

Automating Workforce Information Database (WID) Updates with NCSC Data

The National Crosswalk Service Center (NCSC) hosts over 3,000 files for the WID. Not only do you have all these files to sort through, but many of these files are frequently updated. Wouldn't it be great if you could push a button and have your data updated? Data Transformation Services (DTS), a utility included in Microsoft SQL Server 2000, is the tool for this job. It allows you to create a package that uses File Transfer Protocol (FTP) to download files, connect to most relevant database management systems, and executes tasks on those connections.

There are four general steps necessary to update a WID table with data from NCSC. First you need to use FTP to download the table that you need. Next, that table needs to be imported into a holding table to determine what data is being replaced in the WID table that you are updating and delete that data. Finally, you import the data into WID.

For this example, I created a DTS package that automates the updating of the national Current Employment Statistics (CES) data. Execution of a DTS package automatically executes the tasks contained within it based on conditions defined by the developer. My DTS package contains five tasks.

Task 1) an FTP task downloads the CES22.dbf table.

Task 2) An execute SQL task recreates the holding table. This is an additional step specified to this purpose.

Task 3) A transform data task between a dBase 5 connection and an SQL Server connection is used to import the data from CES22.dbf into the holding table.

Task 4) In another execute SQL task, all records in the WID CES table that match records in the holding table are deleted.

Task 5) Now we are ready to update the WID CES table with the data in the holding table. This is done with a transform data task between an SQL Server connection to the holding database and another one to the WID.

The DTS package diagram below depicts this entire process. Task 1 is labeled XWalkCenter. Task 2 is labeled Recreate CES22. Task 3 is depicted by the arrow between the dBase 5 and the CES connections. Task 4 is labeled Delete from CES. And Task 5 is depicted by the arrow between the CES and the WID connections. After the successful executions of Tasks 1 and 2, Task 3 executes. If Task 3 executes successfully, then Task 4 executes. Finally, Task 5 executes, if Task 4 executed successfully. If any task fails, the process halts.

Setting up this package requires more initial work, but the benefits are obvious. Additionally, once you set one package up, it's easily modified for use on a different table. If you would like to implement this tool in your state and would like assistance, please contact me at SJones2@state.wy.us.

