

Welcome to the CRUK CI Linux & HPC courses Etherpad!

This pad text is synchronized as you type, so that everyone viewing this page sees the same text. This allows you to collaborate seamlessly on documents!

Get involved with Etherpad at <http://etherpad.org>

This pad is at: https://public.etherpad-mozilla.org/p/CRUK_Linux_11_19

<https://tinyurl.com/cruk-linux19>

Course web-site is at: <https://bioinformatics-core-shared-training.github.io/2018-11-13-shell/>
<https://tinyurl.com/Cllinux101>

We will be teaching material from the Dta Carpentry course::

Introduction to the Command Line for Genomics (but the lessons will apply to most command-line usage).

Please enter any questions below and we will endeavour to address them between lectures. Many thanks.

Shell

====

- Anatomy of a linux command i.e.

command -flags <parameters>. e.g. **ls -al training/**

- **man 'command'** to find out what these can be for a given command
- **cd** - change directory
- **pwd** - print working directory
- **cd ..** is change directory one level 'up'
- **cd ~** will take you to your home/login directory
- full paths **/a/b/c/d** is an explicit path to a specific directory whereas if we are at **a/b** then **c/d** is a relative directory from there.
- The wildcard ***** can be substituted for part of a filename to filter what a command uses e.g. **ls fred*** lists all files starting with fred
- Forgotten which commands you have used - **history** command is your friend. Up and down arrow keys will cycle through elements of your history.
- **cat** concatenate file command is to output contents of file(s) and later we will show you how to put the output of this to a new file
- **less** is useful to interactively view a file a page at a time or search with **/**
- **head** and **tail** list top 10 lines or bottom 10 lines of a file
- **cp** command copies files from one place to another
- **mkdir** makes new directories
- **mv** is a command to move files from a source to a destination (or to rename a file)
- We can use **ls -l** to find out the permissions on a file (who can do what to it) we use the **chmod** command to set the permissions
- We use **rm** to remove files (delete them IF we are permitted to)
- **grep** - your new best friend when searching through the contents of files
- redirection e.g. **>** lets us send the output of one command into a file **>>** appends to an existing file
- **wc** counts words, lines and characters in a file use a flag to select only one e.g. **wc -l file**
- **cut** lets us pull columns out of a tabular file

- **sort** orders whatever it gets as an input **uniq** filters to give you distinct items in the input e.g. how many different types of fruit not how many of each type
- **nano** text editor for mortals to create text files
- Others exist e.g. sublime, atom
- We can use a text editor to create our scripts. These enable us to automate tasks that run more than once e.g. thousands of times
- If you call your scripts **.sh** then it helps you remember that it is a shell script. We run our scripts by typing e.g. **bash script.sh**
- shell variables when output have a \$ in it e.g. **\$name**
- Use meaningful names for your files and directories (so in 6 months time you know what's in them)
- Structure your directories e.g. rawdata, data, outputs, scripts etc this will help to protect you from outputting things into the top level and making a mess
- This also helps with reproducibility of your analysis

-
- ***If you are CRUK-CI based and plan to use the cluster to do computationally intensive tasks in Linux then this course (This Friday) will be relevant:***

<https://www.training.cam.ac.uk/event/2659289>

- Site for you to practice/revise your commandline skills http://rik.smith-unna.com/command_line_bootcamp/?id=kp9lx7qdss
- A linux cheatsheet <https://github.com/bioinformatics-core-shared-training/2018-11-13-shell/blob/gh-pages/unixref.pdf> You can find

more examples with a Google search

- If you use Windows 10 and want to work through this material it is possible to install a bash linux environment under Windows

<https://www.windowcentral.com/how-install-bash-shell-command-line-windows-10>

- Please fill in the survey after/at the end of the course

Data Carpentry Pre-Workshop Survey: https://www.surveymonkey.com/r/dcpworkshopassessment?workshop_id=2018-11-13-shell

Data Carpentry Post-Workshop Survey: https://www.surveymonkey.com/r/dcpworkshopassessment?workshop_id=2018-11-13-shell

Your questions...

HPC
===