**Stat 587: Hints & tips for downloading and reading data files**

One of the hardest things in this course is getting data files downloaded and read into your choice of software. This is a browser and operating system problem, not a statistical problem. Various folks struggled with this in lab. Mostly the problems seem to be with Windows; I may not have noticed similar Mac issues because many fewer folks were using Macs.

Contributing factors are:

by default, Windows does not display the file extension (the .txt, .csv, .r, …) with the file name

only the icon indicates the file type

by default, browers displaying text want to save the window contents as a .txt file type

even if the original file (on the class web page) has a .r extension

the software program is looking for a file in one folder; the browser saved it somewhere else

Here are some hints and tips to make reading files work more smoothly. Once you find a process that works for you, this won't be a problem for the rest of semester (and longer).

**file extension issues:**

Most (all?) browsers provide two ways to download a file from a web site:

1) click the link to file to display it in the browser window.

Then use file/save or right-click / save page as to save the file.

By default, this will want to save the file as text. Look right below the file name box.

That shows File type or Save as Type (browser-specific label) Text (or .txt)

even if the original file had a different extension

If you save the file this way, creativity.r gets saved as creativity.r.txt

and since the operating system doesn't show the file type explicitly, all you see is creativity.r

Fix: if you use file/save or save page as, change the file type option to all files.

that stops the browser from adding .txt to the file name.

The file is then saved with its original name, which includes the correct extension

2) A simpler fix is to use the second way to download a file

Don't left click on the link and view the file.

Instead, right click on the link. That menu includes Save link as. Select that option.

You get a file save box again, but now the file type is (by default) set to the correct type.

So creativity.r appears with the file type option set to \*.r (which is the correct one for R code)

**Folder issues:**

I strongly recommend you create a folder to hold all your stat 587 work.

Not necessary, but useful in the future and very useful if you have lots of stuff in my documents.

JMP, R, and SAS all have the concept of a working directory.

this is default location when you want to read a file (either data or code)

and (usually) the default location when you store results

Setting the working directory to your stat 587 folder will simplify finding files when you need them

Specifics:

Saving a file using your browser:

make sure you navigate to your stat 587 folder before saving the file

Setting the working directory (details depend on your software):

JMP:

when you use File / Open and navigate to a folder, that folder becomes your working directory.

SAS:

Look at the ribbon on the very bottom of the SAS window

You will see the system path to a folder, something like c:/Users/SOMETHING,

where SOMETHING is your name on that computer, probably your netid

on my laptop, it is c:/Users/pdixon

Sometimes this is towards the center-right of the ribbon; sometimes on the very left

If you hover the mouse over that folder name, you will (should) see Change current folder

Double-left-click on that folder name and you will get a dialog box that allows you to select

the desired folder.

If you set the current folder to your Stat 587 folder immediately after you start SAS,

then your stat 587 folder will be the default one.

The next time you start SAS, you will need to reset the default folder.

R:

The default working directory is your My Documents folder.

The path to this folder depends on the operating system.

on my Windows 7 machine, it is C:/Users/pdixon/Documents

To change this to your stat 587 folder, there are two ways:

1) If your stat 587 folder lives in your My Documents folder:

type setwd('Stat 587') into the command window.

replace Stat 587 by the name of your stat 587 folder

This works because setwd() starts with the current working directory

so it finds My Documents/Stat 587.

2) Or use the menu options to select any directory as your working directory

Either: file / change dir, or type setwd(choose.dir())into the command window

both bring up a dialog box to choose your working directory

Then navigate to the desired folder, starting at the very top of the file system.

R is operating system agnostic, so you have to drill through folders you usually don't see.

Most folks will start with c:, then choose Users, then choose you, then My documents,

then (and this is what you're really trying to do) choose your stat 587 folder

Other operating systems put My Documents in another place, so the drill down is different.

RStudio:

The default when you start RStudio is My Documents

Again, multiple ways to set the working directory:

1) If you want a folder under My Documents, you can use 1) above: setwd('Stat 587')

where Stat 587 is your name for your Stat 587 work

2) Use Session / Set Working Directory / Choose Directory to choose your directory

This opens in the current working directory (probably My Documents), so it is faster to

navigate to the folder you want.

3) setwd(choose.dir()) also works; it is R operations and RStudio is just an interface.

but that starts at c: so it requires more clicking to drill down to My Documents then to the

the folder you want.