REU Boot Camp Pt 1 Command Line, Github, and Jupyter Notebooks

A brief introduction

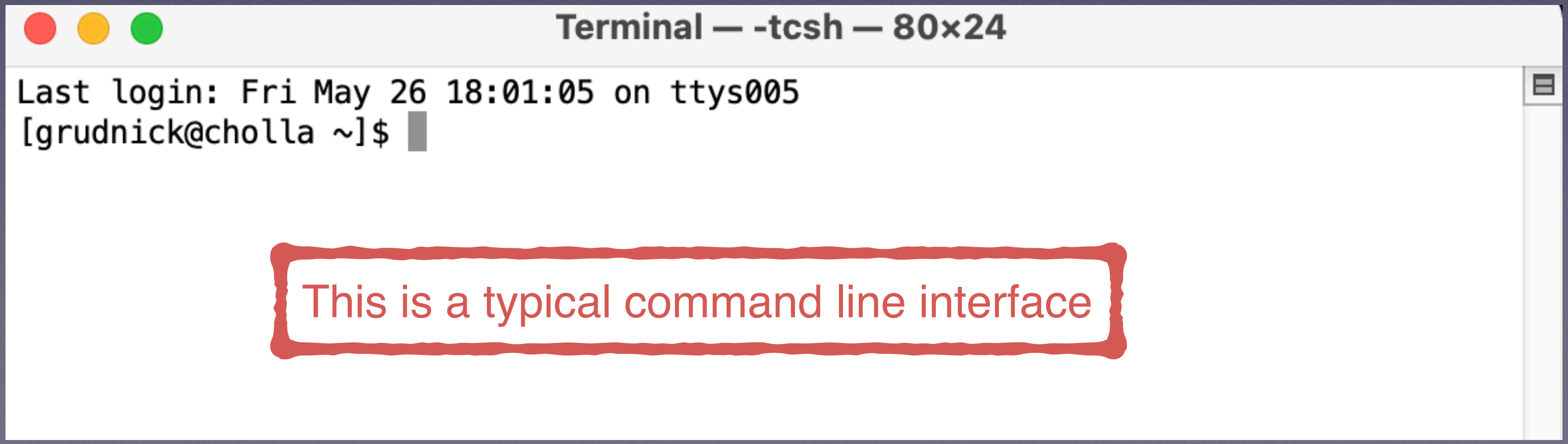
Prof. Gregory Rudnick
2056B Malott
grudnick@ku.edu

What we are going to do today

- Learn how to work from the command line
- Learn how to set up a python environment
- Learn how to use GitHub for version control, backups, and sharing work
- Learn about Jupyter Notebooks as a means of coding, documenting code, and sharing code and results.

What is the command line?

- Command line is an interface that accepts written commands.
- It is the standard interface for many research activities.
- It is a powerful and complex interface but the basics are straightforward.
- I am going to run you through some of the most fundamental commands. I will share these with you by the end of the class.
- Login to your terminal with your KU credentials - **Do not update Ubuntu if prompted.**
- Open "Terminal" app. This is the command line interface.



```
Last login: Fri May 26 18:01:05 on ttys005
[grudnick@cholla ~]$
```

The “Shell”

- Is a program running in the background on the computer
- It interprets everything written at the command line.
- There are different kinds, each with their own syntax peculiarities: bash, csh, tcsh, zsh, ksh. The most commonly used today is **bash**
- Most commands are the same among shells. The differences are more for advanced users.

Common commands
try these at your terminal.
All Case Sensitive

- Navigation commands
- Files can always be specified by their full location or "path", or by their name if in the current directory.
  - man <command> - brings up manual page about every command
  - pwd - shows current dir
  - ls - list contents of dir.
  - ls -la - verbose listing. **modifiers with -<modifier> change command options**
  - mkdir NewDir - makes new directory (i.e., folder) /home/username/NewDir/
  - rmdir NewDir - removes directory NewDir/
  - cd NewDir or cd /home/username/NewDir - change to new dir.
  - cd .. - go back ("up") one dir.
  - cd or cd ~ - return to home dir.
  - cd - - go to previous directory

Common commands
try these at your terminal.
All Case Sensitive

- Things you do to files
  - cp SomeFile NewFile - copies file to another file
  - cp SomeFile NewDir/ - copies file to new dir.
  - cp Dir/SomeFile . - copies file in Dir to current dir. (.)
  - cp ../SomeFile .   - copies file in one dir. up to current dir. (.)
  - mv SomeFile NewDir/ - moves file to new dir.
  - mv SomeFile ../../ - moves file up 2 dirs.
  - mv *.txt NewDir/ - moves all files ending in .txt to new dir.
  - rm SomeFile - deletes file

Common commands try these at your terminal. **All Case Sensitive**

- Ways to see ASCII file contents
  - wc -l SomeFile - how many lines in file
  - more SomeFile - scroll thru file, carriage return for 1 line at time, space bar for many lines at a time
  - less SomeFile - similar to more but arrow keys move you up and down
  - tail -9 SomeFile - shows last 9 lines of file
  - head -9 SomeFile - shows first 9 lines of file
  - cat SomeFile - lists full content of file to screen
  - grep string SomeFile - finds every occurance of "string" in File
  - file SomeFile - tells you the file type

Wildcards and dangers to use with caution

- using wildcards
  - ls *.cat - will list all files ending with ".cat"
  - rm * - will remove all files in a directory
- Dangers
  - **rm is not recoverable. If it's removed with "rm" it is gone.**

Setting up a Python Environment

- A python environment contains a fully independent installation of python, including packages that can be different for every environment.
- Open a Terminal Window
- type `/opt/conda/bin/conda init bash`
- close your terminal window and reopen it
- type `conda create -n python3env anaconda`
- type `conda activate python3env`
- python3env is called your "Conda environment" or "python environment"
- This should take about 6-8 minutes. Leave window alone while it's working.
- This will need to be done every time you use a new computer but will stick around once you install it on a given computer.

GitHub

- Git is a version control and backup protocol for storing files remotely, keep track of changes, and allowing multiple people to access files.
- GitHub is a web service that allows to create, modify, update and share “repositories”
- I have my students put their code on github.
- When they come to my office I can pull their most recent version of the code to my computer.
- Can also store non-ASCII files.
- GitHub is free.

GitHub intro and terminology

- Repository - A directory structure that contains files. It can exist locally or on the github.com servers
- You make a repository at the web page and include a README
- You can clone (or copy) that repository onto any computer
- Using the web interface you can add new files or update current ones using the add file button
- It is **your responsibility** to put your files from your local computer onto the server.
- There are ways to do this from the command line using the push and pull commands, but you need to store a large token which is a pain to type unless you have a "token manager"
- We will be doing it just from the web interface
- Github can also allow multiple people to work on the same code and "check out" copies. We won't worry about that here.

What are Jupyter Notebooks

- The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text.
- Runs in your browser
- Is stored on your disk in a *.ipynb file

GitHub and Jupyter notebook tutorial

We will retrieve a Jupyter Notebook from Github.

1. Make sure you are in your home directory with "cd ~"
2. make a directory called REUBootcamp using "mkdir REUBootcamp" then "cd REUBootcamp"
3. With a browser make a Github account at https://github.com
4. Go to the following URL https://github.com/grudnick/REU_bootcamp_KU_2023
5. Push the green "Code" button and copy the URL you see.
6. At your command line, type "git clone " followed by the text you copied, pasted by pressing the middle mouse button.
7. This creates a "clone" of the REU_bootcamp_KU_2023 GitHub repository on your computer that contains all the most recent versions of the files in that repository

GitHub and Jupyter notebook tutorial

Once your python environment.has finished installing, do the following.

8. cd into the new REU_bootcamp_KU_2023 directory and type
9. ls to see what is in the directory.
10. Start your conda environment with conda activate python3env
11. Then type jupyter lab undergrad_summer2023_notebook_tutorial.ipynb
12. This will open a window in the Browser.
13. Follow the instructions in the notebook.
14. **Ask questions!**