

Introduction to CLI

- Adapted from Software Carpentries: <https://datacarpentry.github.io/shell-genomics/01-introduction.html>

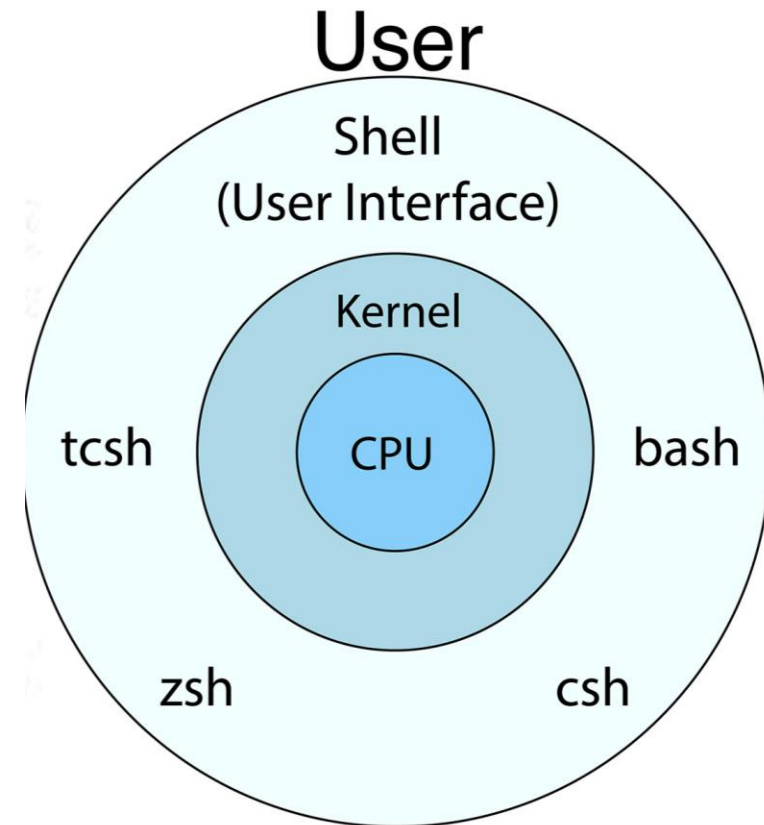
Ammar Aziz – VIDRL

Objectives

- Describe key reasons for learning shell.
- Navigate your file system using the command line.
- Access and read help files for bash programs and use help files to identify useful command options.
- Demonstrate the use of tab completion, and explain its advantages.

What is a shell?

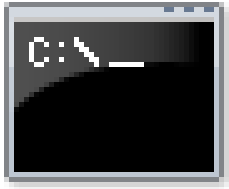
- A *shell* is a computer program that presents a command line interface which allows you to control your computer using commands entered with a keyboard instead of controlling graphical user interfaces (GUIs) with a mouse/keyboard/touchscreen combination.



Why should (you) care?

- Many bioinformatics tools can only be used through a command line interface.
- Automate your work.
- Less error-prone and more reproducible
 - It also gives you a way to communicate unambiguously what you've done, so that others can inspect or apply your process to new data.
- Many bioinformatic tasks require large amounts of computing power.
 - These tasks require large compute clusters

How to access the shell - Windows



- Command Prompt – deprecated?
 - Not useful for bioinformatics



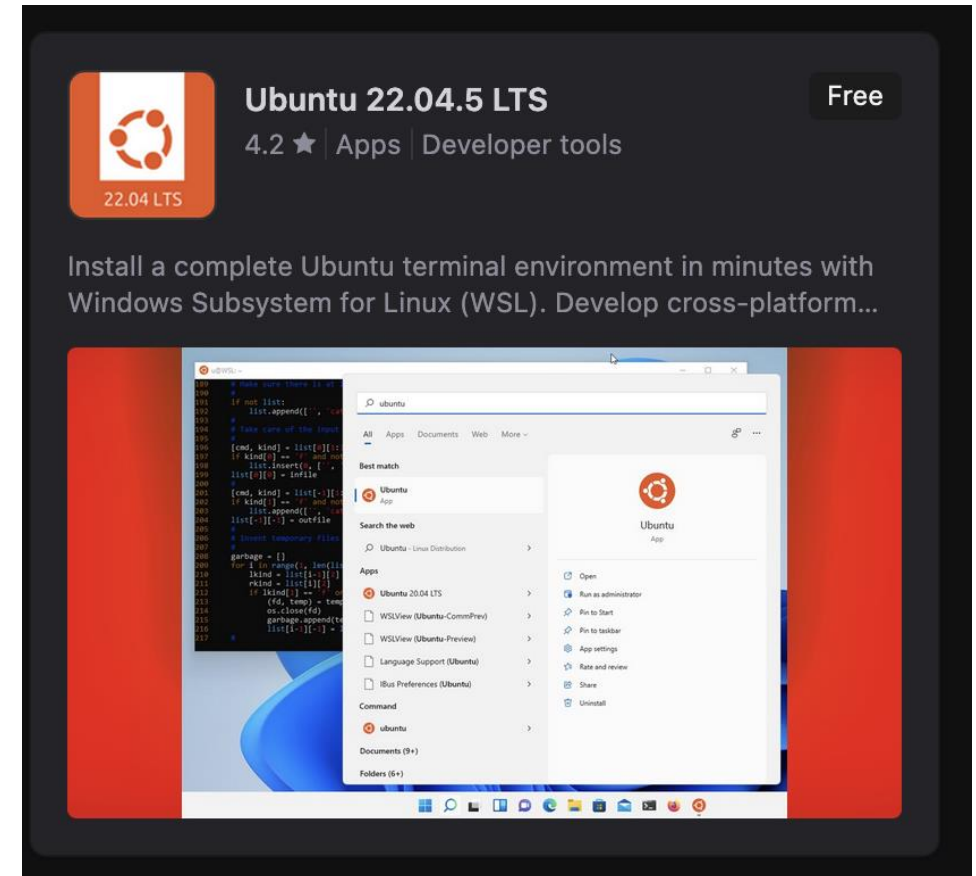
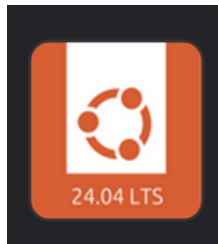
- **PowerShell – use this one**
 - Can run Conda/CLI tools
 - More modern, similar commands and behavior to Linux
 - Not posix compatible but good fine for conda



- Windows Terminal - Ignore
 - Runs all the above shells and more in one program
 - Ignore

How to access the shell – Linux in My Windows?

- Windows introduced a Linux compatibility layer which allows running Linux inside Windows.
- Linux terminal in Windows! Perfect for bioinformatics.
- Install from Windows Store



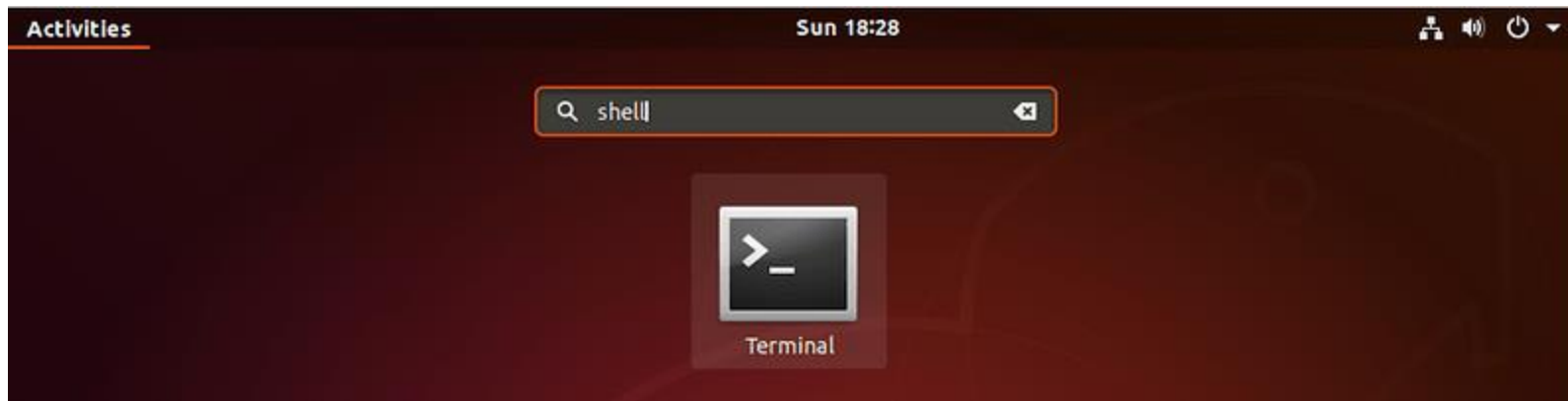
How to access the shell - MacOS



- Simply called **Terminal**
- Can run nearly all bioinformatics CLI tools
- Flavor: zsh
- Note on architecture:
 - The 'silicon' models of Macbooks run ARM architecture, different than 'standard' Intel/AMD x86.
 - Some bioinformatics tools will not run on ARM due to missing software
 - Rapidly improving due to cross compilation to ARM

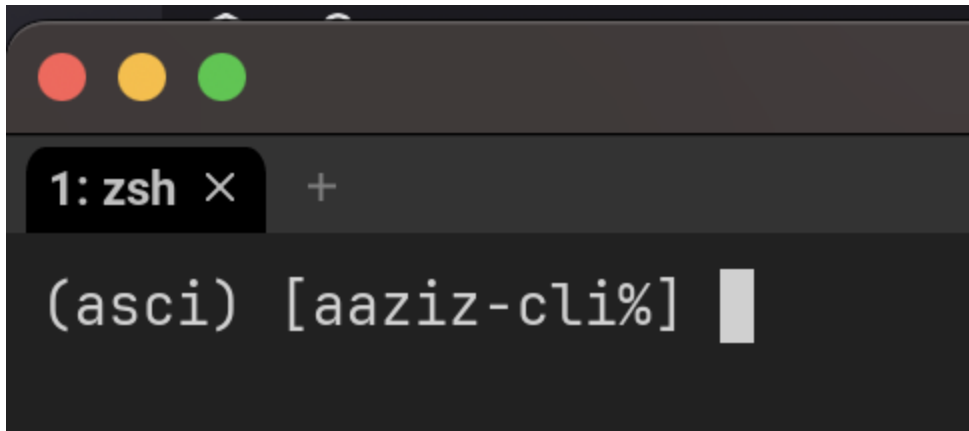
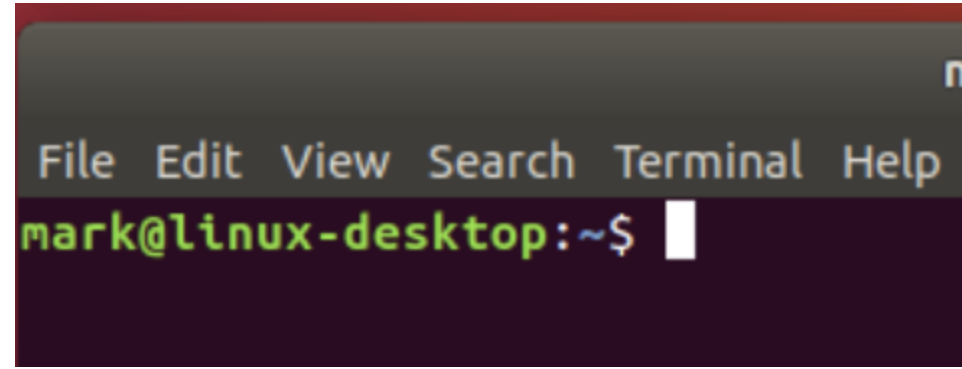
How to access the shell – Linux (Ubuntu)

- Called **Terminal**



- Flavor: Bash
- We will focus on Linux (Ubuntu) terminal - concepts apply to all

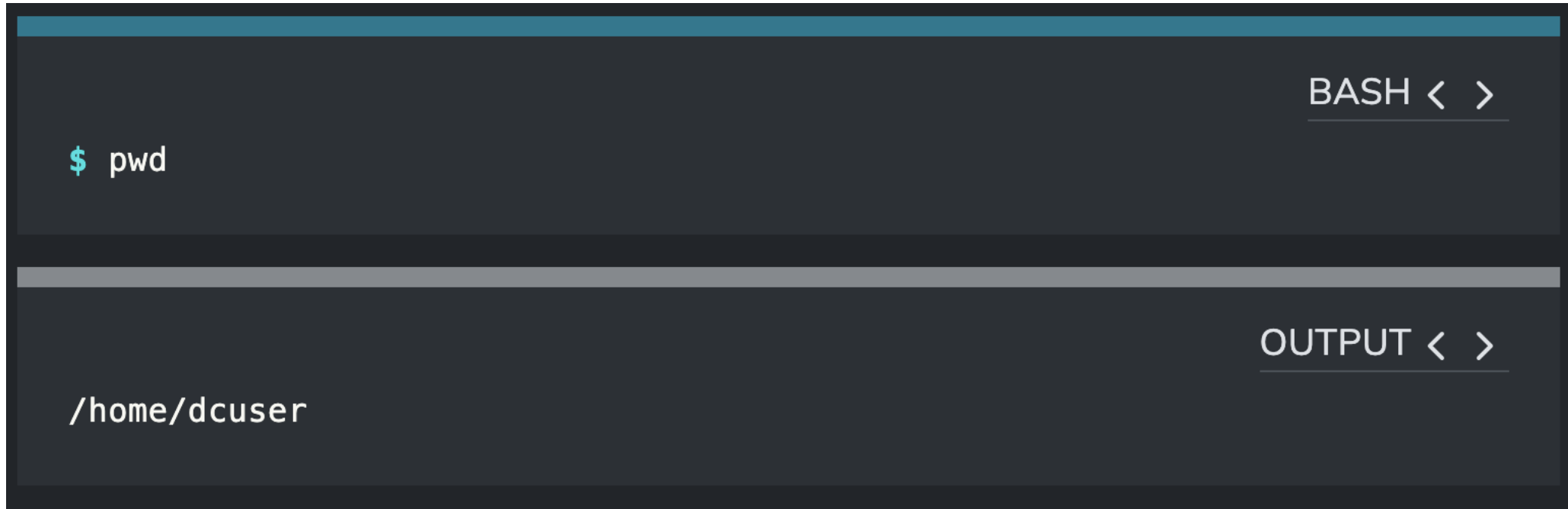
Terminal Basics - Prompt

A screenshot of a terminal window on a macOS system. The window has a title bar with red, yellow, and green window control buttons. Below the title bar, there is a tab labeled '1: zsh' with a close button (X) and a plus sign (+). The terminal content shows the prompt '(asci) [aaziz-cli%]' followed by a white cursor bar.A screenshot of a terminal window on a Linux desktop. The window has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal content shows the prompt 'mark@linux-desktop:~\$' followed by a white cursor bar.

- The dollar sign is a **prompt**, which shows us that the shell is waiting for input; your shell may use a different character as a prompt and may add information before the prompt. When typing commands, either from these lessons or from other sources, do not type the prompt, only the commands that follow it.
- Prompts can be changed!

Navigating - pwd

- Enter 'pwd'
- You should see a directory path printed out (probably something like /home/YOUR_USERNAME)

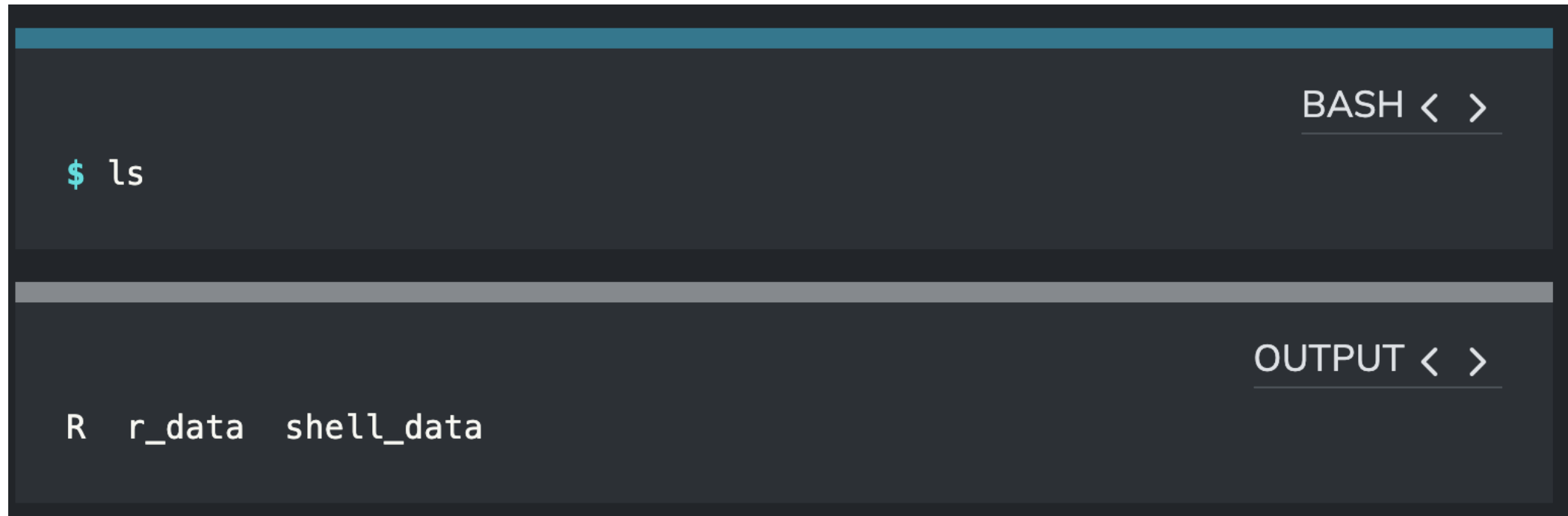


```
BASH < >
$ pwd
/home/dcuser
OUTPUT < >
```

A terminal window with a dark background and a blue header bar. The terminal shows a prompt '\$' followed by the command 'pwd'. Below the command, the output '/home/dcuser' is displayed. The terminal has a scroll bar on the right side.

Navigation - ls

- The command 'ls' short for listing shows the contents of the current working directory

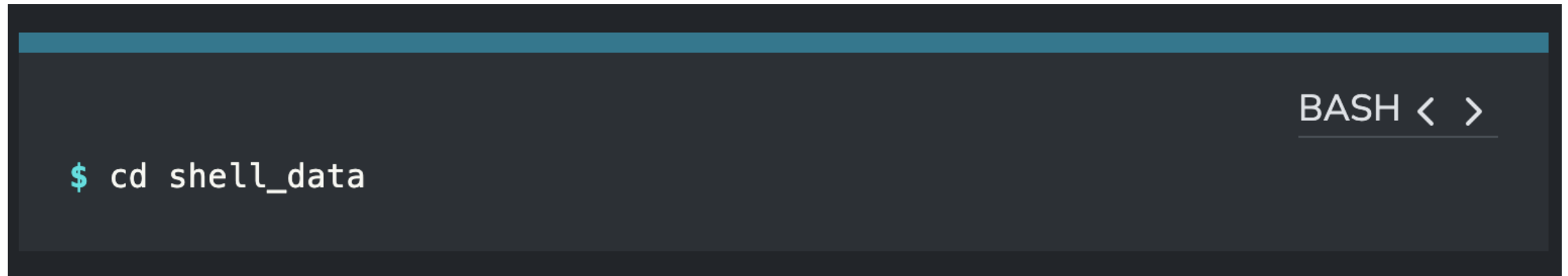


A terminal window with a dark background and a blue header bar. The header bar contains the text "BASH < >". The terminal shows the command "\$ ls" entered. Below the command, the output "R r_data shell_data" is displayed. The terminal has a light gray horizontal bar separating the command input area from the output area. The output area has a header bar containing the text "OUTPUT < >".

```
BASH < >
$ ls
R r_data shell_data
OUTPUT < >
```

Navigation - cd

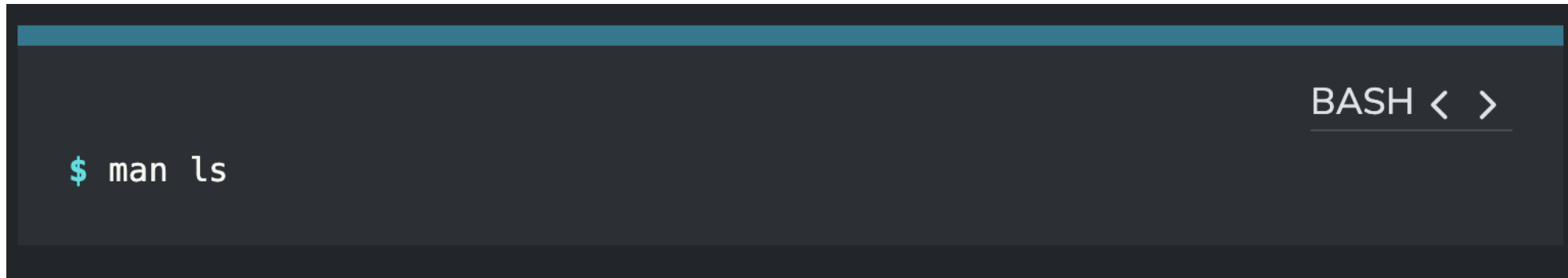
- The command 'cd' (changed directory) allows us to move into a folder – changing the working directory

A terminal window with a dark background and a blue header bar. The text 'BASH < >' is in the top right corner. The command '\$ cd shell_data' is entered in the terminal.

```
$ cd shell_data
```

Commands - help

- Use the command 'man' (short for manual) to show the help pages

A terminal window with a dark background and a blue title bar. The prompt is '\$' and the command 'man ls' is entered. The window title is 'BASH < >'.

```
$ man ls
```

Demonstration – listing files and entering directories



Tab completion

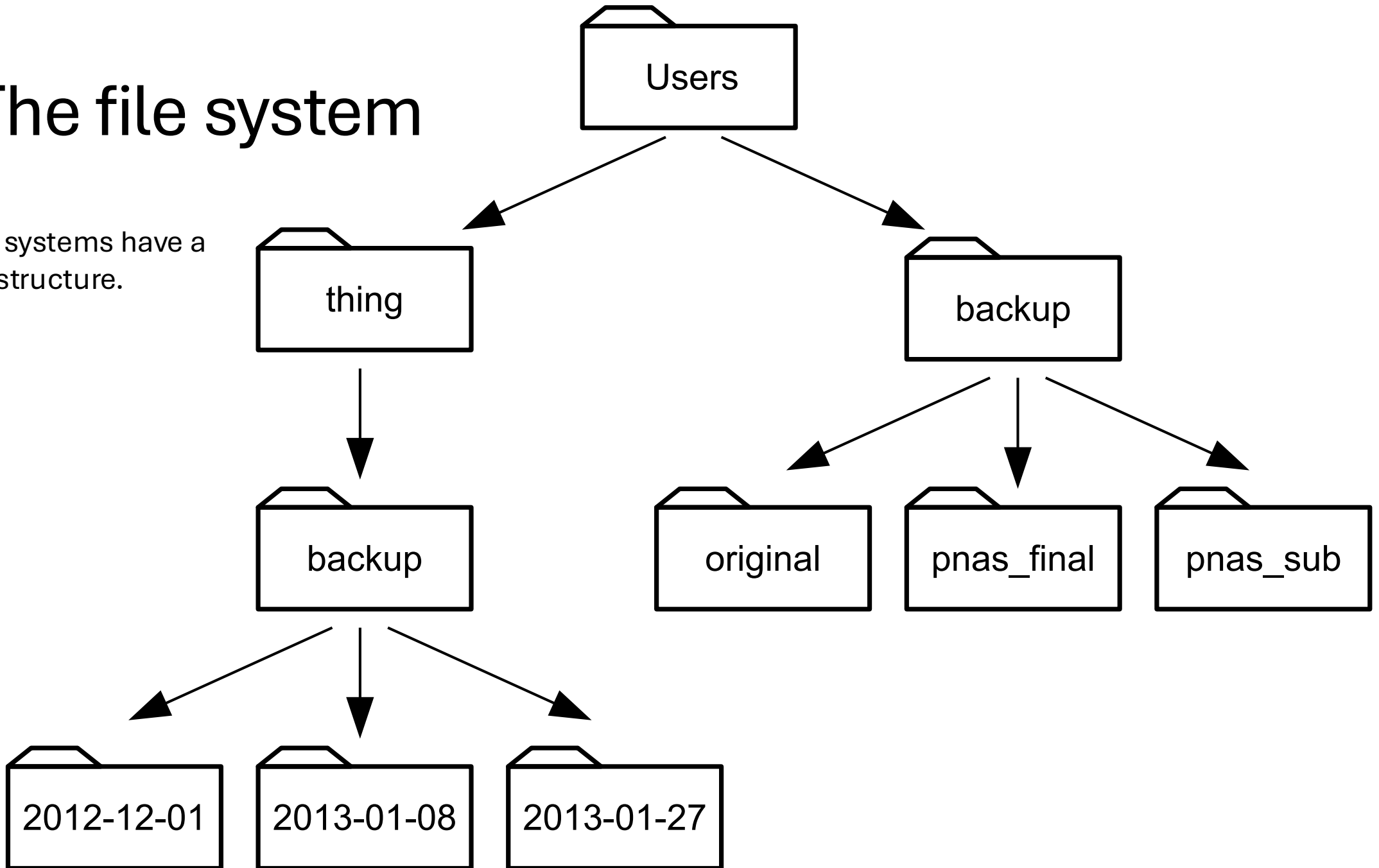


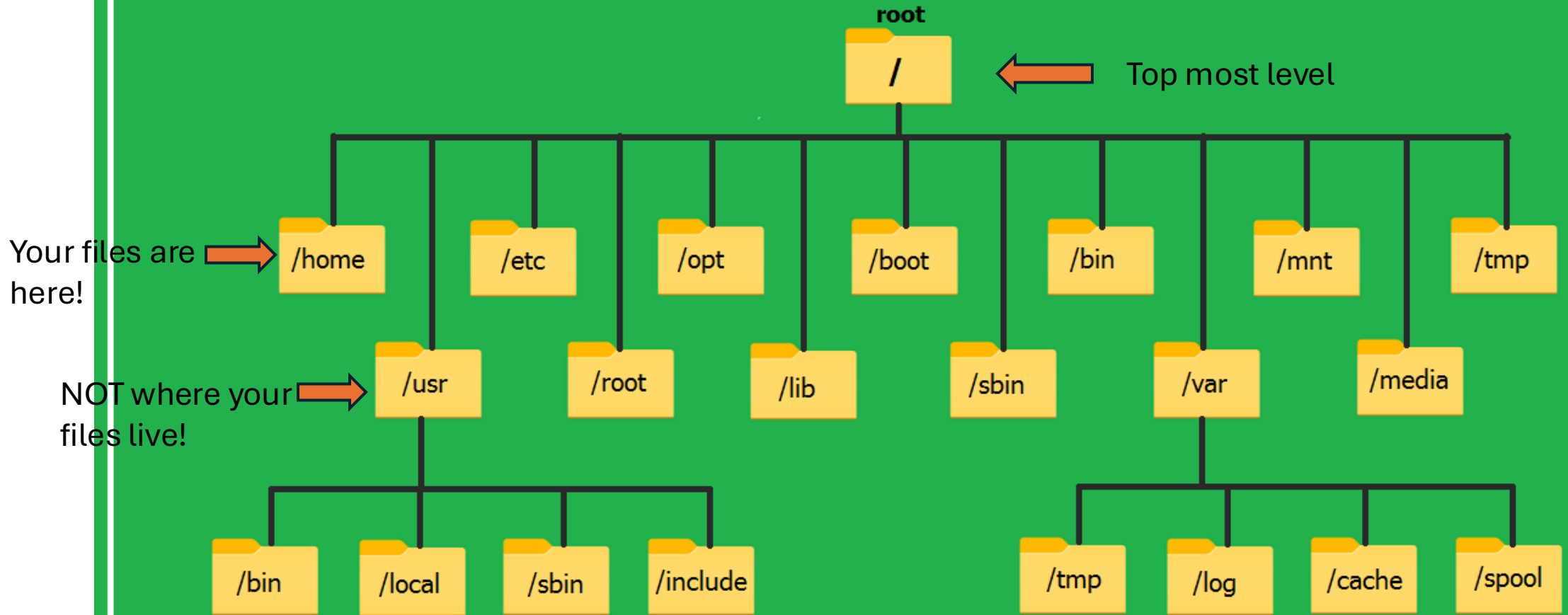
File system (FS)

- File system controls the location and access to files on your system:
 - Location of the files
 - Size
 - Metadata
 - Permissions
 - And much more
- Understanding the FS system is important for navigating and manipulating files

The file system

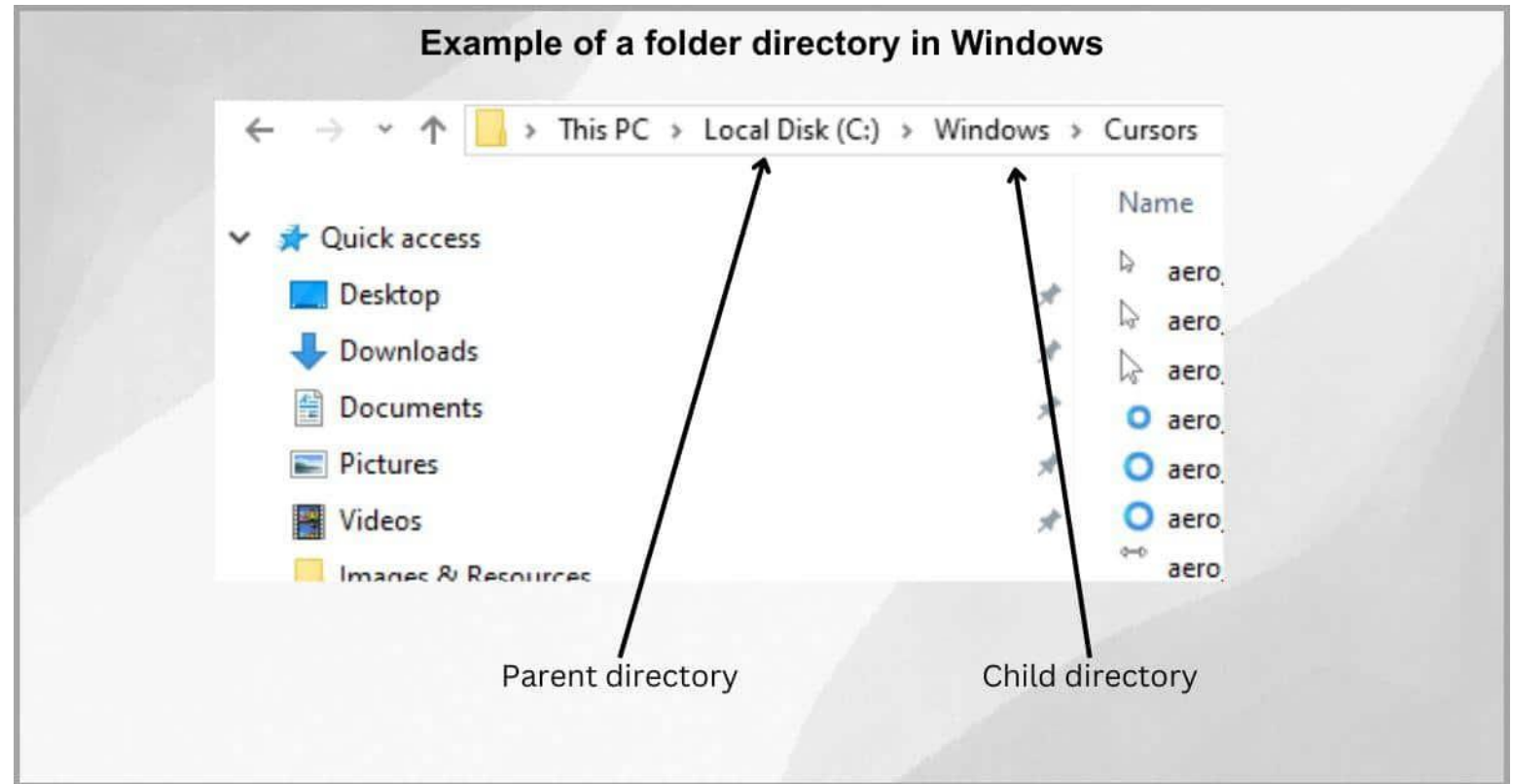
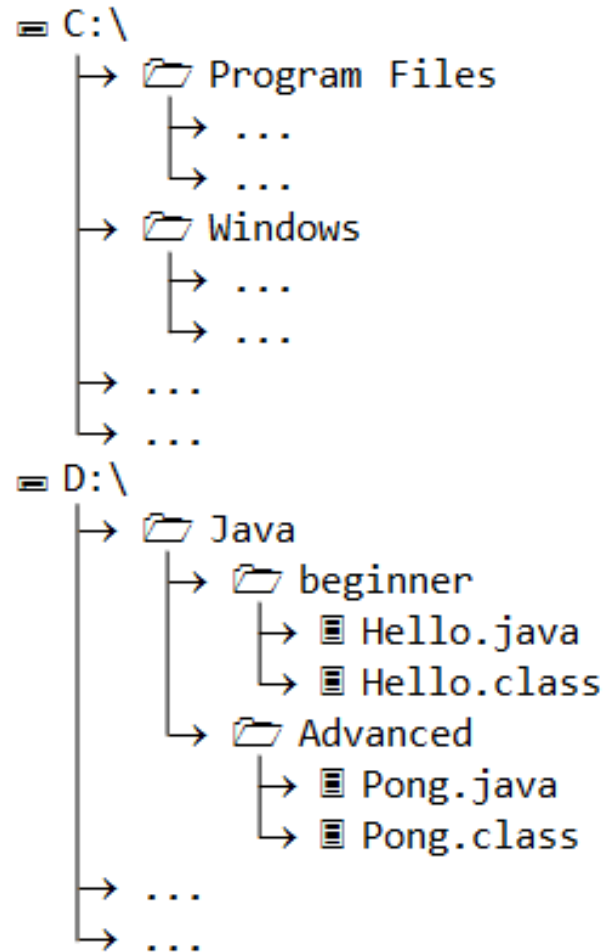
Most file systems have a tree like structure.





File System Hierarchy(FHS) of Linux

Windows FS



Finally

- The shell gives you the ability to work more efficiently by using keyboard commands rather than a GUI.
- Useful commands for navigating your file system include: `ls`, `pwd`, and `cd`.
- Most commands take options (flags) which begin with a `-`.
- Tab completion can reduce errors from mistyping and make work more efficient in the shell.

Questions? + Resources for future learning

- Data Carpentry is excellent:
 - <https://datacarpentry.github.io/shell-genomics/01-introduction.html>
- <https://sandbox.bio> (based on Carpentries):
 - <https://sandbox.bio/tutorials/carpentries-shell-novice/>
 - In browser interactive learning – check it out!
 - Much more bioinformatics training at <https://sandbox.bio/>
- Test your skills at <https://cmdchallenge.com>
- Hardcore: Aaron Quinns University course:
 - <https://github.com/quinlan-lab/applied-computational-genomics>
 - Slides, videos, homework