Loading Data into OLAP using For Each Loop Container & File System Task

Table of Contents

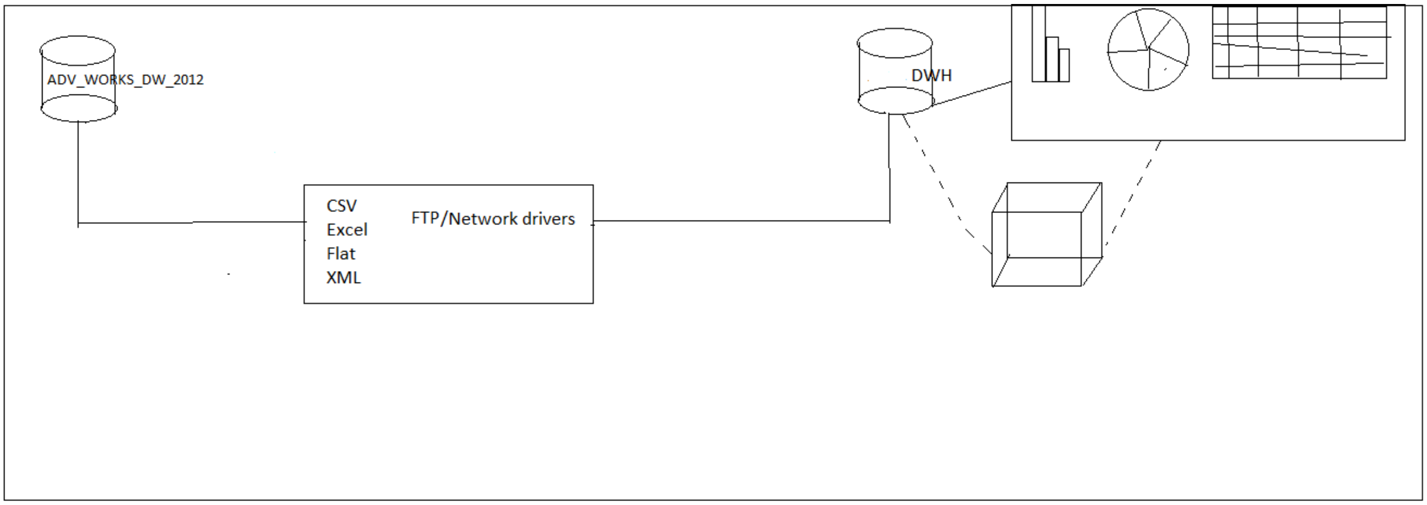
Introduction

This document will guide you to load the Countries list into Data Warehouse respective Country tables, with step by step guide lines.

Project data flow architecture diagram

We do have the following ETL flow in the project. Namely the staging areas are

- AdventureWorksDW2012 (F-D model RDBMS)
- Files area (Daily refresh data through these files)
- OLAP_DB (Final F-D model destination)



OLAP (Final F-D model destination)

This is our DWH, also final destination. Data will be merged /append from above Dest_Staging to OLAP. Data loading strategy will be INCREMENTAL LOAD.
In our practice model, below is the details about the OLAP

```
Data Source=VISSU; Initial Catalog=OLAP_DB=SQLNCLI11.1; Integrated
Security=SSPI;Auto Translate=False;
```

ETL frame work objects(Tables/Procedures/Functions)

ETL Frame work tables

Create the destination tables with structures in the OLAP namely **Dbo.Australia, Dbo.Canada , Dbo.France, Dbo.Germany , Dbo.UnitedKingdom , Dbo.United States**

Script:

```
SELECT TOP 0 * INTO OLAP_DB.dbo.Australia FROM  AdventureWorksDW2012.dbo.SalesVw
SELECT TOP 0 * INTO OLAP_DB.dbo.Canada FROM  AdventureWorksDW2012.dbo.SalesVw
SELECT TOP 0 * INTO OLAP_DB.dbo.France FROM  AdventureWorksDW2012.dbo.SalesVw
SELECT TOP 0 * INTO OLAP_DB.dbo.Germany FROM  AdventureWorksDW2012.dbo.SalesVw
SELECT TOP 0 * INTO OLAP_DB.dbo.[United Kingdom] FROM  AdventureWorksDW2012.dbo.SalesVw
SELECT TOP 0 * INTO OLAP_DB.dbo.[United States] FROM  AdventureWorksDW2012.dbo.SalesVw
```

ETL Frame work procedures

Use the **InsertSSISBatch2, UpdateSSISBatch2, InsertSSISTask2, UpdateSSISTask2** procedures from the Package10 Design Document.

ETL Frame work Views

[dbo].[SalesvW]

Create a view **dbo.SalesvW** by joining 5 specified tables from AdventureworksDW2012 from the following script

Script:

```
CREATE  VIEW SalesVw
AS
SELECT
 GETDATE() AS LoadedDate
,DD.DateKey
,DD.CalendarYear
,DD.CalendarSemester
,DD.CalendarQuarter
,DD.MonthNumberOfYear
,DD.EnglishMonthName
,DD.FullDateAlternateKey
,DP.ProductKey
,DP.EnglishProductName
,DP.EnglishDescription
,DG.GeographyKey
,DG.EnglishCountryRegionName
,DG.StateProvinceName
,DG.City
,DC.CustomerKey
,DC.FirstName
,DC.LastName
,DC.MiddleName
,DC.Gender
,DC.EmailAddress
,FIS.SalesAmount
,FIS.TaxAmt

FROM [dbo].[FactInternetSales] FIS
INNER JOIN [dbo].[DimCustomer] DC
ON DC.CustomerKey = FIS.CustomerKey
INNER JOIN [dbo].[DimGeography] DG
ON DG.GeographyKey = DC.GeographyKey
INNER JOIN [dbo].[DimProduct] DP
ON FIS.ProductKey = DP.ProductKey
INNER JOIN [dbo].[DimDate] DD
ON DD.DateKey = FIS.OrderDateKey
```

ETL Frame work functions

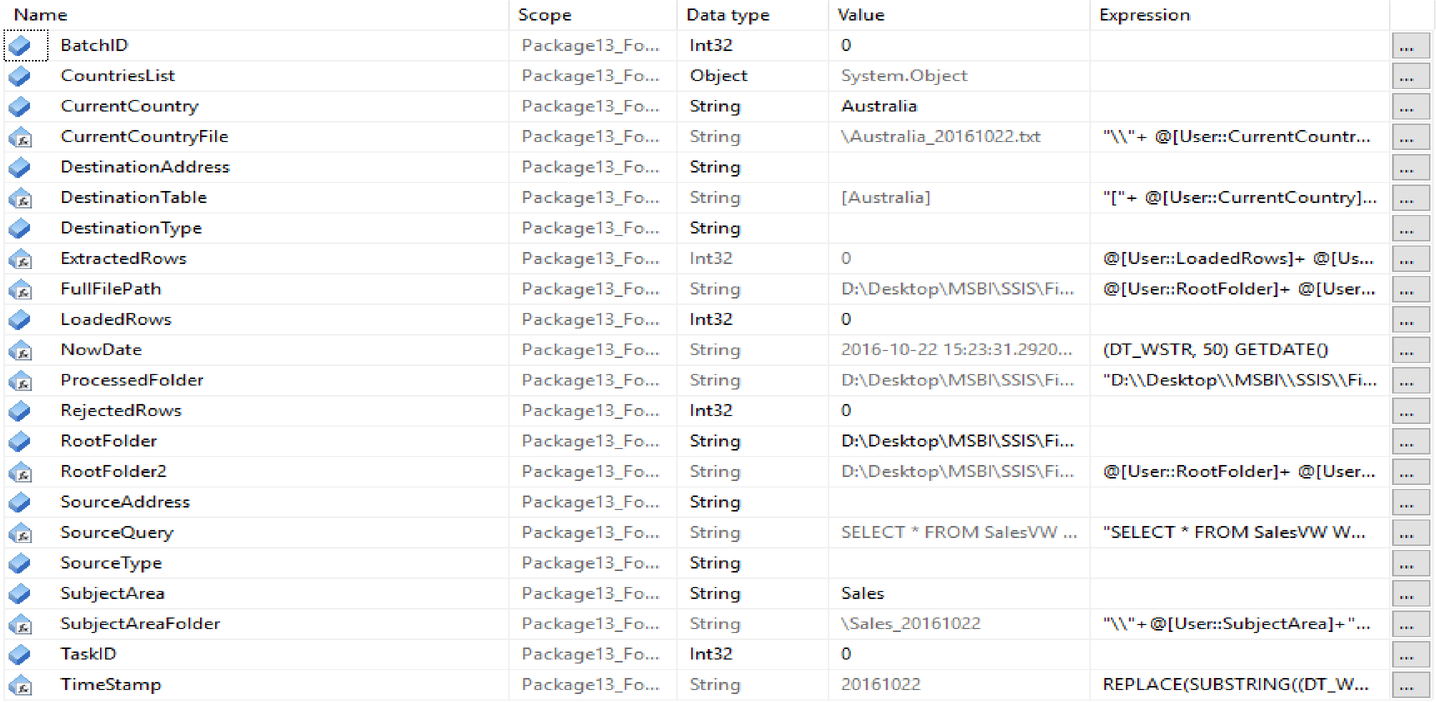
[dbo].[DateAndTimeNarration]

This function will accept two arguments (Start & End times), and returns the elapsed time, which helps us to close the audit (Task & Batch) record.

Script:

dbo.DateAndTimeNarration.sql

Important Package variables we use



Name	Scope	Data type	Value	Expression	
BatchID	Package13_Fo...	Int32	0		...
CountriesList	Package13_Fo...	Object	System.Object		...
CurrentCountry	Package13_Fo...	String	Australia		...
CurrentCountryFile	Package13_Fo...	String	\Australia_20161022.txt	"\\"+ @[User::CurrentCountr...	...
DestinationAddress	Package13_Fo...	String			...
DestinationTable	Package13_Fo...	String	[Australia]	"["+ @[User::CurrentCountry]...	...
DestinationType	Package13_Fo...	String			...
ExtractedRows	Package13_Fo...	Int32	0	@[User::LoadedRows]+ @[Us...	...
FullFilePath	Package13_Fo...	String	D:\Desktop\MSBI\SSIS\Fi...	@[User::RootFolder]+ @[User...	...
LoadedRows	Package13_Fo...	Int32	0		...
NowDate	Package13_Fo...	String	2016-10-22 15:23:31.2920...	(DT_WSTR, 50) GETDATE()	...
ProcessedFolder	Package13_Fo...	String	D:\Desktop\MSBI\SSIS\Fi...	"D:\\Desktop\\MSBI\\SSIS\\Fi...	...
RejectedRows	Package13_Fo...	Int32	0		...
RootFolder	Package13_Fo...	String	D:\Desktop\MSBI\SSIS\Fi...		...
RootFolder2	Package13_Fo...	String	D:\Desktop\MSBI\SSIS\Fi...	@[User::RootFolder]+ @[User...	...
SourceAddress	Package13_Fo...	String			...
SourceQuery	Package13_Fo...	String	SELECT * FROM SalesVW ...	"SELECT * FROM SalesVW W...	...
SourceType	Package13_Fo...	String			...
SubjectArea	Package13_Fo...	String	Sales		...
SubjectAreaFolder	Package13_Fo...	String	\Sales_20161022	"\\"+@[User::SubjectArea]+"...	...
TaskID	Package13_Fo...	Int32	0		...
TimeStamp	Package13_Fo...	String	20161022	REPLACE(SUBSTRING((DT_W...	...

The Variables **NowDate , Source Type ,Destination Type,BatchID ,TaskID, ExtractedRows, LoadedRows, RejectedRows** are taken from the Package10 Design Document

The remaining variables are defined as following

CountriesList variable

This variable Countries List is Object Data Type. it can be used as variable name in Get list of countries Execute SQL Task and ADO Object source in For Each Loop.

CurrentCountry variable

It is a String variable and defaulted to Australia it can it's value at runtime in the for each loop container

Timestamp variable

It is string variable evaluated to the current date using the following expression

Expression
REPLACE(SUBSTRING((DT_WSTR, 50) GETDATE(),1,10),"-","")

RootFolder variable

RootFolder is a string type variable and that holds the path of the folder to which the flat files are created andalso processe.

It is static and value as

```
D:\Desktop\MSBI\SSIS\Files\Flatfiles\DailySalesDetails
```

Source and Destination address variable

Source address is a string variable that is default null stores the source address. It is evaluated during execution in the script task.
Destination address is a string variable that is default null the stores the destination address. It is evaluated during execution in the script task

Source Query variable

This is a string type variable. This variables evaluates the sql query over the source oledb database using the following expression

Expression

```
"SELECT * FROM SalesVW WHERE EnglishCountryRegionName = '"+ @[User::CurrentCountry]+"'"
```

Destination Table variable

This is a string type variable. This variables holds the table name in the destination evaluaed dynamically

```
"["+ @[User::CurrentCountry]+"]"
```

RootFolder2 variable

The RootFolder2 is Dynamic variable of String data type that is evaluated from the following expression
Expression:

```
@[User::RootFolder]+ @[User::SubjectAreaFolder]
```

ProcessedFolder variable

It is also static variable of sting data type. It is used to move sales data that is generated today is processed(moved) to that folder the folder pathis as follows

```
D:\Desktop\MSBI\SSIS\Files\Flatfiles\ProcessedFiles
```

Open BatchID

Drag an execute SQL task name it as Open BatchID. Used in the package 10 design document
Drag a file system task
Under general tab
Set operation = create directory
Source variable= User::RootFolder2
Drag a execute sql task name it as List of countries
Under general tab
Set ResultSet = full result set
Connection type=OLEDB
Connection= vissu.AdventureWorksDW2012
Sql Statement as

```
SELECT DISTINCT
EnglishCountryRegionName
FROM
AdventureWorksDW2012.dbo.SalesVw
```

ETL Task-1 (AdventureWorksDW2012 to FlatFile)

Before you open the ETL sequence, drag and drop a for loop container
Under collection tab
Set enumerator = for each ADO enumerator
Ado object source variable = user ::CountriesList
Enumeration mode = Rows in the first table
Under variables mapping tab
Set variable set::CurrentCountry
Index= 0

Now drag and a script task and select variables as
User::destination address, user;;destination Type, User :: Source Address , UserSource Type
Now edit the
Script:

```
     Dts.Variables["User::SourceType"].Value = "RDBMS";
            Dts.Variables["User::SourceAddress"].Value = "Data Source=VISSU;Initial
Catalog=AdventureWorksDW2012;Provider=SQLNCLI11.1;Integrated Security=SSPI;Auto
Translate=False;";
            Dts.Variables["User::DestinationType"].Value = "FILES";
            Dts.Variables["User::DestinationAddress"].Value =
Dts.Variables["User::FullFilePath"].Value;
```

Open TaskID

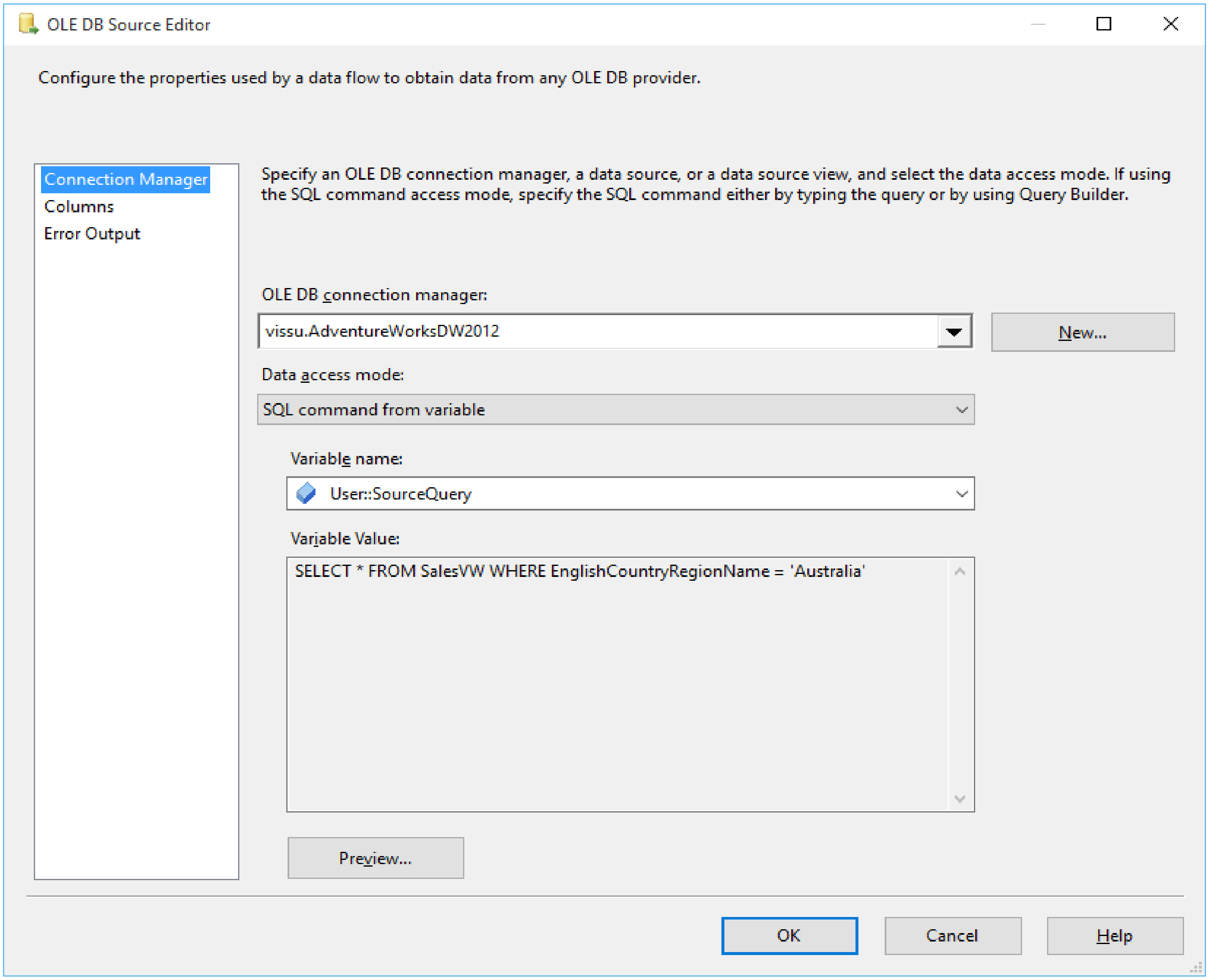
Drag an execute SQL task name it as Open TaskID. Used in the package 10 design document

Mapping steps for Task-1

Drag and drop a Data flow task, and choose “OLEDB“ (Object Linking Embedded DB) type source and ”Flat File” type destination.

Source Configuration:

Drag and drop an OLEDB source. Choose vissu. AdventureWorksDW2012. Choose the data access mode as “SQL Command from variable”, and then map the “SourceQuery” variable.



OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager: vissu.AdventureWorksDW2012

Data access mode: SQL command from variable

Variable name: User::SourceQuery

Variable Value:

```
SELECT * FROM SalesVW WHERE EnglishCountryRegionName = 'Australia'
```

Destination Configuration:

Drag a follow-up flat destination to the above row count transformation, and choose the destination as. Choose the dynamic expression for the flat file connection as

Expression:
@[User::FullFilePath]

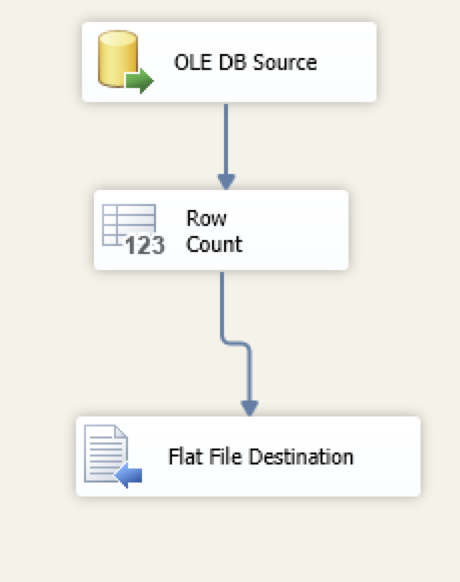
Static Path for **Flat file destination connection** is as follows

Path:

D:\Desktop\MSBI\SSIS\Files\Flatfiles\DailySalesDetails\Sales_20161022\Australia_20161022.txt

The finalized ETL task 1 screen shot is as follows

Finalized ETL screen shot



Close TaskID

Drag an execute SQL task name it as close TaskID. Used in the package 10 design document

Task-2 (Flat file to OLAP_DB)

Open TaskID

Drag an execute SQL task name it as Open Task Used in the package 10 design document

Mapping steps for Task-2

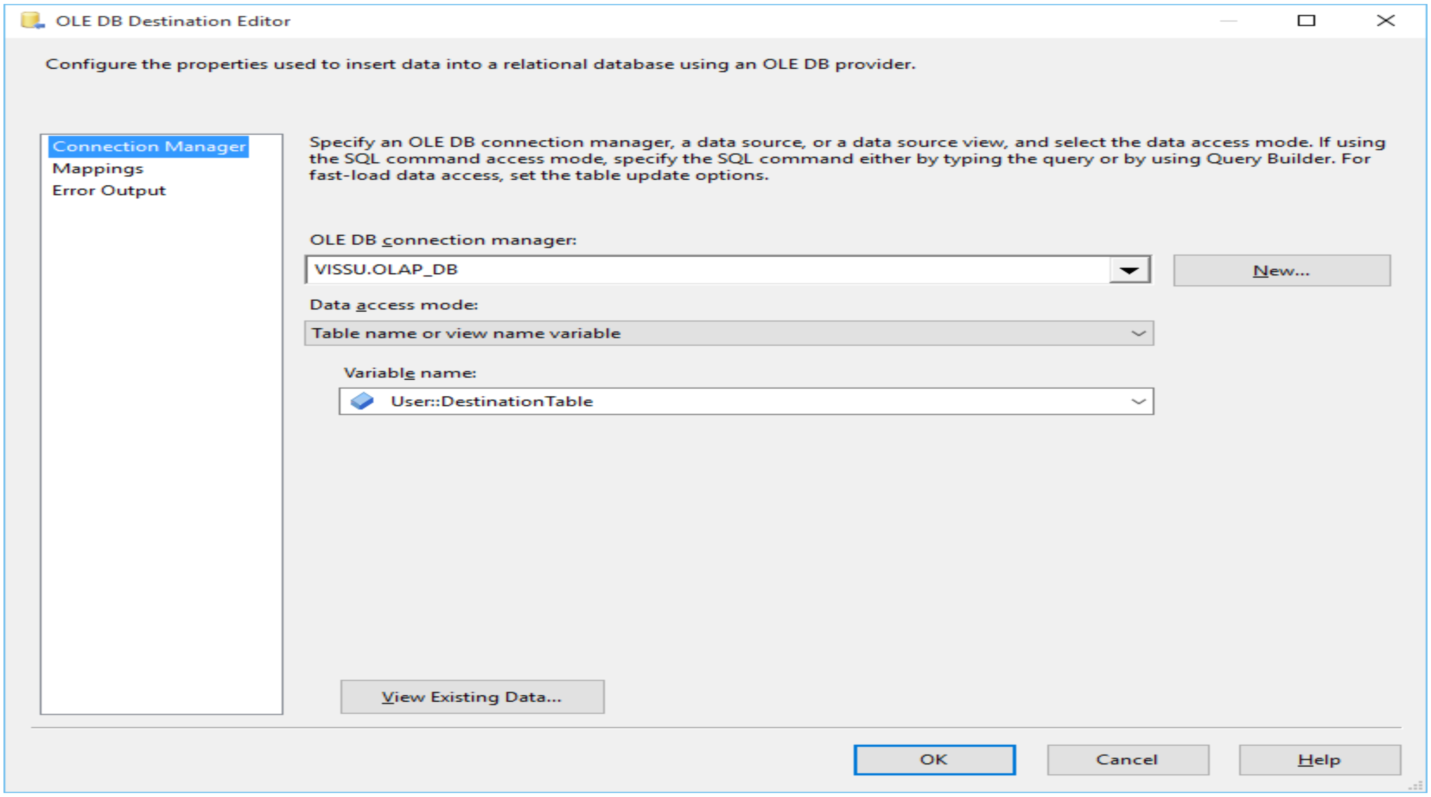
Drag and drop a Data flow task, and choose “FlatFile Source” and “OLEDB” type destination.

Source Configuration:

Configure the Flat File source as from ETL task1 same here.

Destination Configuration:

Drag and drop an OLEDB destination. Choose **vissu.OLAP_DB**. Choose the data access mode as “Table name or view variable”, and then map the “**DestnationTable**” variable.



OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

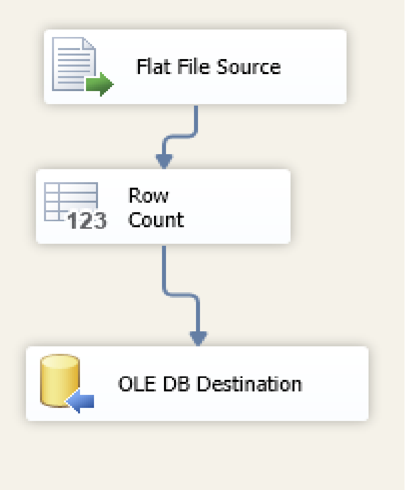
OLE DB connection manager: VISSU.OLAP_DB

Data access mode: Table name or view name variable

Variable name: User::DestinationTable

The finalized ETL Task 2 is as follows

Finalized ETL screen shot



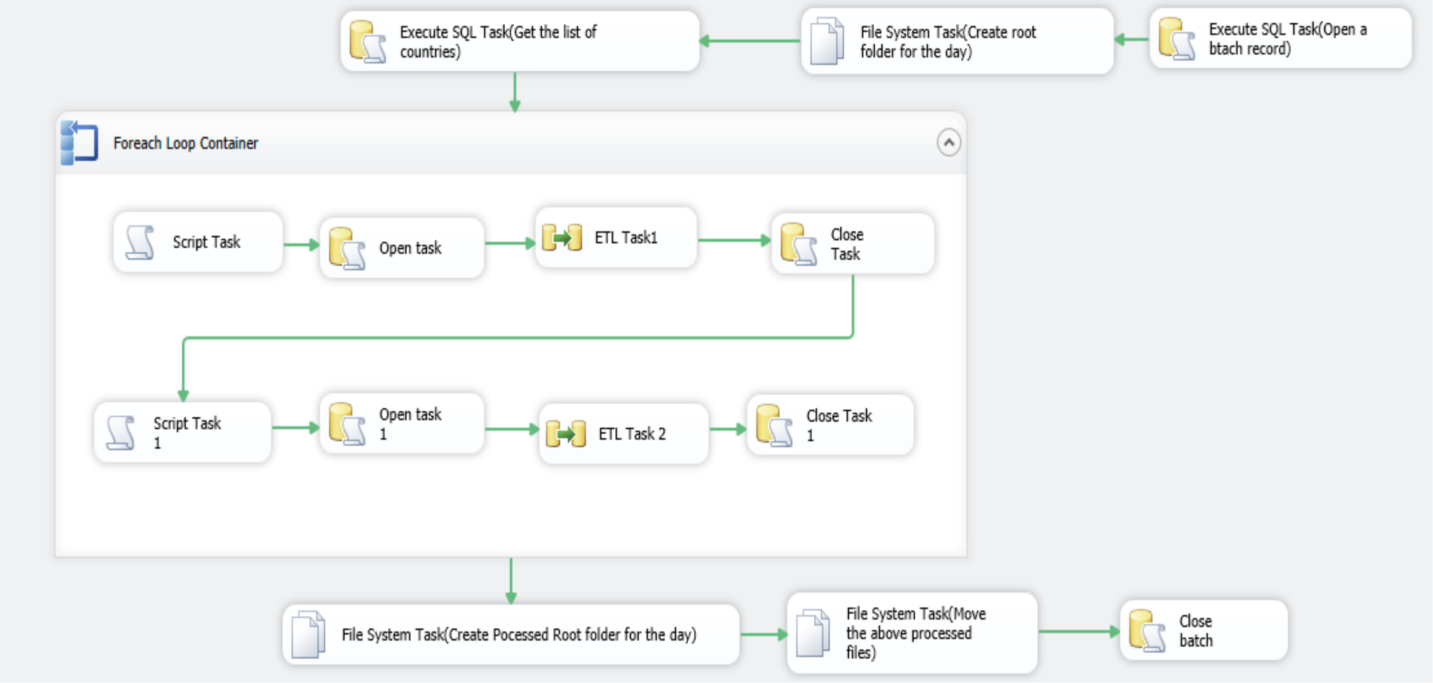
Close TaskID

Drag an execute SQL task name it as Close Task1 Used in the package 10 design document

Close BatchID

Drag an execute SQL task name it as "Close Batch" Used in the package 10 design document

Finalized Package screen shot


Execute SQL Task(Get the list of countries)
File System Task(Create root folder for the day)
Execute SQL Task(Open a btach record)
Foreach Loop Container
Script Task
Open task
ETL Task1
Close Task
Script Task 1
Open task 1
ETL Task 2
Close Task 1
File System Task(Create Pocessed Root folder for the day)
File System Task(Move the above processed files)
Close batch