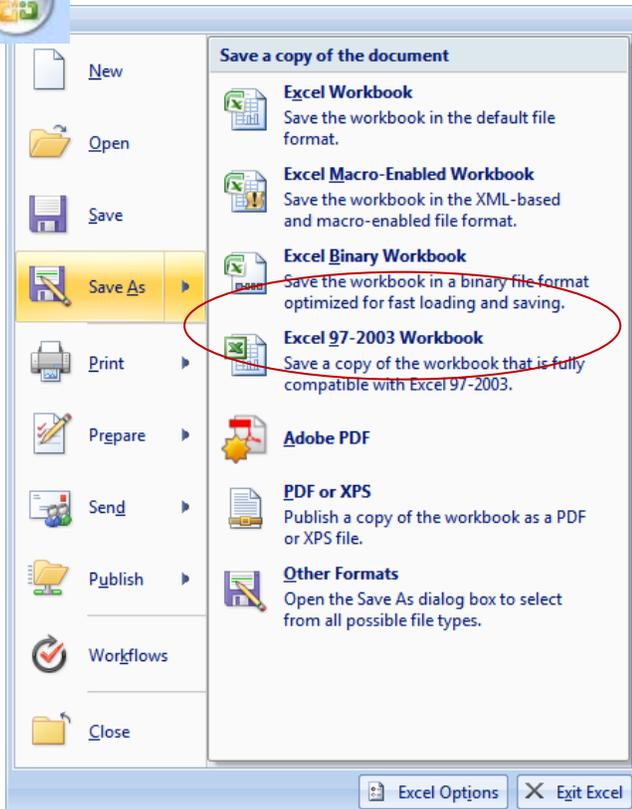# Best Practices –Preparation of Excel Data for Use in GIS

File Types

Excel files in the Office 2007 format (.XLSX) are not recognized in ArcView. However you can save these back to the older version.
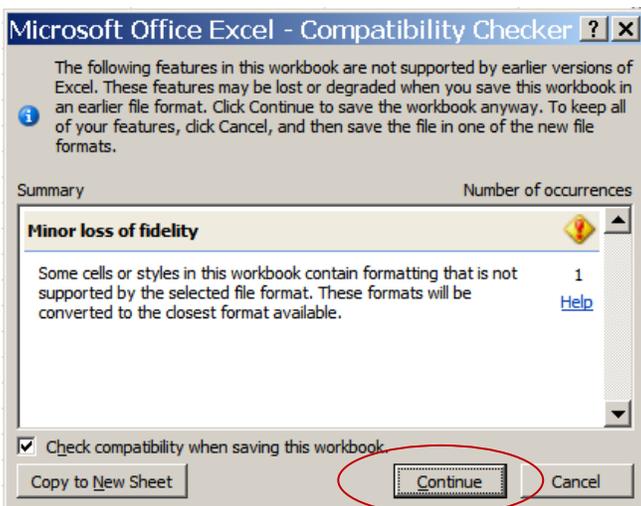
Click the Icon in the upper left to activate the menu options.

Select Save As > Excel 97-2003 Workbook

This will allow you to save the file back to the format (XLS). Specify a file name that is the same (or very similar) to the original file name.

You may get a possible compatibility problem; you should just click Continue from this dialog.
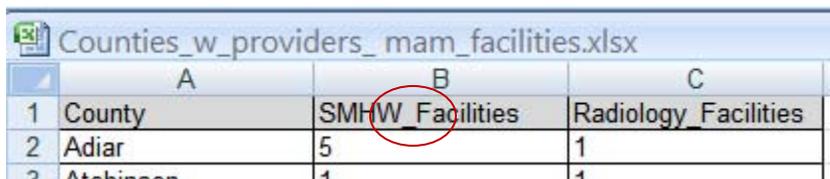
File Preparation - Excel

Here are some guidelines for making your Excel file 'GIS ready'. If you follow these guidelines

The most common problem with Excel files when used in ArcMap is a poorly crafted column heading.

- NO spaces in the column names (you can use capital letters to distinguish words. Example: PctTotal02)
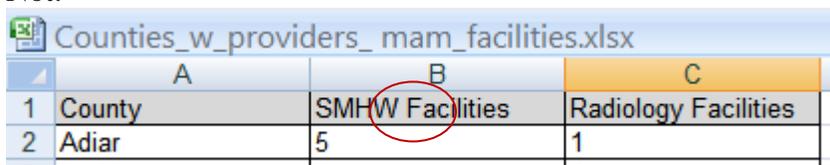
- NO punctuation except for underscore (_).

Examples:
For column headings, this:

	A	B	C
	Counties_w_providers_ mam_facilities.xlsx		
1	County	SMHW_Facilities	Radiology_Facilities
2	Adiar	5	1
3	Atchison	1	1

Not:

	A	B	C
	Counties_w_providers_ mam_facilities.xlsx		
1	County	SMHW Facilities	Radiology Facilities
2	Adiar	5	1

In addition, keep these best practices in mind:

- Keep a copy of your original file.
- Use only one line for your column headings, don't create 'stacked' headings
- The first line must contain ONLY the names of the fields; brief headings are preferable.
- There should be NO additional header, footer or sidebar comments in the document. This includes summary lines, totals, references, citations, etc.
- Headings should start with a letter, not a number.
   Example: 2002rate should be changed to rate02 or rate2002
- Brief headings are preferable.
- Field names should be in all UPPER case or all mixed/proper; either is OK, as long as it is consistent.
- The same consistency applies within the data. Text fields should contain data in all UPPER case (e.g. COLE, OSAGE, BOONE) or all mixed/proper case (e.g. Cole, Osage, Boone).

- For number data do not use 1000 separators. This can be changed under Format Cells > Number Uncheck 'Use 1000 separator (,)'.
- If you are using a whole number, set the number of decimal places to 0. This can be changed under Format Cells > Number. Use the down arrow next to number found in 'Decimal places:' or key in '0'.
- For character/text data, make sure there are no leading blanks at the front of the data. Text should appear justified to the left, without spaces before the first character.
- Do not use calculated cells. These do not transfer properly to other file formats. If necessary, create a new column and use the Paste Special command to copy only the Values. An example of this is shown later in this document.
- **Bold** text, boxes around cells and other cosmetic formatting can be left as-is. This formatting will be ignored.
- **Data that was copied from a web page** should always first be saved in the TXT file format. This will remove any formatting (HTML) that was copied along with the data. You can then save the text file back into Excel again.

Here is what your file should look like after you are done editing it:

BEFORE

	A	B	C	D	E	F	G	H	I
1	AGE OF THE POPULATION 60 YEARS AND OVER [100]						*Source: U.S. Census - 2000*		
2		Universe: Population 60 years and over							
3	[Based on a sample. Rounded data. Geographic areas are omitted if they do not meet the specified population threshold.]								
4	Geographic Name	Total:	60 to 64 years:		65 to 84 years:		85 years and over:		
5			Total	Percent of ≥ 60 Population	Total	Percent of 65 - 84 Population	Total	Percent of ≥ 85 Population	
6									
9	Adair	3850	800	0.2078	2590	0.6727	460	0.1195	
10	Andrew	3000	630	0.2100	2070	0.6900	300	0.1000	
11	Atchison	1640	290	0.1768	1090	0.6646	260	0.1585	
12	Audrain	5570	1130	0.2029	3870	0.6948	560	0.1005	
13	Barry	7280	1800	0.2473	4860	0.6676	610	0.0838	
14	Barton	2650	580	0.2189	1770	0.6679	300	0.1132	

AFTER

	A	B	C	D	E	F	G	H
1	COUNTY	TOTPOP	AGE60_64	PCT60_64	AGE65_84	PCT65_84	OVER85	PCTOVER85
2	ADAIR	3850	800	0.2078	2590	0.6727	460	0.1195
3	ANDREW	3000	630	0.2100	2070	0.6900	300	0.1000
4	ATCHISON	1640	290	0.1768	1090	0.6646	260	0.1585
5	AUDRAIN	5570	1130	0.2029	3870	0.6948	560	0.1005
6	BARRY	7280	1800	0.2473	4860	0.6676	610	0.0838
7	BARTON	2650	580	0.2189	1770	0.6679	300	0.1132
8	BATES	3690	790	0.2141	2510	0.6802	390	0.1057

Here are some things to watch for in files of county data:
- Suggested field name for county names is County or COUNTY.
- Make sure there are no misspellings in your county names. The most common misspellings are:

    Daviess – sometimes has the 2nd 'S' missing

    DeKalb, McDonald – The 'K' and 'D' should be capitalized

Ste. Genevieve – sometimes is missing the 'E' in Ste.

- Do not include the word 'County' in the name of the county.
- For St. Louis – it is St. Louis and St. Louis City – you should not add the word 'County' after St. Louis to denote the county of St. Louis.
- All of the county names that start with Saint should be abbreviated as St. (be sure and get the period at the end). The abbreviation for Sainte Genevieve is Ste. (feminine form of the saint abbreviation).
- If you have the Federal Information Processing Standards or FIPS code, this may be used instead of the county name.
- St. Louis City is equivalent to a county for Federal and State government purposes. Consequently, St. Louis City is both a city and a county. Kansas City is NOT a county or equivalent. Kansas City actually falls within four separate counties
-

Here is something to watch for in files of ZIP code data:

ZIP codes should be stored as text field types. The same goes for any ID code that has only numbers and may have a zero as the first character, such as three digit county FIPS codes. The leading zero will be lost if the field is stored as a numeric field (integer, float, single or double).

If you want to change all your data to the same case (either mixed or upper)

Converting data, such as county names, to all upper case is one way to create consistency in your data set. This conversion can be accomplished with a function in Excel.

Create a new empty column next to the column you need to convert. Select the column to the left of the column you need to convert, click Insert > Columns. A column will be added to the file.

Create a function for this new column. Select the first empty cell at the top of the new column. Use the = sign (next to the place where you can type at the top). Type the function:

Copy and paste this function from the first cell into the rest of the cells in the empty column. This will

change the format of the data to all upper case. This column will only exist temporarily.

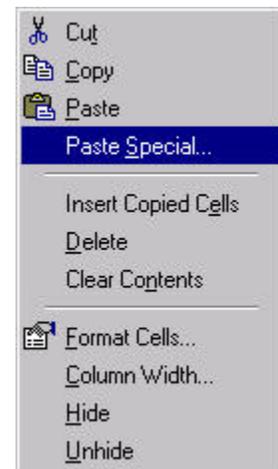|	A	B	C	D
1	COUNTY	COUNTY		TOTAL
2	Adair	ADAIR		3850
3	Andrew	ANDREW		3000
4	Atchison	ATCHISON		1640
5				

Create another empty column next to the column you just created.

You will now copy the column containing the calculated functions to this new empty column. Right-click on the column heading of the calculated function (in the example, column B), select Copy.

Right click on the new empty column (in the example, column C) and Select Paste Special.

You only want to copy the values for the county names, not the formula used to create them.

Make sure to click the choice for Values in the dialog. (This is not the default.)

C2		=	ADAIR	
	A	B	C	D
1	COUNTY	COUNTY	COUNTY	TOTAL
2	Adair	ADAIR	ADAIR	3850
3	Andrew	ANDREW	ANDREW	3000
4	Atchison	ATCHISON	ATCHISON	1640

This will look identical to you, and visually they are the same. The difference is that one column (column B in this case) contains a function and not the actual data. Column C contains only the text values.

	A	B	C	D
1	COUNTY	COUNTY	COUNTY	TOTAL
2	Adair	ADAIR	ADAIR	3850
3	Andrew	ANDREW	ANDREW	3000
4	Atchison	ATCHISON	ATCHISON	1640
5				

To complete the conversion, you must delete the original unformatted column and the temporary column.
Select the original column and the temporary column. Select Edit > Delete.

	A	B
1	COUNTY	TOTAL
2	ADAIR	3850
3	ANDREW	3000
4	ATCHISON	1640
5		

This will leave only one column of properly formatted values.

If you need to convert UPPER case to Proper case, the steps are very similar. Use the function Proper rather than Upper, and otherwise follow the same steps above.

= =proper(a1)

When converting from UPPER to Proper, you must go back and fix DeKalb, McDonald, making sure the 'K' and the 'D' are capitalized.

Created by Tracy Schloss
Last Updated 12/9/2008