Session 2: Working with files and directories (Part II) (Andrew Weisman)

Notes:

- The notes below are based on the "bottom" part of this webpage
- The page you are reading can be found here: https://github.com/CBIIT/p2p-datasci/blob/master/workshop_materials/2021-09-21-introduction_to_linux/instructors_notes/working_with_files_and_directories_part_2-andrew.md
  - This includes "Homework" exercises
- Copy that link into the chat, including the link for the course website: https://cbiit.github.io/p2p-datasci/2021-09-09-introduction_to_linux

Create a text file ctd.

Creating Files a Different Way

- `cd` to `~/Desktop/shell-lesson-data/thesis`
- We just saw how to create files using `nano`
- We can also create files using `touch`, e.g., `touch my_file.txt`
- If we inspect the file using `ls -l`, we see that it is 0 bytes
  - E.g., in a text editor, you get a blank file
- Creating a file this way allows you to create a placeholder file that you can edit further using the appropriate program for that extension
- Note you can create a file with any extension
  - Windows by default hides the extension

Moving files and directories

- Return to `shell-lesson-data` by either `cd`ing up a directory or `cd`ing to `~/Desktop/shell-lesson-data`
- Change the name of `thesis/draft.txt` to `thesis/quotes.txt` using `mv`
  - This has the effect of renaming the file
  - Generally, it has the effect of *moving* a file
- Confirm using `ls`: `ls thesis/quotes.txt`
  - Using `ls` with an argument only lists the argument if it exists
- In general, `mv` silently overwrites the target if it exists
- `mv` also works on directories
- Move `quotes.txt` to the current working directory: `mv thesis/quotes.txt .`
- Confirm using `ls` that the `thesis` directory is empty and that the current directory contains `quotes.txt`
- Confirm that the previous command `ls thesis/quotes.txt` no longer works

Exercise: Moving Files to a new folder

Copying files and directories

- `cp` works like `mv` except it copies a file instead of moving it
- Use it to copy `quotes.txt` to `thesis/quotations.txt`
- Check using `ls` with both arguments (same exact arguments as `cp` command) to see that the file exists in both places
- Copy a directory and all its contents using the recursive option `-r`, e.g., to back up a directory: `cp -r thesis thesis_backup`
    - Without `-r`, since `thesis` is a directory, the copy will not work because `cp` otherwise essentially expects the first argument to be a file
- `ls` the `thesis` and `thesis_backup` directories together to see that `quotations` was copied successfully

Exercise: Renaming Files

Exercise: Moving and Copying

Removing files and directories

- Return to `shell-lesson-data`
- Tidy things up by removing (using `rm`) the `quotes.txt` file
- Confirm using `ls`
- Deleting is forever: there is no trash or recycle bin
- Try removing the `thesis` directory using `rm` (not `rmdir` or `rm -r`)
- It fails because `rm` by default only works on files, not directories
- To do it successfully, use `rm -r`
- To do it successfully without prompting for each file, use `rm -rf`

Operations with multiple files and directories

- You can move, copy, and remove multiple files and directories at once
- Do this by provide lists of files/directories
- Enter the `data` subdirectory of `shell-lesson-data`
- Create a `backup` directory and copy multiple files to it: `cp amino-acids.txt animals.txt backup/` and show what happened using `ls`
    - This shows that you can have multiple sources but only a single destination at the end
- When there is more than two arguments to `cp`, the command expects a directory as the final argument
    - So, when you do e.g. `cp amino-acids.txt animals.txt morse.txt`, an error is thrown
    - This rarely comes up in practice, because you would rarely enter a command like this (think about it... why would you do it?)

Using wildcards for accessing multiple files at once

- You can also perform operations by specifying patterns using wildcards
- `*` matches zero or more characters
- It's often used with `mv`, `cp`, and `rm`, but first tested with `ls`
- `cd` into `shell-lesson-data/molecules` and do an `ls`
- `ls *.pdb` matches all `.pdb` files
- `ls p*.pdb` matches only those that begin with 'p'
- `?` matches exactly one character

    - Not used as frequently as `*`
- `ls ?ethane.pdb` maches `methane.pdb` whereas `ls *ethane.pdb` matches both `ethane.pdb` and `methane.pdb`
- You can use multiple `?` characters to specify a precise number of characters
- E.g., `ls ???ane.pdb` matches `cubane.pdb`, `ethane.pdb`, and `octane.pdb`
- Try `ls *.pdf` to show that wildcards that result in files that don't exist return errors as usual

Homework: List filenames matching a pattern

Homework: More on Wildcards

Exercise: Organizing Directories and Files

Exercise: Reproduce a folder structure

Key Points