

Excel: Importing and Exporting Data

Importing a text file into multiple columns

Types of text file

We'll be importing data from files that contain "records" in single lines of text. (If your records extend to multiple lines, importing data gets really complicated.) The different items of information about each record will either be separated in some way with spaces or punctuation marks or (less common nowadays) with a fixed number of text characters for each item.

- **Delimited**

Often used for exchanging data between programs, it can also be used to extract data from text documents. Items can be separated by spaces, tab characters, commas, or other characters. More than one type of delimiter can be used in the same line, though this is normally useful only in extracting data from documents.

- **Fixed width**

Once common in the days of keypunches, and still used by some data acquisition systems, each item occupies a fixed location in the line (i.e., characters 1-10, 11-25, 26-27, 28, etc.)

Opening a text file in Excel

- **File -> Open**
- **Right-click and choose "Open with..."**
- **Text Import Wizard**

1. **Choose file type**

- ***Delimited***

Select the characters that will delimit a line into columns (tab, space, semicolon, comma, and/or your choice of character). Should consecutive delimiters indicate blank columns (usually tab or comma delimited files) or should they be lumped together into a single delimiter (multiple spaces in text tables)?

- ***Fixed Width***

Insert break lines between data columns by clicking on the ruler; remove them by double-clicking. Drag the line to move it.

2. Choose a data format each column.

- **General**
Converts numbers into numbers, date values to dates and everything else into text.
- **Text**
Converts everything into text even if it looks like a number or a date.
- **Date**
Converts the column to date values using the format from the drop-down box. Good when the date is a numeric (like MMDDYYYY) that might be ambiguous.
- **Do not import columns**
Skip these columns of the line altogether. Good for removing data you don't really need.

3. Click Finish.

Parsing data already in the worksheet

1. Highlight the data
2. Choose Data -> Text to Columns
3. See the rules above for the Text Import Wizard

Concatenating data from multiple columns into a single column

In a cell to one side at the top of the data, use the **CONCATENATE** function to group multiple columns in one cell

Either:

=CONCATENATE(A4,B4,C4,D4,E4,F4,G4)

Or:

=A4&B4&C4&D4&E4&F4&G4

Then:

1. Use **Edit -> Fill -> Down** to copy that cell to the entire column
2. With the column highlighted, choose **Edit -> Copy**.
3. Without moving the cursor, choose **Edit -> Paste Special** and select **Paste Values**.
4. (Optional) Delete the original data from the sheet.

Saving as alternate formats

File -> Save As...

Text (tab delimited) (.txt)

This writes each row of the worksheet out to a (potentially very long) line in a text file with the columns separated by tab characters. Common imported format for many programs that can't handle Excel files directly. There are also special versions for Macintosh (where the end-of-line character is different) and MS-DOS (which uses a slightly different character set than Windows).

CSV (comma delimited) (.CSV)

This writes each row of the worksheet out to a (potentially very long) line in a text file with the columns separated by commas. If a text item contains commas, it's surrounded with double quotation marks. This is the format generated by the **WRITE** command and read with **INPUT** in Visual BASIC, QuickBASIC and good old GWBASIC. Many programs accept these as input.

Formatted Text (Space Delimited) (*.prn)

This writes the worksheet out as a text file, but uses multiple spaces to attempt to maintain the appearance of the worksheet. Nice for pasting into text-only email.