Data File Handling Suggestions

File Type and Size

A large volume of data is included in this publication in a single comma-separated file (CSV) file. In an effort to assist the user, software and file handling suggestions are presented below.

File Handling

The size of the data file(s) may be large so data users are urged to first save the file to their computer and then open using one of the options presented below. The CSV data files can be imported into Microsoft Excel or to a database program such as Microsoft SQL for use. Current versions of Excel have row limitations of 1.04 million rows per worksheet. In order to import directly into Excel, the data file will need to be split into manageable blocks prior to import. Alternatively, the Excel add-in "PowerPivot" can be used with Excel 2013 to import the data file without splitting. Tips for handling the CSV file using these software options are described below.

The steps are:

1. Save zip file to a local hard drive
2. Open zip file from the hard drive
3. Unzip and extract the files to a local hard drive

Option 1: Split the CSV

Steps to split the CSV file for use with Spreadsheet Software

1. Determine the maximum number of rows supported by your version of Excel or other spreadsheet software.
2. Use a CSV Splitter to split the CSV file into smaller CSVs (ex. Free Huge CSV Splitter software is available at http://sourceforge.net/projects/splitcsv/).
3. CSV Splitter will produce multiple files with a sequential value appended to the end of the file name. Only the first file will contain the header row.
4. Open the first CSV file in Excel or other spreadsheet software.
5. Copy and paste the header row from the first CSV file into each of the other split CSV files.

Option 2: Load to Database Server

Steps to import the CSV file to a Database Server (ex. Microsoft SQL)

1. Create an empty "Chemistry Data" table with the desired column types (attached is an example create table script for users with a database server such as SQL).
2. Import the CSV file using the SQL Import Wizard with a field DataType of string [DT_STR] (255) for all the columns to a new table. Tip: if you select multiple columns in the advanced dialog, you can change their width at the same time.
3. Insert into the "Chemistry Data" table all the rows from the imported table to covert the columns to the appropriate data types.

Option 3: Link CSV file to Microsoft Access database

Steps to link the CSV file to an Access database

1. Create and open a new blank database in Access
2. Select External Data>Text File (in Import & Link section)
3. Specify the CSV file as the data source and select the "Link to the data source by creating a linked table."
4. Select Next
5. Toggle "First Row Contains Field Names" and select Advanced options
6. Specify date order as "YMD" and update Data Types within the Field Information section.
7. Select OK and then finish
8. CSV file is now linked to Access database and can be queried.

Option 4: PowerPivot

Steps to load the CSV file to PowerPivot for Excel 2013

1. Activate the Add-in "PowerPivot" in Excel 2013.
2. Select PowerPivot > Data Model > Manage.
3. Select Get External Data > From Other Sources.
4. Select Text File, and then click Next.
5. Type in the File Path, and then click Finish.

Note: Excel may incorrectly assign format types to data contained in the CSV files (e.g., CAS No). Data users are urged to review accuracy of field formatting after import into Excel.