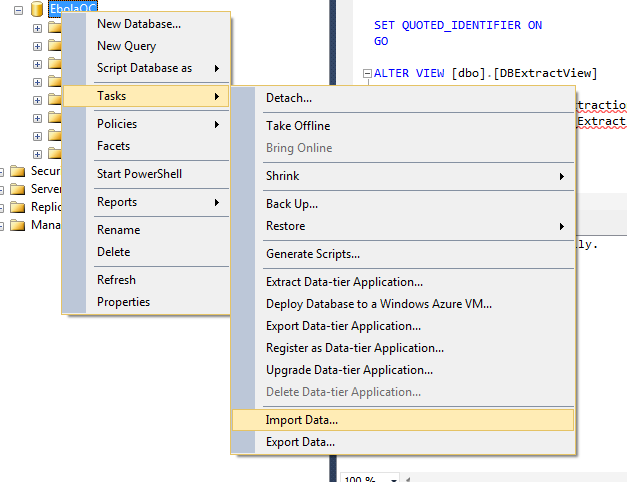
# SQL Data load instructions

V0.1 21/03/2015 – Brian Lee

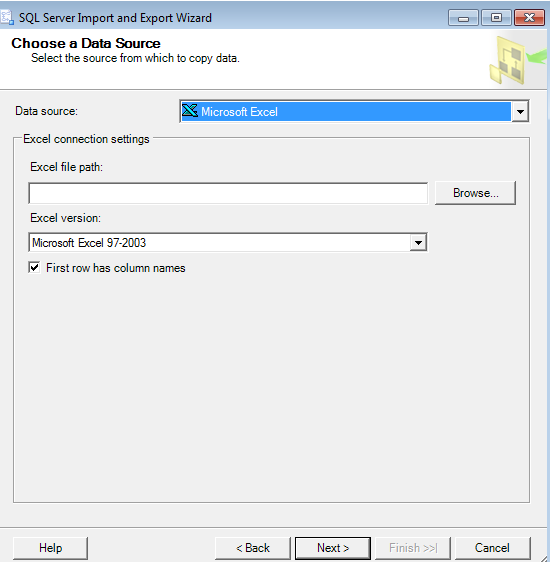
Instructions for loading the VHF daily extract excel spreadsheet into SQL for analysis using SQL/T-SQL. Similar import processes could be used for SAS, Stata or R if you prefer to use these tools instead of SQL.

Note: this document contains no sensitive information and all examples use synthetic or fake data. For accurate data, please refer to the [WHO Sitreps](http://www.afro.who.int/en/clusters-a-programmes/dpc/epidemic-a-pandemic-alert-and-response/sitreps.html).

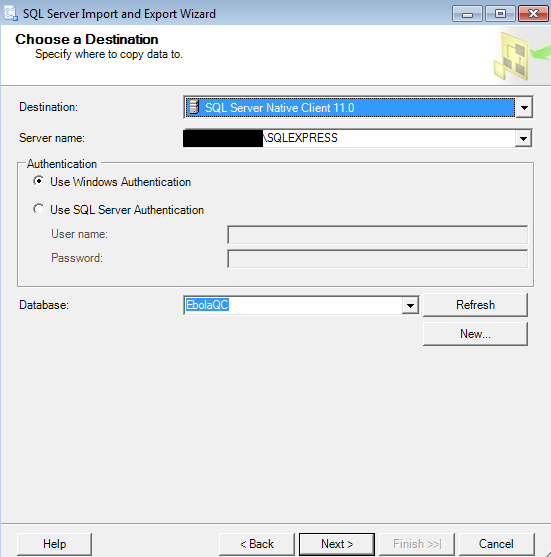
1. Start with VHF extraction in Excel format
2. Remove fields not needed. There is a limitation/bug in SQL (and SAS) that restricts the number of usable fields imported from Excel to 255. So get the fields in the extract to 255 or less. I wrote a macro to do this that is available in Dropbox (GUINEE\14. Analyse\prepExcelForSAS) and Github.
3. Install a version of SQL Server on your laptop ([SQLServer 2014 Express](http://www.microsoft.com/en-us/download/details.aspx?id=42299) is free and what I use). Note you could use any sql database (sqllite, mysql, etc.), I just already had this one installed.
4. Work with SQL using the SQL Server Management Studio (accessible by Start|All Programs|Microsoft SQL Server 2014|SQL Server 2015 Management Studio)
5. Create a database with all default parameters. Only you will use this database. Name it whatever you want, I chose “EbolaQC.” Replace “EbolaQC” in all scripts and instructions with the name you choose.
6. Import the Excel daily extract subset using the SQLServer data import task
   1. Launch the SQL Server Import and Export Wizard: Right click on your database, choose Tasks, then choose Import Data



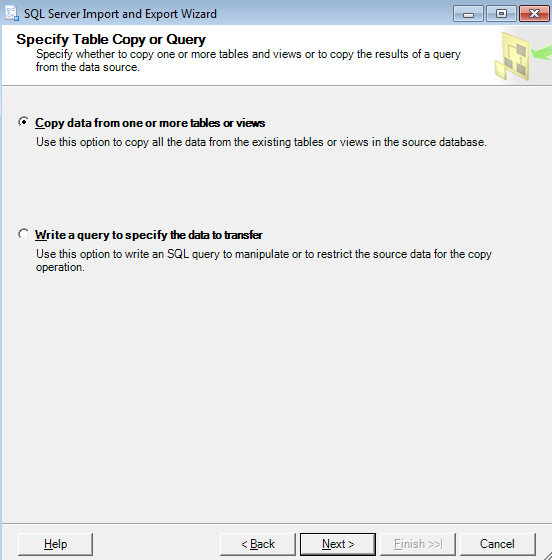
* 1. Step 1: Choose a Data Source, select “Microsoft Excel” in the data source dropdown. Browse to the extract subset file from step 2 in the Excel File Path field. Choose “Microsoft Excel 97-2003” for the Excel version. Click Next.



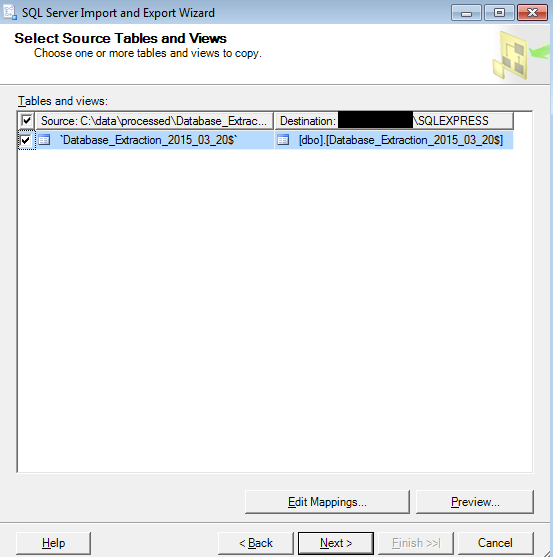
* 1. Step 2: Choose a Destination, select “SQL Server Native Client 11.0” in the Destination field. Select your laptop name plus “\SQLEXPRESS” in the server name field. Choose “Use Windows Authentication” and select your database name. Click Next.



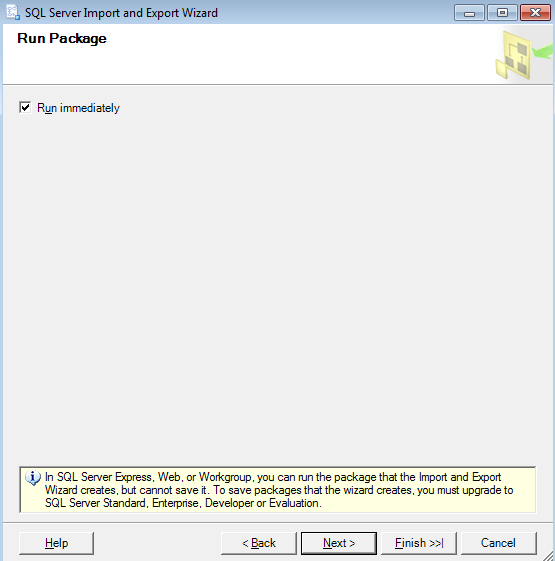
* 1. Step 3: Specify Table Copy or Query, choose “Copy data from one or more tables or views.” Click Next.



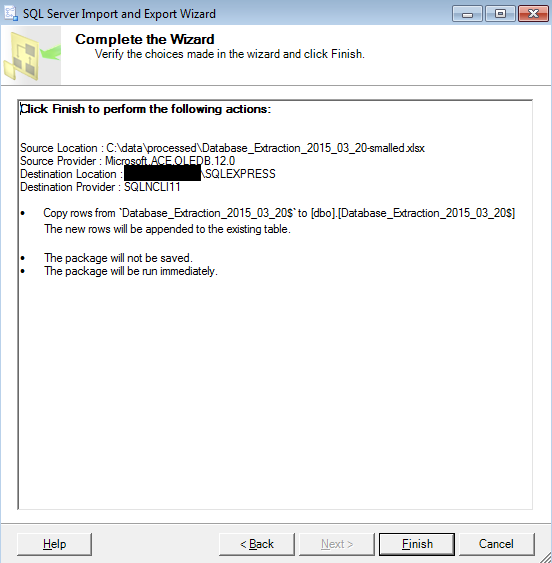
* 1. Step 4: Select Source Tables and Views, check the box next to source and destination. Click Next.



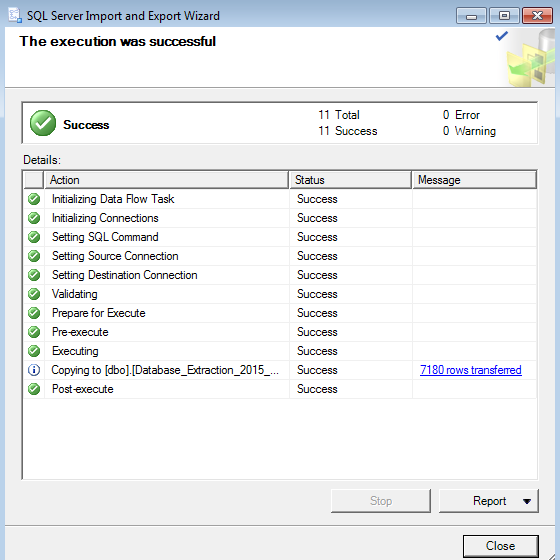
* 1. Step 5: Run Package, check the “Run immediately” checkbox. Click Next.

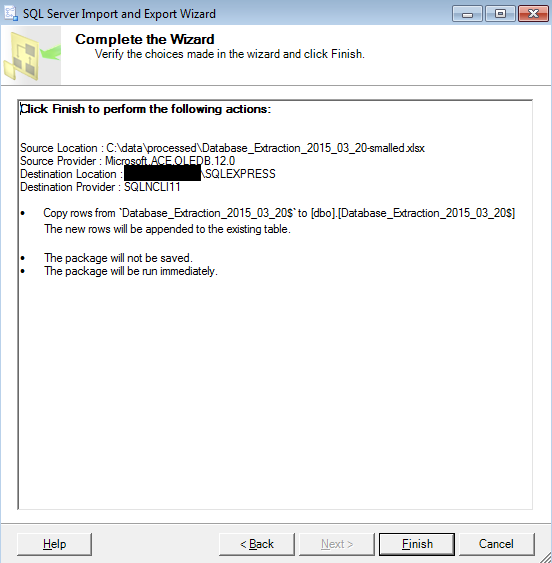


* 1. Step 6: Complete the Wizard (final). Click Finish to run import.



* 1. Complete. This is what a successful import looks like. If something goes wrong, email [balee3@cdc.gov](mailto:balee3@cdc.gov).





1. Run your scripts as desired against the new table created. I found it useful to create a view called “DBExtractView” so I don’t have to rename my scripts with each new import table. Each day I alter the view using “ALTER VIEW DBExtractView AS SELECT \* FROM [Name of Extraction Table]”
2. All of my example scripts at <https://github.com/leebrian/ebola/tree/master/sql> for reference. Note, no sensitive information should ever be stored in source code and these scripts are shared only as an example.