no 10, connect derived column to excel destination.

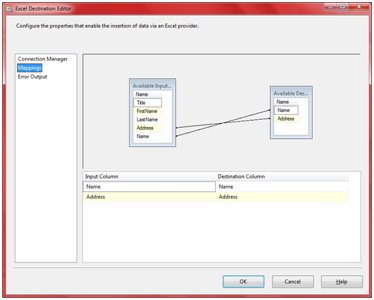
**Step 14. Configure Excel Destination**

14.1 Double click the Excel destination, popup looks like follow.



14.2 Set connection Manager to “ExcelConnectionManager”, Data Access mode to “Table or View” and Name of the excel sheet to “Datasheet1”.

14.3 Click on mapping and make sure its proper, if not make sure to do it before proceeding.

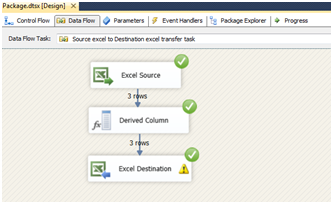


14.4 Click ok.

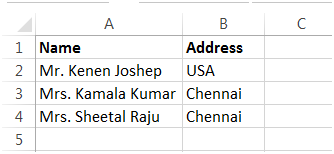
**Step 15. Execute package**

15.1 Press F5.

On successful execution we will get a screen something like this.



16.2 Open the Result.xlsx file and confirm the output.



**References**:

1. W. H. Inmon, "Building the Data Warehouse", 3rd edition.

2. Anahory and Murray .,Data warehousing in the real world , Pearson Education/Addison Wesley.

3. https://docs.oracle.com/cd/B10501\_01/server.920/a96520/concept.htm