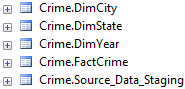
Crime Data Mart Design:

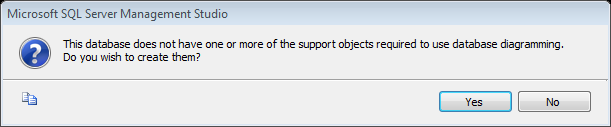
# Download the Data Set

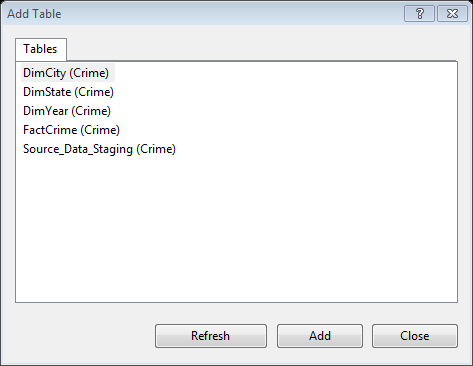
* Download and extract the contents of the “**Crime-Data-Mart-Files.zip**” archive from this GitHub Repository. Navigate in the unarchived directory and locate the “2006 – 2008 Crime in the United States (Data.gov)” data set.
* Save the .csv file in a location that you have full rights to edit (ex. My Documents, Desktop).

# Build the Data Mart Design

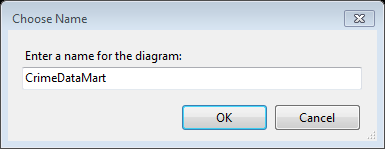
* Open Microsoft SQL Server Management Studio and connect to your local instance of SQL Server.
* Create a new Database called “Crime”
  + Right-click on the “Databases” folder and choose “New Database …” from the secondary menu drop-down. This will open the “New Database” dialog
  + In the “Database” field type “Crime”. Leave all other fields with their default entries and click the “OK” button to create the database.
  + You should now see the “Crime” database created under the “Databases” folder of your server: 
* Create a new Schema called “Crime” to hold all objects created for this Data Mart.
  + Expand the “Crime” database. Expand the “Security” folder.
  + Under the “Security” folder, Right-click on the “Schemas” folder and choose “New Schema …” from the secondary menu drop-down. This will open the “Schema – New” dialog.
  + In the “Schema name” field type “Crime”. Leave all other fields with their default entries and click the “OK” button to create the Schema.
  + Expand the “Schemas” folder and you should now see “Crime” listed as an available Schema in the list: 
* Create 5 tables using the 5 SQL scripts included with this document:
  + Using SQL Server Management Studio, click on the “File” 🡪 “Open” 🡪 “File” option from the secondary menu drop-down.
  + Navigate to the directory where you have downloaded the 5 SQL scripts and choose the “CrimeDimState.sql” script. **NOTE:** **Be sure to execute these scripts in the order described in this document.**
  + Now execute the rest of the scripts in the following order:
    - CrimeDimCity.sql
    - CrimeDimYear.sql
    - CrimeFactCrime.sql
    - CrimeSourceDataStaging.sql
  + Expand the “Tables” folder under the “Crime” database and you should see the 5 tables created:



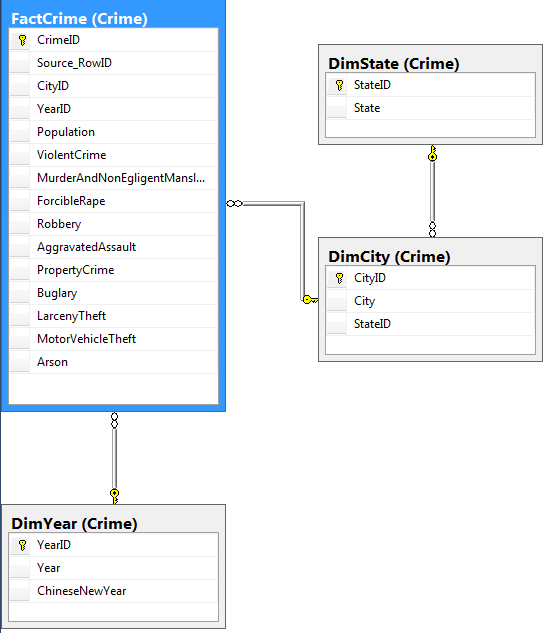
* Create a “Database Diagram” so that you can see the Data Mart Design
  + Under the “Crime” Database, right-click on the “Database Diagrams” folder and choose “New Database Diagram”. You may receive this message: 
  + Click “Yes” to this message if it is displayed. This will open the “Add Table” dialog.
  + Select one table at a time listed in the “Add Table” dialog (**NOTE: Do NOT add the “Source\_Data\_Staging” table**) and click the “Add” button to add them to the diagram:



* + Click the “Close” button once you’ve added all tables (except the “Source\_Data\_Staging” table).
  + Now click the “Save” button (  ) in the upper left hand side of the upper menu bar. This will open the “Choose Name” dialog.
  + Type “CrimeDataMart” in the “Enter a name for the diagram” field:

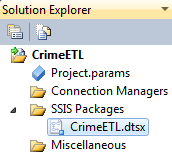


* + Click the “OK” button to save the diagram.
  + Your diagram should look like this (see next page):

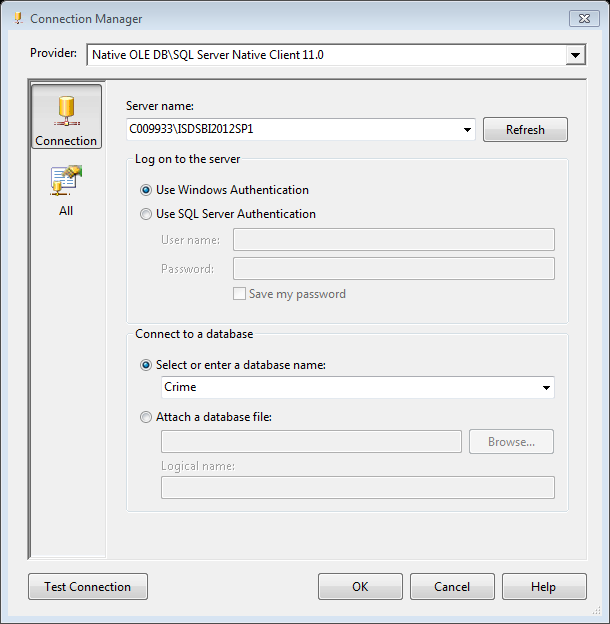


# Execute the ETL Package

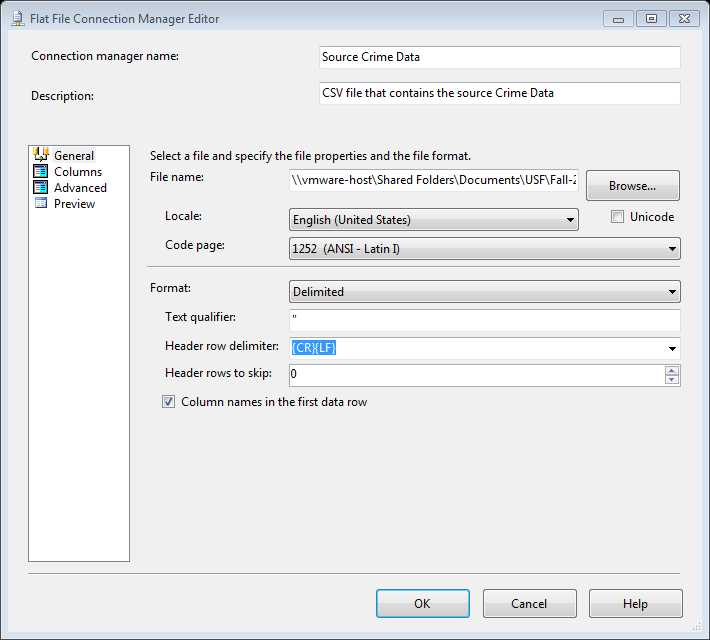
* Be sure you have downloaded and “unzipped” the “CrimeETL.zip” folder.
* From your Programs menu expand “Microsoft SQL Server 2012” and launch SQL Server Data Tools (SSDT)
* Once you have launched SSTD click on the “Open Project” link: 
* Navigate to the “CrimeETL” Solution file: 
* Click “Open” to load the “CrimeETL” Solution in SSDT.
* Once the Solution is open, double-click on the “CrimeETL.dtsx” file located under the “SSIS Packages” section of the “Solution Explorer”:



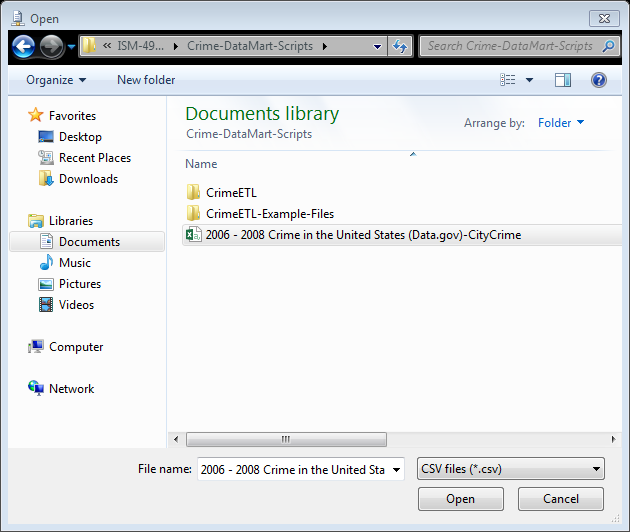
* The “CrimeETL” package will then load within SSDT in “Design” mode.
* Update the “Connection Managers” within the package:
  + Locate the SQL server connection called “localhost.Crime” in the “Connection Managers” Section: 
  + Double-click the “localhost.Crime” connection to load the “Connection Manager” dialog.
  + While in the “Connection Manager” dialog change the “Server name” field to the name of your local SQL Server Instance. If you’re not sure what to type as your “Server name”, type the exact name of the server listed in the “Object Explorer” of SQL Server Management Studio.
  + If you have the correct “Server name” listed in the field, you should be able to click the drop-down in the “Select or enter a database name” field and choose the “Crime” database.
  + Here is what your “Connection Manager” should look like (your Server name will be different):



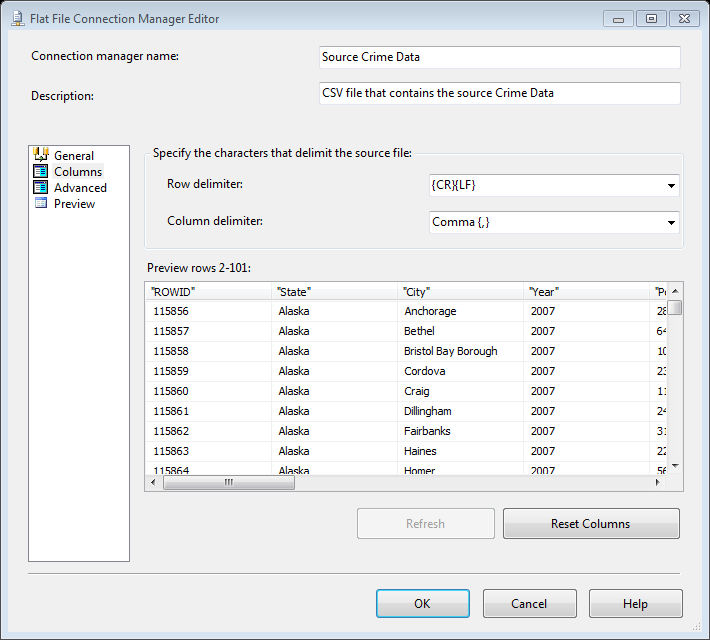
* + Click the “Test Connection” button to ensure that you are able to connect to the “Crime” database. Then click the “OK” button to close the dialog.
  + Now you need to double-click the open the “Source Crime Data” data source in the “Connection Managers” section and update its connection to the “2006 - 2008 Crime in the United States (Data.gov)-CityCrime.csv” file. This will open the “Flat File Connection Manager Editor”.



* + Click the “Browse” button to locate your local csv file that contains the data.



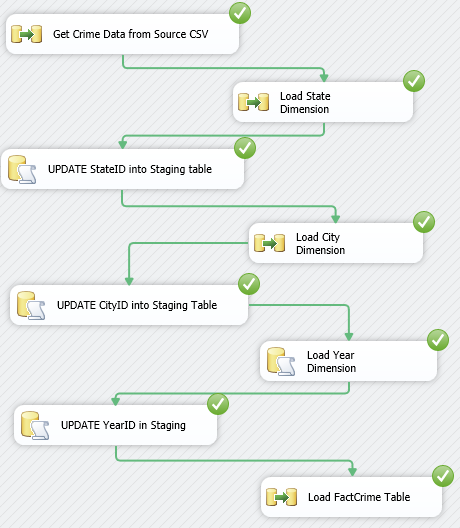
* + While browsing for your .csv file, be sure to change the file type drop-down in the lower right-hand corner of the dialog to “CSV files (\*.csv)”.
  + Choose your .csv file and click the “Open” button.
  + If you receive a warning dialog after clicking “Open”, just agree to the dialog.
  + While still in the “Flat File Connection Manager Editor” click the “Columns” tab on the left side of the dialog. Review the Columns listed.



* + Click “OK” button to close the “Flat File Connection Manager Editor”.
  + Your “Connection Managers” are now set.
* At the bottom of your project in the “Error List” section of Visual Studio you may see 2 Warning similar to the following. Ignore these warnings. They are displayed because you are re-using an SSIS package that was created on another PC.



* The “CrimeETL” SSIS package is now ready for execution.
  + Locate and click on the “Start Debugging” button in the menu bar at the top of Visual Studio: 
  + Wait for the SSIS package to complete all steps. Upon completion your screen should look like this:



# Test your Data Mart

* Open SQL Server Management Studio and connect to your local SQL Server Instance.
* Click on the “New Query” button found at the top menu of the interface to open a new query tab in Design mode: 
* Paste and Execute the following code in your query tab:

|  |
| --- |
| USE Crime;  SELECT  CrimeState.[State],  CrimeCity.City,  CrimeYear.ChineseNewYear,  Crime.Robbery  FROM Crime.FactCrime AS Crime  INNER JOIN Crime.DimYear AS CrimeYear ON Crime.YearID = CrimeYear.YearID  INNER JOIN Crime.DimCity AS CrimeCity ON Crime.CityID = CrimeCity.CityID  INNER JOIN Crime.DimState AS CrimeState ON CrimeCity.StateID = CrimeState.StateID  WHERE CrimeState.[State] = 'Florida'  AND CrimeYear.ChineseNewYear = 'Rat'  ORDER BY CrimeCity.City ASC |

* You should see 298 Rows of data return that look like this:

