Imports System

Imports System.Data

Imports System.Math

Imports Microsoft.SqlServer.Dts.Pipeline.Wrapper

Imports Microsoft.SqlServer.Dts.Runtime.Wrapper

Imports System.Text.RegularExpressions ' Se importa esto para poder usar el método Regex.Split()

<Microsoft.SqlServer.Dts.Pipeline.SSISScriptComponentEntryPointAttribute>

<CLSCompliant(False)>

Public Class ScriptMain

Inherits UserComponent

' >>> Escrito por Ernesto Moscoso Cam

' >>> Octubre 2016

' >>>

' >>> Basado en el código C# hecho por Afu Nang Tse Mundaca: http://r3xet.blogspot.pe/2013/12/obtener-el-tipo-de-cambio-de-sunat-del.html

' Esta el método que devuelve las filas del Script Component

Public Overrides Sub CreateNewOutputRows()

Dim objDataTable As Data.DataTable ' Tabla auxiliar para guardar los datos de la Web

Dim objDataRow As Data.DataRow

' Primero se configura la tabla para que tenga las 3 columnas necesarias...

objDataTable = New Data.DataTable

objDataTable.Columns.Add("Fecha", GetType(Date))

objDataTable.Columns.Add("Compra", GetType(Double))

objDataTable.Columns.Add("Venta", GetType(Double))

' Se manda poblar la tabla...

ObtenerDatos(objDataTable)

' Si toda ha salido bien y objDataTable tiene la data, entonces se copian los datos al buffer...

For Each objDataRow In objDataTable.Rows

SUNATTCBuffer.AddRow()

SUNATTCBuffer.Fecha = CType(objDataRow(0), Date)

SUNATTCBuffer.Compra = CType(objDataRow(1), Decimal)

SUNATTCBuffer.Venta = CType(objDataRow(2), Decimal)

Next

End Sub

Private Sub ObtenerDatos(ByRef objDataTable As Data.DataTable)

Dim objWebClient As System.Net.WebClient

Dim strHTML As String = ""

Dim strURL As String

Dim intAño As Integer

Dim intMes As Integer

Dim dtDateAux As Date

objWebClient = New System.Net.WebClient()

dtDateAux = DateSerial(Year(Now), Month(Now), 1).AddDays(-1) ' Este es el último día del mes anterior

' Primero: Obtener el TC del mes anterior

intAño = Year(dtDateAux)

intMes = Month(dtDateAux)

' La página del TC de la SUNAT acepta un Query String donde se le indica el año y el mes

strURL = "http://www.sunat.gob.pe/cl-at-ittipcam/tcS01Alias?mes=" & intMes & "&anho=" & intAño

If CargarDocHTML(objWebClient, strURL, strHTML) Then

ProcesarDocHTML(strHTML, intAño, intMes, objDataTable)

End If

' Segundo: Obtener el TC del mes en curso

intAño = Year(Now())

intMes = Month(Now())

strURL = "http://www.sunat.gob.pe/cl-at-ittipcam/tcS01Alias?mes=" & intMes & "&anho=" & intAño

If CargarDocHTML(objWebClient, strURL, strHTML) Then ' Si la función retorna falso, entonces no se procesa

ProcesarDocHTML(strHTML, intAño, intMes, objDataTable)

End If

objWebClient.Dispose()

objDataTable.AcceptChanges() ' Commit

End Sub

' Esta rutina sólo se encarga de conectarse a la web y obtener el HTML

Private Function CargarDocHTML(ByRef objWebClient As System.Net.WebClient, ByVal strURL As String, ByRef strHTML As String) As Boolean

Dim bAux As Boolean = False

Try

strHTML = objWebClient.DownloadString(strURL) ' Esto podría fallar (debido a problemas con la red, o que el servidor esté fuera de linea, etc)

bAux = Not (InStr(strHTML, "No existe Informaci") > 0) ' Esta función retorna False si no hay datos

Catch ex As Exception

' Si hay una excepción no se hace nada

End Try

Return bAux

End Function

' Esta rutina "procesa" el HTML como String

' Lo que hace es recorrer el código HTML de la página y obtener de allí subcadenas...

Private Sub ProcesarDocHTML(strHTML As String, intAño As Integer, intMes As Integer, ByRef objDataTable As Data.DataTable)

Dim strTable As String

Dim strFilas As String()

Dim strCeldas As String()

Dim intFirstPosition As Integer

Dim intLastPosition As Integer

Dim i As Integer

Dim j As Integer

Dim dtFecha As Date

Dim dblCompra As Double

Dim dblVenta As Double

intFirstPosition = InStr(strHTML, "form-table") ' Primero se obtenie la posición de la cadena "form-table"

intFirstPosition = InStr(intFirstPosition, strHTML, "<tr>") - 1 ' Una vez que el elemento table se ha localizado, se avanza el puntero hasta la primera fila

' En una tabla HTML, las filas se identifican con <tr>

intLastPosition = InStr(intFirstPosition, strHTML, "</table>") - 7 ' Se obtiene la posición del elemento de cierre "</table>"

strTable = strHTML.Substring(intFirstPosition, intLastPosition - intFirstPosition) ' Se extrae una subcadena que contiene todas las filas de la tabla

strFilas = Regex.Split(strTable, "</tr>") ' Se usa Split() para obtener un arreglo de filas. Ojo que la primera fila contiene las cabeceras de la tabla

For i = 1 To strFilas.GetUpperBound(0) ' El bucle empieza desde 1 en vez de 0 para saltearse la fila de cabeceras

strCeldas = Regex.Split(strFilas(i).Trim, "/td>") 'Se usa Split() para obtener un arrego de celdas

' ...............................................Al dejar el caracter '<' se puede delimitar luego el contenido de la celda

' strCeldas tiene TODAS las celdas de una fila, pero se necesitan leer las celdas 3 a la vez...

For j = 0 To strCeldas.GetUpperBound(0) Step 3 'Las celdas siempre aparecen en grupos de 3: ("Dia", "Compra", "Venta")

If strCeldas(j) <> "" Then ' La función RegEx.Split a veces deja una cadena nula al final del arreglo. Por eso se itera sólo donde hay datos

ObtenerTriplete(strCeldas, j, intAño, intMes, dtFecha, dblCompra, dblVenta) ' Un triplete se define como (dtFecha, dblCompra, dblVenta)

AddRow(objDataTable, dtFecha, dblCompra, dblVenta)

End If

Next

Next

End Sub

' Valores posibles para intIndice: 0, 3, 6, 9

' Un triplete es un grupo de 3 celdas, una para la Fecha, otra para el valor de Compra y finalmente otra para el valor de Venta

Private Sub ObtenerTriplete(strCeldas As String(), intIndice As Integer, intAño As Integer, intMes As Integer, ByRef dtFecha As Date, ByRef dblCompra As Double, ByRef dblVenta As Double)

Dim intDia As Integer

intDia = CInt(ObtenerValorCelda(strCeldas(intIndice)))

dtFecha = DateSerial(intAño, intMes, intDia)

dblCompra = CDbl(ObtenerValorCelda(strCeldas(intIndice + 1)))

dblVenta = CDbl(ObtenerValorCelda(strCeldas(intIndice + 2)))

End Sub

' Esta función extrae el Valor del elemento <td>

Private Function ObtenerValorCelda(strCelda As String) As String

Dim strValorCelda As String

Dim intFirstPosition As Integer

Dim intLastPosition As Integer

intFirstPosition = InStr(strCelda, "<strong>") ' El elemento "<strong>" sólo se encuentra en la primera columna del triplete ("Dia")

If intFirstPosition > 0 Then

intLastPosition = InStr(strCelda, "</strong>")

intFirstPosition = intFirstPosition + 7

Else ' Si el elemento "<strong>" no se ha encontrado, entonces estamos en la columna "Compra" o "Venta"

intFirstPosition = InStr(strCelda, "tne10") ' la cadena 'class = "tne10"' sólo se encuentra en las columnas "Compra" o "Venta"

intFirstPosition = intFirstPosition + 7

intLastPosition = InStr(intFirstPosition, strCelda, "<") ' El caracter '<' fue dejado intencionalmente cuando se usó RegEx.Split() para poderse usarsele como marcador aquí

End If

strValorCelda = strCelda.Substring(intFirstPosition, intLastPosition - intFirstPosition - 1).Trim ' Después de obtenerse las posiciones, se extrae la subcadena del valor

Return strValorCelda

End Function

' Agrega una fila en objDataTable

Private Sub AddRow(ByRef objDataTable As Data.DataTable, dtFecha As Date, dblCompra As Double, dblVenta As Double)

Dim objDataRow As Data.DataRow

objDataRow = objDataTable.NewRow

objDataRow(0) = dtFecha

objDataRow(1) = dblCompra

objDataRow(2) = dblVenta

objDataTable.Rows.Add(objDataRow)

End Sub

End Class