Software Carpentry Workshop

Oct 18, 2017

Lumbers 306

Instructors: Remi Daigle and Stephen Childs

Unix

* Computers do 4 things: calculations, display output, input, talks to other computers
* Shell is related to the input and output
* Shell is a read/evaluate/print loop (REPL)
* Shell runs other programs
* Bash is the name of the program that’s running (most common)
* Commands in UNIX
  + Whoami
  + Mycommand- brings an error bc not defined yet
  + Pwd (print working directory)
    - Root is the top-level directory of the computer then / in between to see how far from the root you are and you can go back as far as the root
  + Ls- list all files here (we are at the root so will list all)
  + Ls-F (- is an option within the command)- gives info which are directory and which are files
  + Ls –help (not on Mac)
    - On mac: man ls (gives manual of all options within ls)
    - Exit with q
* ls -F D Desktop/
  + see all that’s on the desktop
  + move around in directory with cd- change directory
* ls -F Desktop/data-shell/
  + cd desktop
    - bring back to desktop directory
  + cd data shell
    - if do cd data shell again it looks for another directory named data shell within data shell so it won’t work
  + can type the absolute path manually from users/dovabrenman/desktop/data-shell/
  + access home directory quickly from ~/Desktop
  + cd ..
    - refers to the last directory above
  + ls –a
    - shows all the files
    - anything that starts with a dot is hidden
  + create directory: mkdir
    - here we called it thesis
    - don’t put spaces in file names bc harder to use
  + tab to autocomplete a name we have already typed or already exists
  + put date at the beginning of the name to auto sort your files
  + also don’t use slashes in names bc those are used for moving between directories
  + can edit a text in unix 🡪 use nano command then name a new text editor
    - save file with crtl O
    - exit with ctrl-X
  + rm command is remove- rarely use this
    - rm –r will delete something that is not a directory like a file
    - rm –r –i will remove something that is a directory – will go through all the levels and will ask if you want a file at each level – goes in a loop to the different levels
  + mv command can be used to rename
    - mv oldfile newfile
  + cp copy a file
    - works like mv but doesn’t move it
    - cp oldfile
  + search files within a directory
    - ls \*.pdf (search for all files that are pdf)
  + wc – word count
    - lines, words, characters
    - wc \*.pdb
      * tells info for all files individually
    - wc p\* anything that starts with a p
    - wc p\*.pdb
    - ? match one character
      * wc \*t??ne.pdb
        + returns both methane and ethane
    - wc -1 \*.pdb
    - can sort- then look at shortest file (head) then longest (tail)
  + cat filename
    - can be used to view file or can merge two files of the same type together
      * cat file1
        + view contents of file 1
      * cat file1 file2 > combinedfile
        + move the contents of file1 and 2 into a combined file
      * cat combinedfile
        + view the new combined file

R and R studio

Can use shell script to call both R and python together

Can use the history command in unix

How to make a project folder- SWC

Projects help keep your data mobile so if you move it to another folder you can still find it

Don’t use abolsute paths in r scripts so that you can move it around

New project button on the top right corner- choose the folder that was created on the desktop

Want to save all excel as csv so other programs can read it

* make sure to avoid spaces and special characters \*?! Or ()

read in data: read.csv ()

* get help by highlighting then press fn+f1
* call file:
  + read.csv(“
* getwd
  + get working directory
  + this will give whole path- don’t use this to call the file bc then you cant send the data around bc its connected to your computer only
  + this way you only use data/cat.csv so only access the data folder which shouldn’t be moved and you could move the whole folder and it would still work
  + csv2- is for French- bc they use “,” for “.” and it would mess up comma separated so it separates by “;” instead of “,”
* double check data is loaded properly by clicking on the environment window that says your data and it will show up in a new window
* can look and individual columns
  + cats$like\_string

R graph catalogue- find the graph you are interested in and click, the code will pop up for ggplot

Oct 19, 2017

* not doing GIT this morning
* today: making custom functions, diplyr, and tidyr, GIT- version control, make own website on Github with R markdown
* r doesn’t have standard error so we can make own
  + newfunctionname<- function(argument 1, argument 2) {

}