

Intro to SSIS

Preliminary work

Create Arcane Code schema.

Create table to hold staging output:

```
use ArcaneCode;
go
create table staging_SalesData
(   OrderDate datetime
    , LineTotal numeric(38,6)
    , ProductLine nvarchar(2)
)
```

Create table to stage Update rows.

```
create table staging_SalesSummary
(   OrderDate datetime
    , LineTotal numeric(38,6)
    , ProductLine nvarchar(2)
)
```

Create table to hold output:

```
create table Sales_Summary
(   OrderDate datetime
    , SubTotal numeric(38,6)
    , ProductLine nvarchar(2)
)
```

Create the Data Sources.

1. Right click on Data Sources.
2. Pick New Data Source.
3. Click next on the opening Data Source Wizard Screen.
4. Click New.
5. Under Server Name, pick N4IXT\SQLEXPRESS.
6. Under Select database name, pick AdventureWorks.
7. Click Next.
8. Enter "AdventureWorks" for the Data Source Name. Click Finish.
9. Repeat for ArcaneCode.

Create Connection Managers.

1. In the Package.dtsx area, right click in the Connection Managers area at the bottom.
2. Select "New Connection from Data Source".
3. Select "Adventure Works" and click OK.
4. Repeat for "Arcane Code".

Create the first DataFlow to move data to staging.

1. Drop a Data Flow Task on the package.
2. Rename it "DFT - Move to staging."
3. Open the Data Flow Task.
4. Drop an OLE DB Source on it. Rename to SRC – Adventure Works Sales
5. Right click and pick EDIT.
6. Change Data Access mode to SQL Command. Change the SQL Command text to:

```
select sh.OrderDate
      , sd.LineTotal
      , p.ProductLine
from Sales.SalesOrderHeader sh
inner join Sales.SalesOrderDetail sd
      on sh.SalesOrderID = sd.SalesOrderID
inner join Production.Product p
      on sd.ProductID = p.ProductID
```

7. Click on Columns, just to show the columns. Click OK to close.
8. Add an OLE DB Destination.
9. Change name to DST – Staging Sales Data.
10. Drag the green arrow from the SRC to DST.
11. Highlight everything. Make sure the "Layout" toolbar is visible. Show left align and widening.
12. Right click and EDIT the DST.
13. Set the connection manager to Arcane Code.
14. Set target to Table or View DO NOT PICK FAST LOAD.
15. Pick dbo.staging\_SalesData from the drop down of the output.
16. Click on Mappings to show mappings.
17. Click OK.
18. Go to the Solution Explorer. Right click on the Package.dtsx and pick "Set as Startup Object".
19. Run the project.
20. Show the DataFlow area to show row output.
21. Show results in SQL Server.
22. Use STOP DEBUGGING button to stop.
23. Now go back and change the SRC to use Table or View – fast load.
24. Re-run and watch the speed difference!

Add SQL Task

1. Above the dataflow, add a SQL Task. Name it SQL – Reset Sales Data Staging Table.
2. Edit. Change the connection to “Arcane Code”.
3. Change the SQL Statement to “truncate table dbo.staging\_SalesData”.
4. Click OK.
5. SSMS: Do a “`select count(*) from staging_SalesData`” to show number of rows.
6. Run again.

#### Add DFT with more functions

1. Add a new DFT. Name it DFT – Move to Warehouse.
2. Drag green arrow from DFT – Move to Staging
3. Edit the DFT.
4. Add an OLEDB Source. Name it SRC – Staging Sales Data.
5. Edit. Change Connection to Arcane Code. Change table to staging\_SalesData.
6. Add a Derived Column transformation.
  - a. Name it DER – Fix Null Product Lines. Edit.
  - b. Add a Derived Column Name of ProductLine. Set Derived column to “Replace ‘Product Line’”.
  - c. Set the expression to read ISNULL(ProductLine) ? “S” : ProductLine.
  - d. Click OK.
7. Add an Aggregate Transformation. Name it AGG – Sales Total By Day. Edit.
  - a. Check the Order Date for the Input Column. Verify the operation is GroupBy.
  - b. Add a second input by checking LineTotal.
  - c. Change the operation to Sum.
  - d. Change the Output Alias to SubTotal.
  - e. Add a third column by checking ProductLine. Verify GroupBy is the operation.
  - f. Click OK to close.
8. Add an OLE DB Destination. Name it “DST – Warehouse”
9. Connect to Aggregate, then Edit.
  - a. Change Connection Manager to ArcaneCode.
  - b. Pick dbo.Sales\_Summary for the table to output to.
  - c. Click Mappings. Verify the fields are mapped correctly.
  - d. Click OK to close.
10. Run.

#### Add Lookup Logic to the flow.

1. Between the two DST’s add a SQL Task, SQL \_ Reset Sales Summary Staging Table. Reconnect the green arrows.
  - a. Edit. Set connection manager to Arcane Code.
  - b. Set SQL Statement to “truncate table dbo.staging\_SalesSummary;” Click OK.
2. Edit DFT – Move to Warehouse.
3. Insert a Lookup just before DST – Warehouse. Name it “LKP – SalesSummary”.

4. Connect the Green Arrow from the AGG to it.
5. Edit the Lookup.
6. Change the connection manager to ArcaneCode
7. Under use table or view, set the table to dbo.Sales\_Summary
8. Click on the Columns tab.
9. On the first line of LookupColumn, pick OrderDate and Replace Order Date for the Operation.
10. On the second line pick ProductLine, and Replace ProductLine for the op.
11. Click Configure Error Output.
12. Under Lookup Output, change from Fail Component to RedirectRow.
13. Close the Lookup dialog.
14. Connect the RED arrow from the LKP to the existing DST – Warehouse.
15. Add a new DST. Name it DST – Staging for Update. Connect GREEN arrow from LKP to it.
16. Set the Connection Mgr to ArcaneCode, the table to dbo.staging\_SalesSummary
17. On the Mappings tab, connect the SubTotal to the LineTotal Line. Click OK.
18. Return to the Control Flow.
19. Add a SQL Task, name it SQL – Update Summary Data
20. Edit. Change the connection to Arcane Code.
21. Add this text:

```

UPDATE dbo.Sales_Summary
SET SubTotal = staging_SalesSummary.LineTotal
FROM dbo.Sales_Summary AS dbo_Sales_Summary
INNER JOIN staging_SalesSummary
    ON dbo_Sales_Summary.OrderDate = staging_SalesSummary.OrderDate
    and dbo_Sales_Summary.ProductLine = staging_SalesSummary.ProductLine

```

22. Save and run.

#### Demonstrate Variables

1. Click on SSIS, Variables or Right Click and pick Variables.
2. Click the “Add new Variable” button.
  - a. Enter “TruncateSummaryStagingSQL” for the variable name.
  - b. Change the variable DataType to String.
  - c. Enter truncate table dbo.staging\_SalesSummary; for the text.
3. Edit the Existing SQL – Reset Sales Summary task.
  - a. Change the SQL Source type to Variable.
  - b. Pick User::TruncateSummaryStagingSQL in the drop down.
  - c. Click OK to close.
4. Add a new SQL Task. Name it SQL – Clean up Summary Data Staging. Edit.
  - a. Change the Connection to Arcane Code.
  - b. Change the SQL Source Type to Variable.
  - c. Change the Variable to “User::TruncateSummaryStagingSQL”.
  - d. Click OK to close.

5. Run.

#### Debugging

1. Open the DFT – Move to Warehouse
2. On the green bar after the AGG function right click. Select DataViewers.
3. Click Add.
4. Show the Grid tab, to demo the fields.
5. Take defaults and click OK.
6. Click OK.
7. Run and demonstrate the grid view.