Data Warehouse Developer

Install the “Data storage and processing” from Visual Studio Installer

From studio 19 and above you need to download the template from the Extension menu

These are the extensions you need to install:

* SQL Server Integration Services Projects
* Microsoft Reporting Services Projects
* Microsoft Analysis Services Projects

31.Create a new SSIS project

A screenshot of a computer

Description automatically generated

32. Key questions to ask when you have a new data source and want to integrate to your ETL solution.

A green and white text on a black background

Description automatically generated

Create a new Data Profile Task (Drag it to the Control Flow tab) and connect to Database/table/column, Address1, to see the Address1 value length (Column Length Distribution Profile Request) and distribution (Column Null Ratio Profile Request).

A screenshot of a computer

Description automatically generated

The Data.txt file should already be existed and select True for Overwrite Destination.

A screenshot of a computer

Description automatically generated

Final result when you click on the Open Profile Viewer:

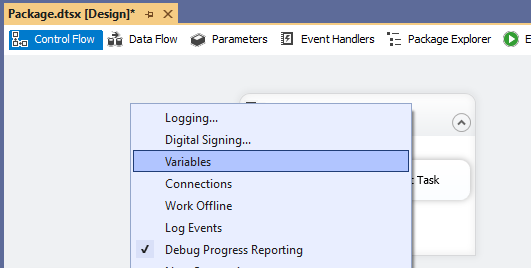
You see that the maximum length is 33 and you need to consider this for the reports you are creating.

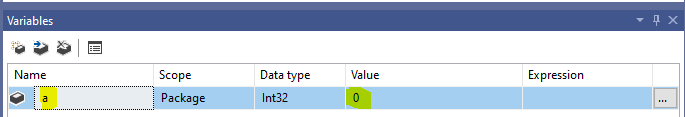
A screenshot of a computer

Description automatically generated

34. Introduction to Control Flow-1

Right click on the design page and create a variable





Drag the For Loop Container from SSIS Toolbox to the design page.The Script Task will be dragged later.

A screenshot of a computer

Description automatically generated

Update as you see.

A screenshot of a computer

Description automatically generated

Drag the Script Task from SSIS Toolbox to the For Loop Container and double click on it and then click on the Edit Script. You’ll see the below new project and update as you see.

A screenshot of a computer

Description automatically generated

Select variable a :

A screenshot of a computer

Description automatically generated

Right click on container and select Execute Container. You’ll see the a values messages pop up.

35. Introduction to Control Flow-2-Conditional Statement

Drag Expression Task and double click on it:

A screenshot of a computer

Description automatically generated

Create two more Expression Tasks

A diagram of a task

Description automatically generated

Put the below condition for Task to Task 1 relation.

A screenshot of a computer

Description automatically generated

And put @[User::a] > 1 for the other relation.

This is what you see when you run the package because 18 > 1

A diagram of a function

Description automatically generated

36. Implementing Data Flow – 1

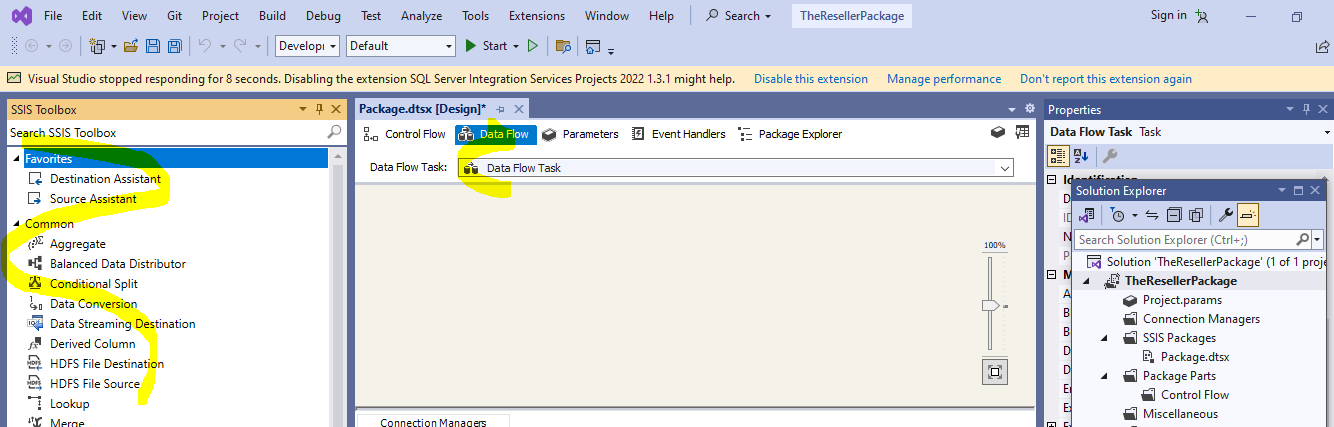
Drag the Data Flow to the design page:

A screenshot of a computer

Description automatically generated

Double click on the Data Flow Task and you’ll notice the tab changes to Data Flow tab.

The Toolbox has been changed too.



Every Data Flow needs a source and destination.

Drag Source Assistant>Flat File>New

A screenshot of a computer

Description automatically generated

You can change the connection name. Select the txt file.

A screenshot of a computer

Description automatically generated

Click OK and you’ll see this:

A white rectangular sign with blue and red arrows

Description automatically generated

Now we need destination:

Drag Destination Assistance and select SQL Server

A screenshot of a computer

Description automatically generated

Here is the result:

A screenshot of a computer

Description automatically generated

The blue arrow represents good data, successfully read or transformed data.

The red arrow represents bad data, not successfully read or transformed data.

Connect the blue arrow to the destination.

A diagram of a computer

Description automatically generated

Double click on the Destination.

Select the destination table or view:

A screenshot of a computer

Description automatically generated

Check the Mapping and it seems fine:

A screenshot of a computer

Description automatically generated

Click OK and you’ll see this:

A diagram of a computer

Description automatically generated with medium confidence

You need to truncate the table data first and then the data will be load from the txt file.

37. Implementing Data Flow – Part 2

A screenshot of a computer

Description automatically generated

Click OK and connect the arrow.

A screenshot of a computer

Description automatically generated

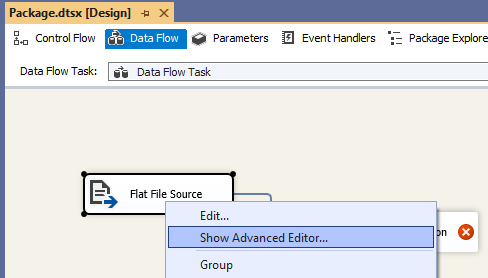
You see there are errors for columns SalesOrderNumber, CarrierTrackingNumber, and CustomerPONumber, when you run the package. It is about data type inconsistency.

A screenshot of a computer

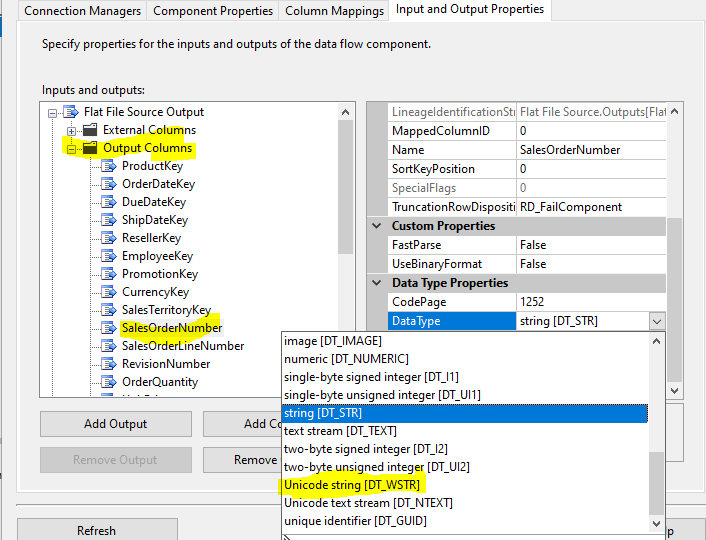
Description automatically generated

To fix it:

Right click on the Flat File Source and click on the Show Advanced Editor.



Select as highlighted for the three columns SalesOrderNumber, CarrierTrackingNumber, and CustomerPONumber and click OK.



Here is the result:

A diagram of a data flow

Description automatically generated

Run the package by click on the Start:

A screenshot of a computer

Description automatically generated

The number of rows are increasing:

A computer screen shot of a diagram

Description automatically generated

And done.

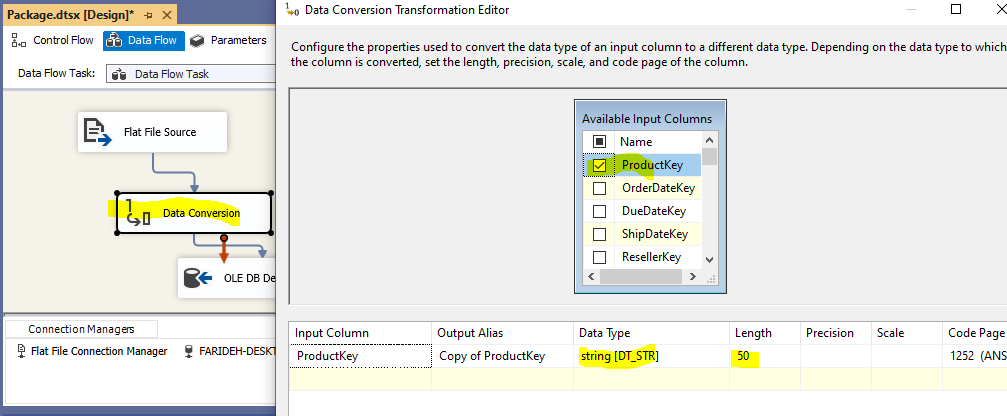
A diagram of a diagram

Description automatically generated

**Data Conversion:**

Drag Data Conversion here between Source and Destination. You can change the data format before transferring.

Double click on the Data Conversion and select an input (ProductKey). It is integer and you see that you can change it to string or change the length.



Run it and you’ll see the same result.

