

Sept 26, 2022

The Unix Command-line

Suzy Strickler
srs57@cornell.edu

Introductions

- Syllabus
- Office hours: Monday 1 - 2 pm on Zoom
- Who knows Linux (work in pairs today)?
- Introductions - Tell us a bit about yourself and why you are here

Why use Unix?

- Free to use.
- Stable and secure
- Easily maintained
- Can edit and enhance source code (Open Source)
- World-wide community of developers
- Can be installed on just about anything
- Many bioinformatics programs will not run on Windows
- Many analysis servers run Linux

Linux Distributions

- Distributions have been created around the Linux kernel
- Many distros available (distrowatch.com)
- Examples:
 - RedHat
 - Debian (we will use this)
 - Ubuntu

Fall 2020

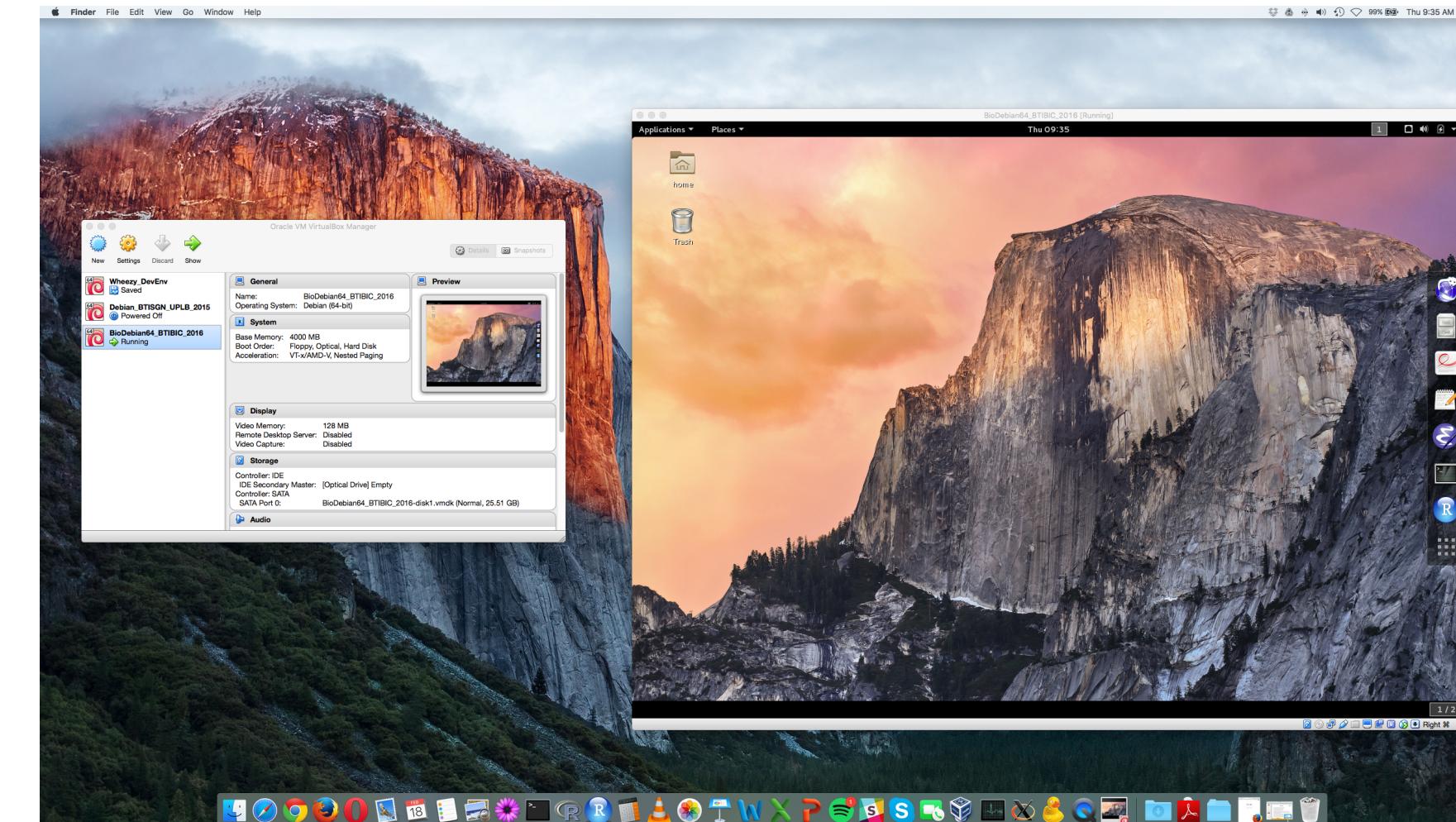
Page Hit Ranking			Page Hit Ranking		
Data span:			Data span:		
Last 6 months			Last 6 months		
Rank	Distribution	HPD*	Rank	Distribution	HPD*
1	MX Linux	3823▼	1	MX Linux	3161▼
2	Manjaro	2640▼	2	EndeavourOS	2627▲
3	Mint	2368▼	3	Manjaro	2229▼
4	Ubuntu	1670▼	4	Mint	1928▲
5	Pop!_OS	1398▲	5	Pop!_OS	1542-
6	Debian	1377▼	6	Ubuntu	1327▼
7	elementary	1341▼	7	Debian	1259-
8	Solus	1050▼	8	Garuda	1152▲
9	Fedora	1000▼	9	elementary	1129-
10	Zorin	921▼	10	Fedora	958▲
11	KDE neon	904▼	11	Zorin	859▲
12	deepin	893▼	12	openSUSE	819▼
13	openSUSE	782-	13	KDE neon	692▲
14	EndeavourOS	764▲	14	Solus	607▼
15	Ubuntu Kylin	723-	15	antiX	543▲
16	Arch	701▼	16	Arch	503▼
17	antiX	662▼	17	Slackware	463▲
18	CentOS	599▼	18	Lite	459▼
19	Linuxfx	563▲	19	Artix	438-
20	ArcoLinux	552▼	20	PCLinuxOS	429-
21	PCLinuxOS	531-	21	Kali	412-
22	Puppy	519▲	22	Puppy	404-
23	Kali	509-	23	deepin	398▼

Class Content

- Introduction to the Virtual Machine
- Terminals & file system navigation
- Anatomy of a UNIX command
- Wildcards, shortcuts and special characters
- File permissions
- Compression UNIX commands
- Networking UNIX commands

What is a Virtual Machine?

- A VM is an operating system which has been installed inside a simulated environment.
- Unlike emulators, virtual machines interface with real hardware.

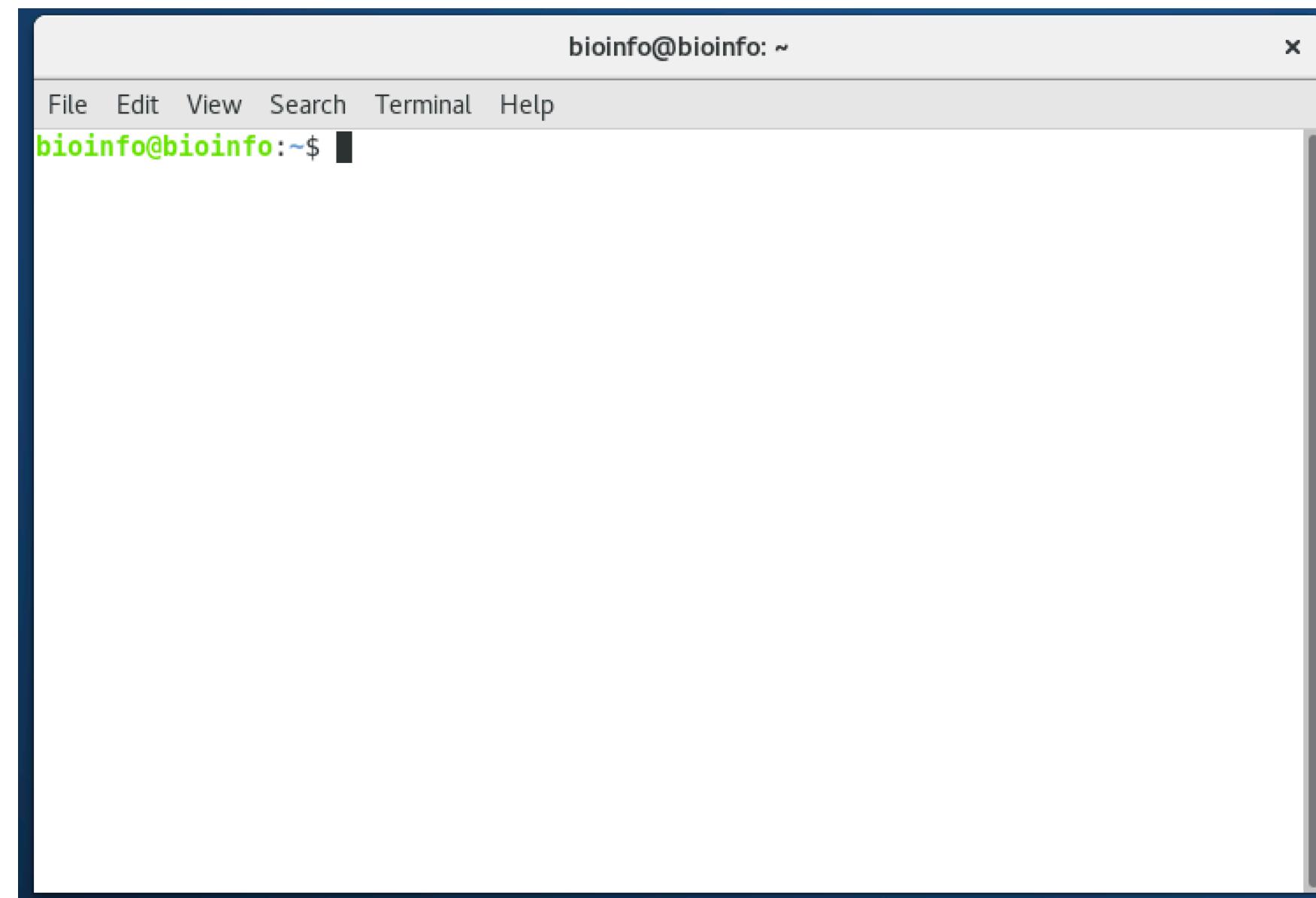


Exercise 1:

- Start your PLSCI7202 Virtual Machine and log in!
- (The password is the same as the username)

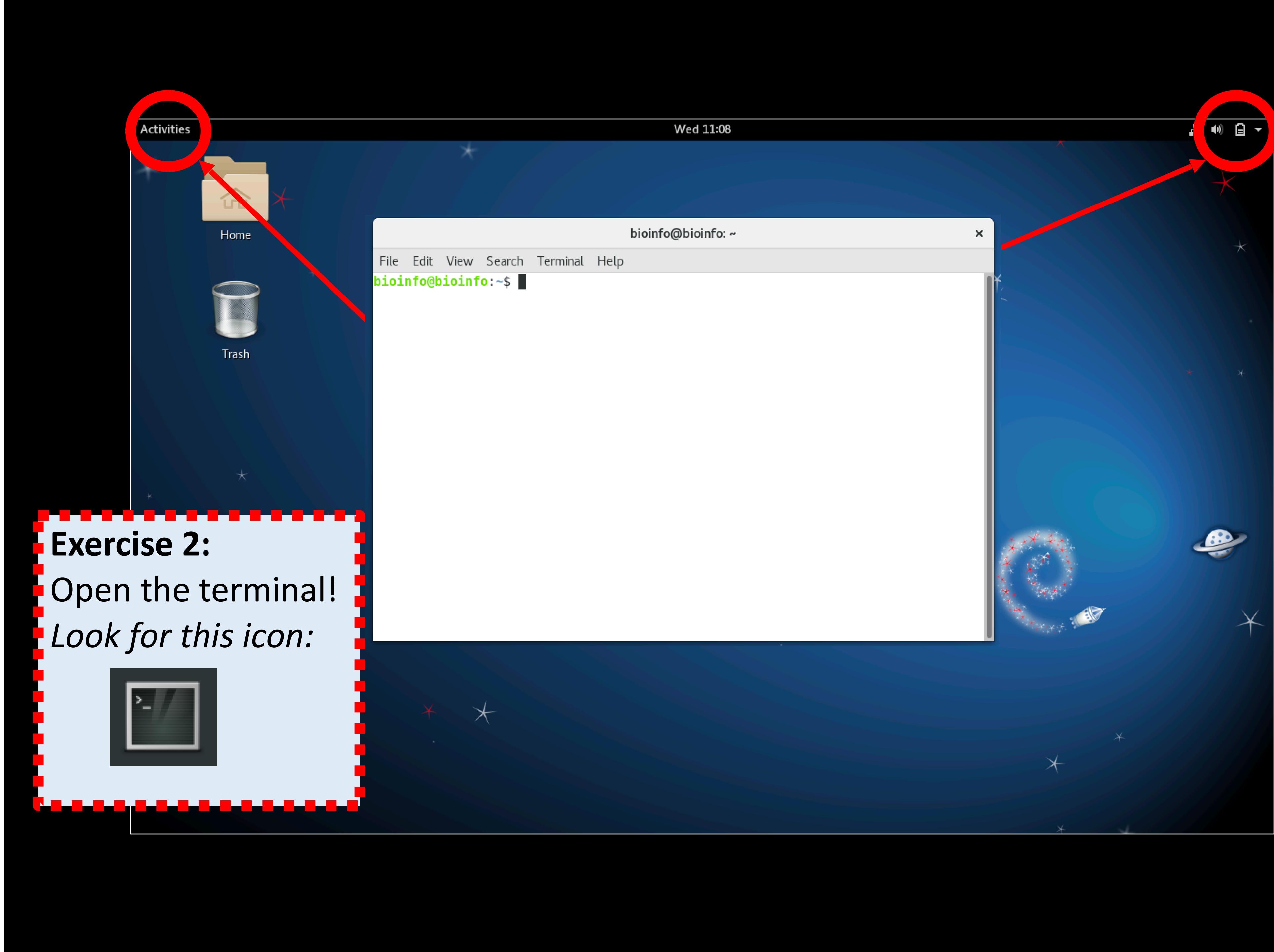
What is a Terminal?

- A terminal is a textual interface for interacting with a computer.
- Using the terminal, one can issue powerful and concise command-line instructions for the computer to follow.



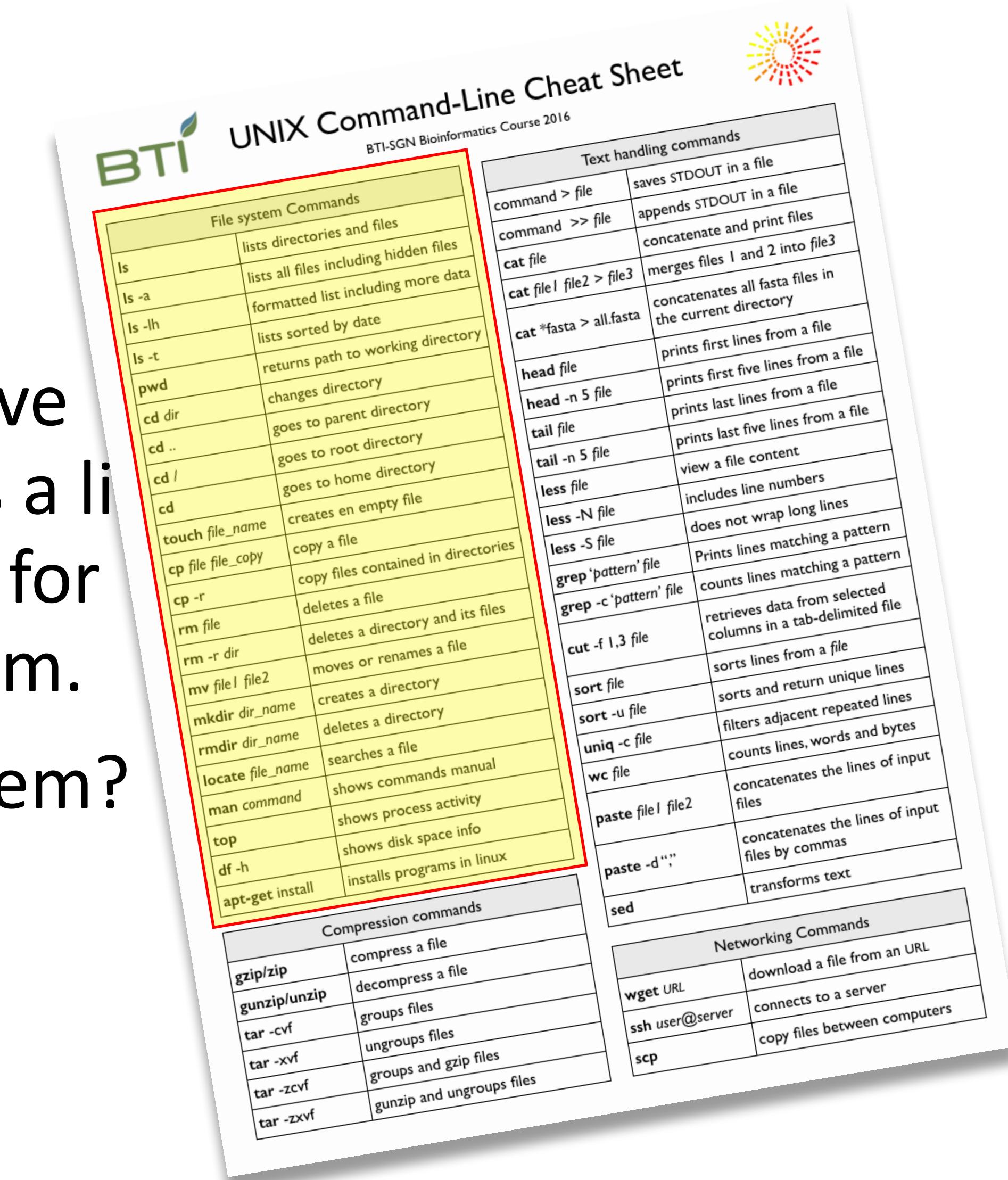
Why use the command-line (terminal)?

- Large data files can not be opened or loaded in most graphical software and web sites.
- Software for biological data analysis is often used through UNIX command-line operations.
- Many of the servers for biological data analysis use Linux/Unix as their operating system.
- Data analysis on calculation servers are much faster since we can use more CPUs and RAM than in a PC or laptop (e.g. BTI's "Boyce" server has 64 cores and 1TB RAM)

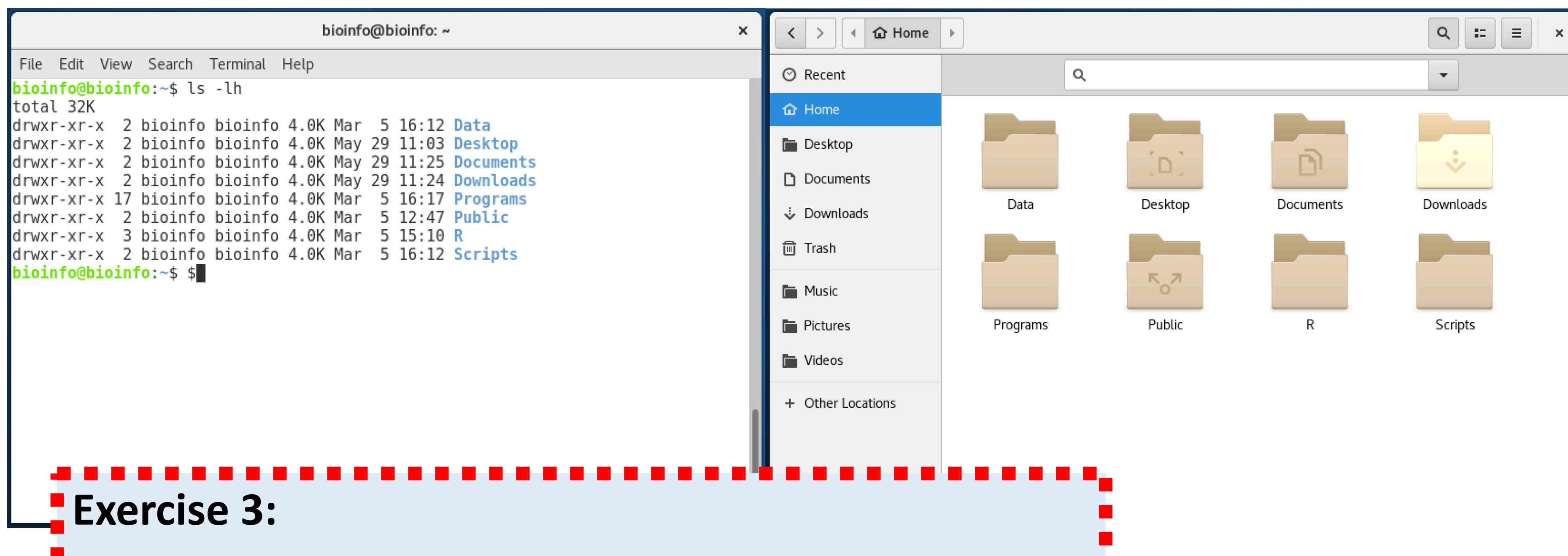
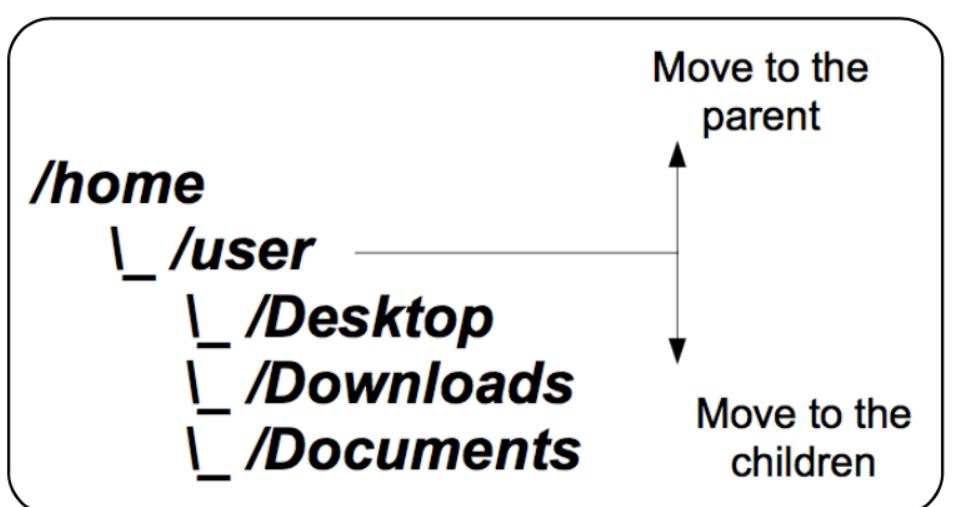


Command-line File System Navigation.

- The cheat sheet you have been provided contains a list of common commands for navigating the file system.
- But what *is* the file system?



The File System



Exercise 3:

Type “ls -lh” into the terminal and press enter.

Home and Root Directories

```
Noes-MacBook-Pro:~ Noe$ ls -lht
total 0
drwx-----+ 29 Noe  staff  986B May 31 11:24 Desktop
drwx-----@  8 Noe  staff  272B May 31 08:26 Dropbox
drwx-----+ 54 Noe  staff  1.8K May 30 16:01 Downloads
drwx-----+  8 Noe  staff  272B May 28 21:06 Pictures
drwxr-xr-x  18 Noe  staff  612B May 17 11:12 BTI
drwxr-xr-x   5 Noe  staff  170B May  8 11:44 programs
drwx-----+ 15 Noe  staff  510B Apr 10 08:33 Documents
drwxr-xr-x   6 Noe  staff  204B Mar 18 09:22 VirtualBox VMs
drwxr-xr-x   8 Noe  staff  272B Mar 14 19:26 py_devel
drwx-----@ 51 Noe  staff  1.7K Mar 11 15:08 Library
```

Home directory

```
/home/bioinfo
/home/noe
/home/noe/Desktop
```

```
noe@debian-virtualbox:~$ ls -l /
total 108
drwxr-xr-x  2 root root  4096 Sep 26  2012 bin
drwxr-xr-x  3 root root  4096 Nov  9  2012 boot
drwxr-xr-x 15 root root  3140 May 31 12:46 dev
drwxr-xr-x 130 root root 12288 May 31 12:45 etc
drwxr-xr-x  5 root root  4096 Feb 28 13:54 export
drwxr-xr-x  4 root root  4096 Nov  7  2012 home
lrwxrwxrwx  1 root root   30 Sep 26  2012 initrd.img
drwxr-xr-x 12 root root 12288 Nov  9  2012 lib
drwxr-xr-x  2 root root 12288 Nov  9  2012 lib32
lrwxrwxrwx  1 root root    4 Sep 26  2012 lib64 -> /|
drwx-----  2 root root 16384 Sep 26  2012 lost+found
drwxr-xr-x  3 root root  4096 Sep 26  2012 media
drwxr-xr-x  2 root root  4096 May  1  2012 mnt
drwxr-xr-x  2 root root  4096 Sep 26  2012 opt
dr-xr-xr-x 134 root root    0 May 31 12:45 proc
drwx----- 10 root root  4096 Nov 15  2012 root
drwxr-xr-x  2 root root  4096 Nov  9  2012 sbin
drwxr-xr-x  2 root root  4096 Jul 21  2010 selinux
drwxr-xr-x  2 root root  4096 Sep 26  2012 srv
drwxr-xr-x  13 root root    0 May 31 12:45 sys
drwxrwxrwt 11 root root  4096 May 31 19:56 tmp
drwxr-xr-x  11 root root  4096 Sep 26  2012 usr
drwxr-xr-x  14 root root  4096 Sep 26  2012 var
```

Root directory

/bin, /lib, /usr	code and code libraries
/var	logs and other data
/home	user directories
/tmp	temporary files
/etc	configuration information
/proc	special file system in Linux

Anatomy of a UNIX Command

- Every UNIX command is made up of a series of space-separated strings.
- The first of these strings is always the command you will run.
- The user can also provide options (shown in red) and arguments (shown in blue)
- A command can require arguments, as can an option.
- In this example, the “-A” option is being given the argument “3” and the command is being given the argument “file.txt”

```
grep -c -A 3 --ignore-case file.txt
```

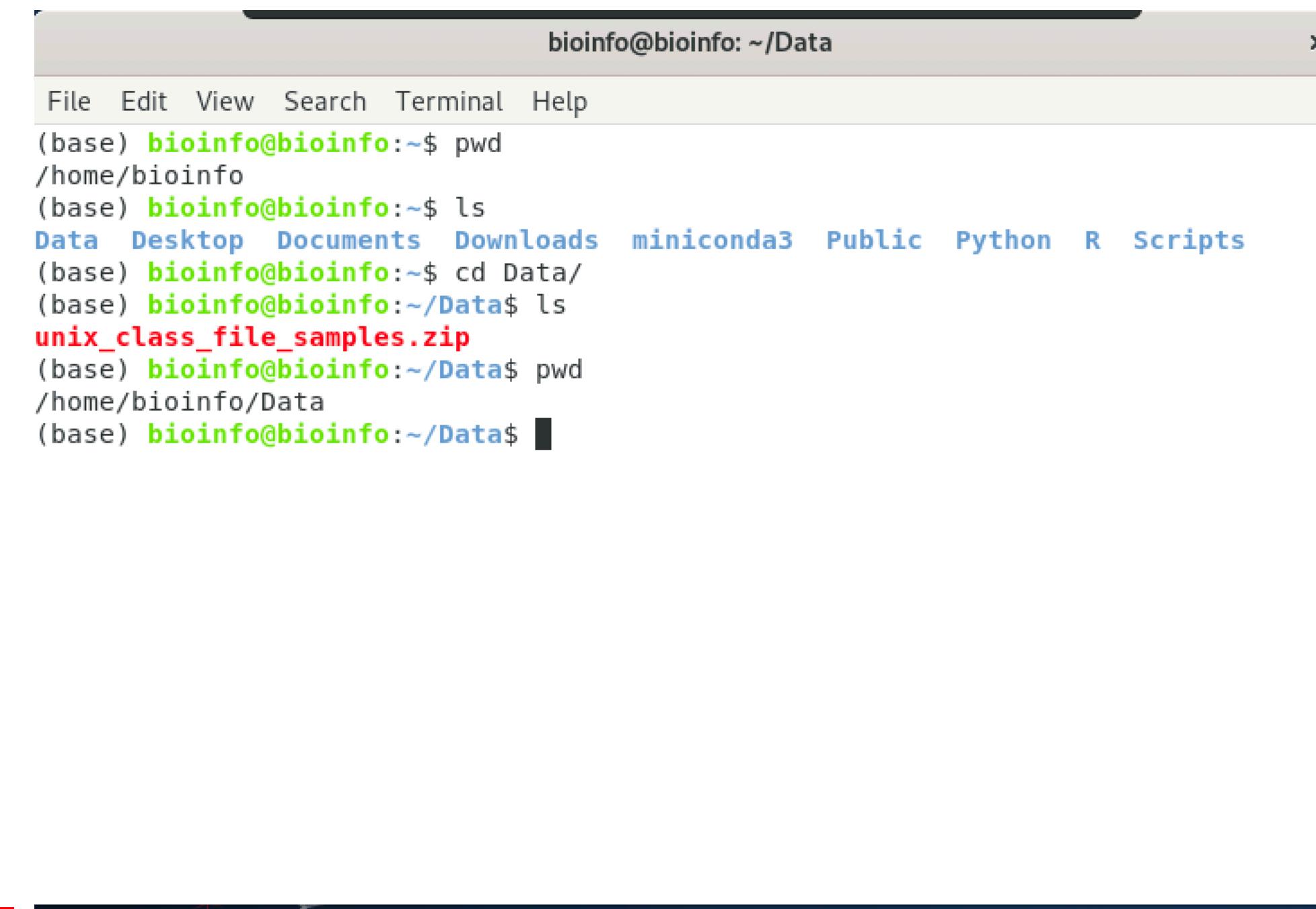
```
grep -c -A 3 --ignore-case file.txt
```

```
grep -c -A 3 --ignore-case file.txt
```

- Options often have a short form (starting with “-”) and a long form (starting with “--”).
- If an option flag does not have any arguments, then can often be concatenated
(e.g. “ls -h -l -a” is equivalent to “ls -hla”)

pwd, ls, and cd

- **pwd** — where am I?
- **ls** — what is inside the current directory?
- **cd** — change my current directory.



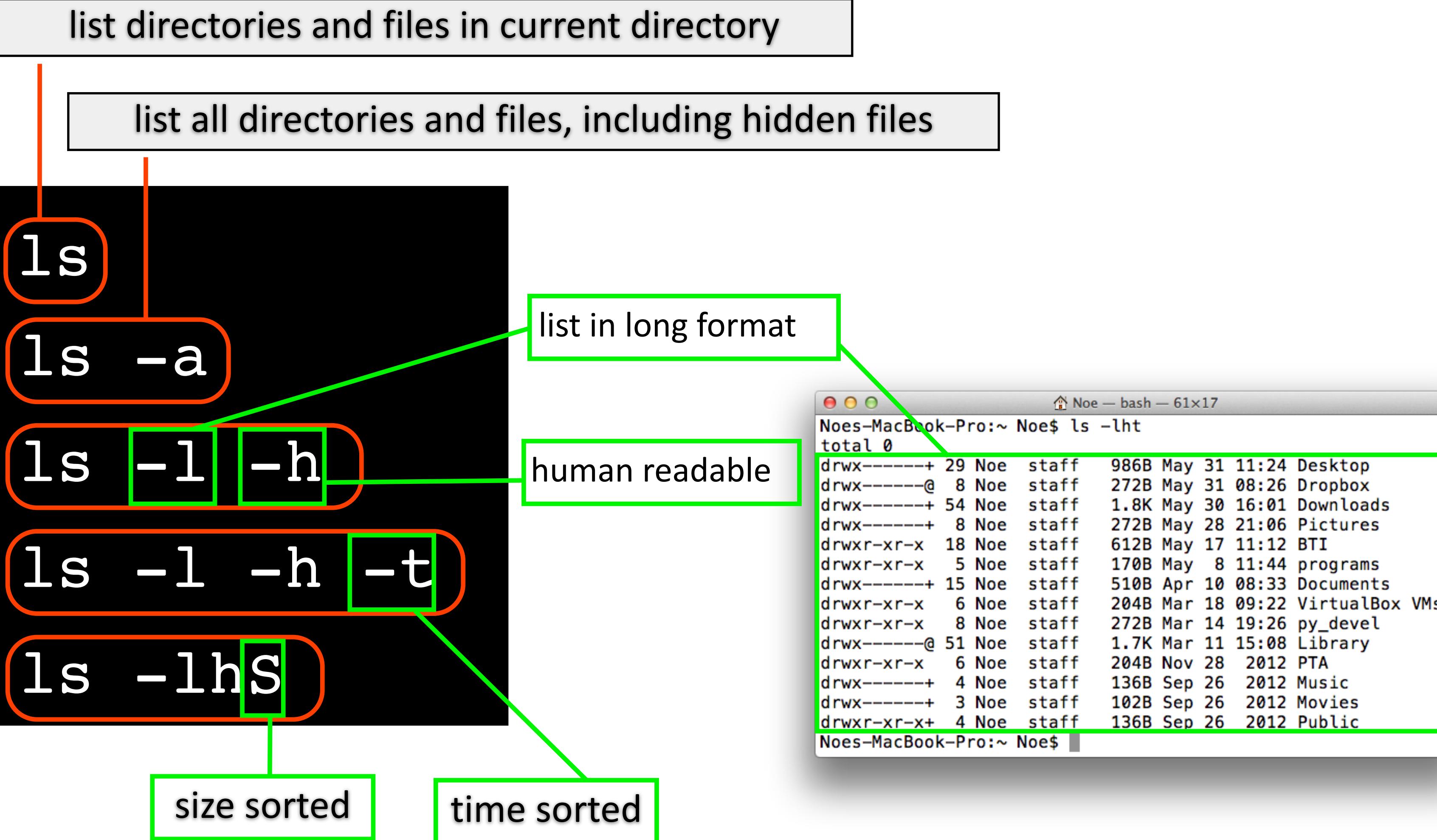
A screenshot of a terminal window titled "bioinfo@bioinfo: ~/Data". The window has a standard OS X-style title bar with a close button. Below the title bar is a menu bar with "File", "Edit", "View", "Search", "Terminal", and "Help". The main area of the terminal shows a command-line session:

```
bioinfo@bioinfo:~/Data
File Edit View Search Terminal Help
(base) bioinfo@bioinfo:~$ pwd
/home/bioinfo
(base) bioinfo@bioinfo:~$ ls
Data Desktop Documents Downloads miniconda3 Public Python R Scripts
(base) bioinfo@bioinfo:~$ cd Data/
(base) bioinfo@bioinfo:~/Data$ ls
unix_class_file_samples.zip
(base) bioinfo@bioinfo:~/Data$ pwd
/home/bioinfo/Data
(base) bioinfo@bioinfo:~/Data$
```

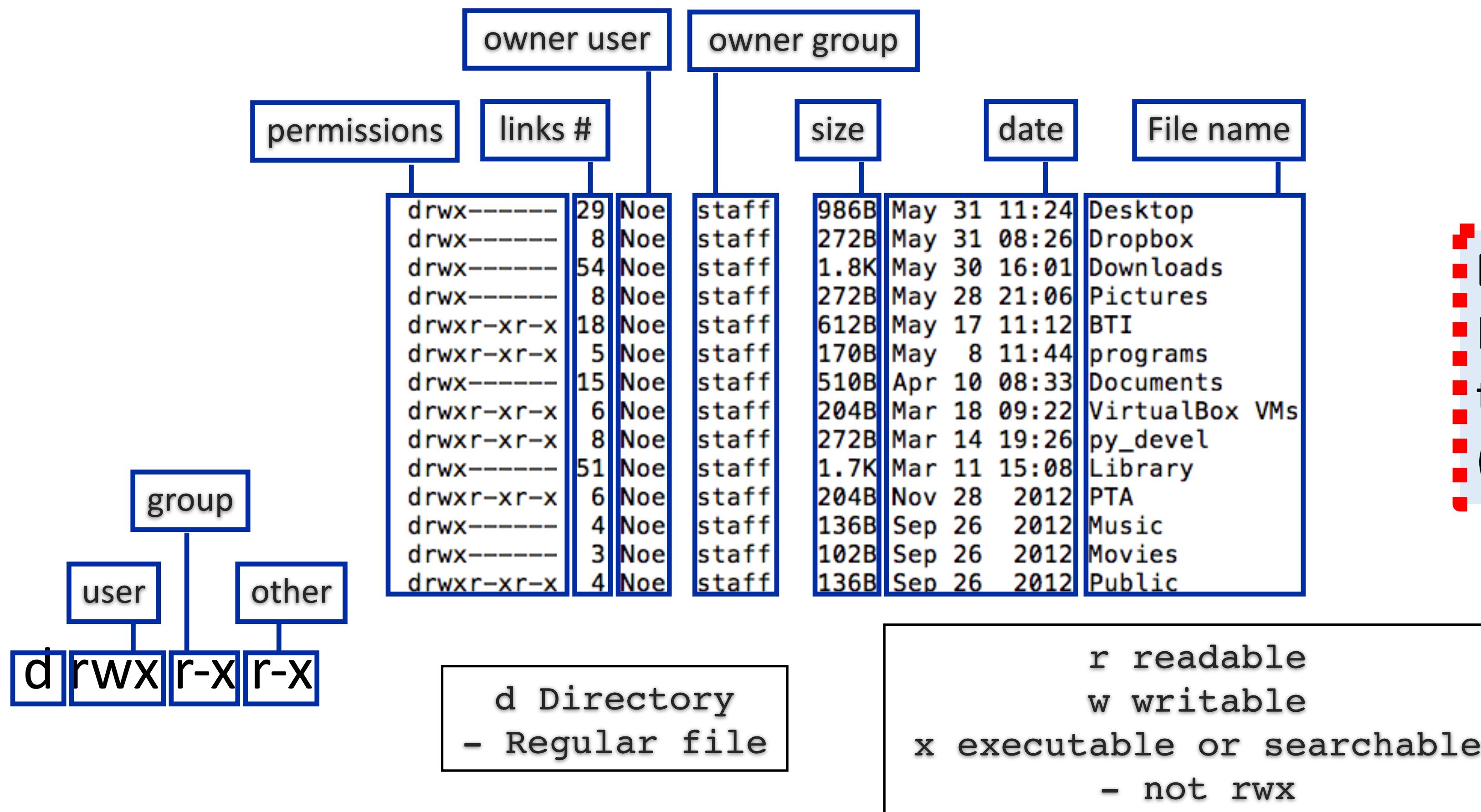
Exercise 4:

- Change into the “Data” directory and list the files found there.

ls lists files and directories



The ls list output



Exercise 5:

- List ALL files present in the root directory (including hidden files)

Wildcards, Shortcuts, and Command History

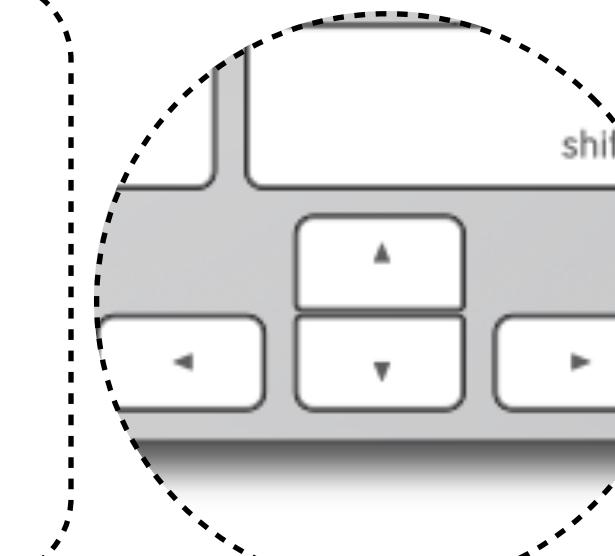
```
ls *txt
```

list all txt files in current directory

```
ls P*s
```

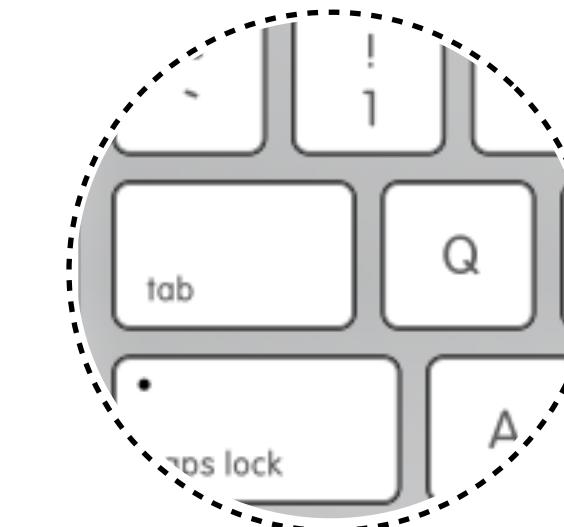
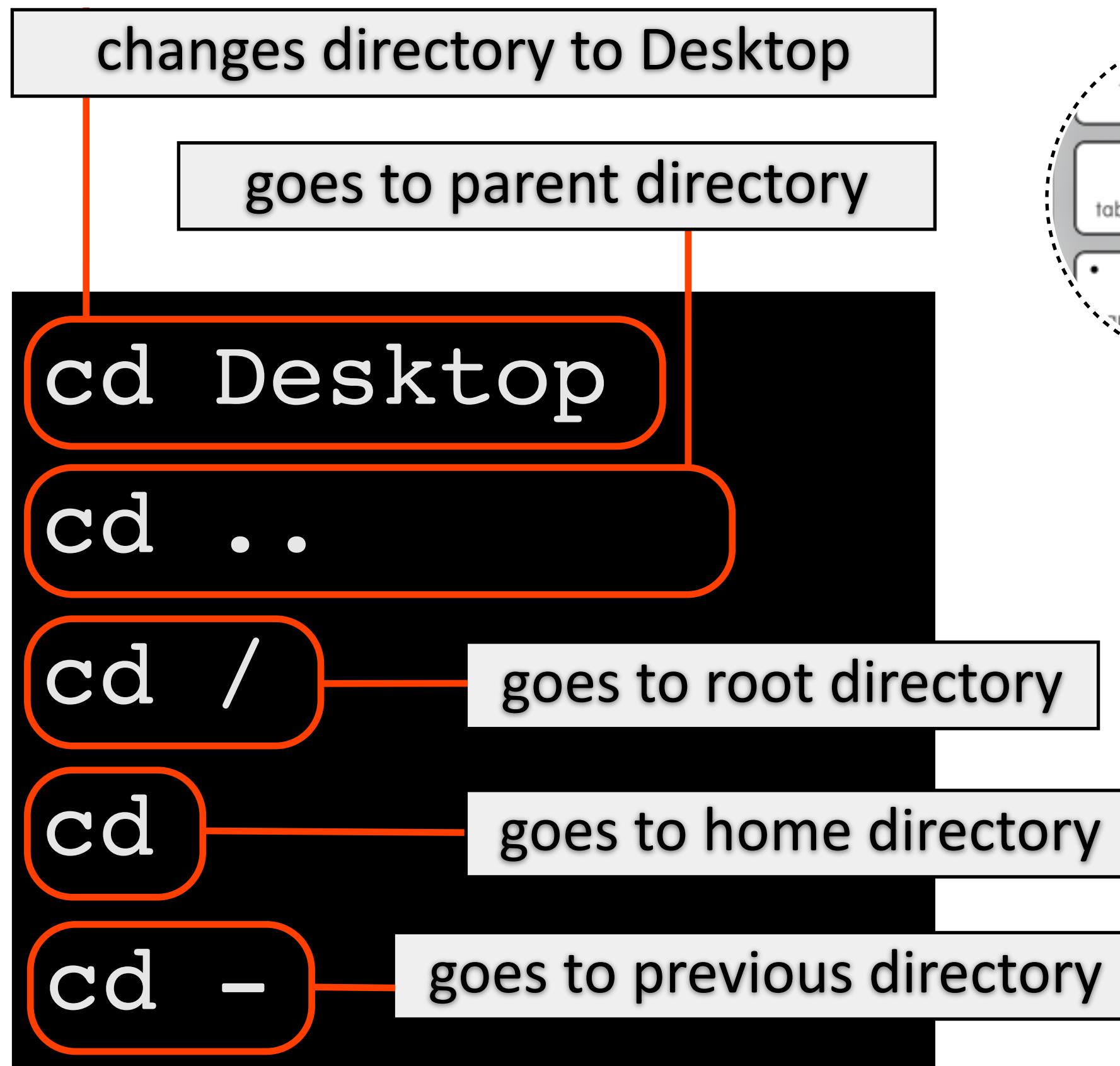
list files starting with P and ending with s,
e.g.: Pictures, Photos, Programs ...

- ctrl-c stop process
- ctrl-a go to begin of line
- ctrl-e go to end of line
- ctrl-r search in command history

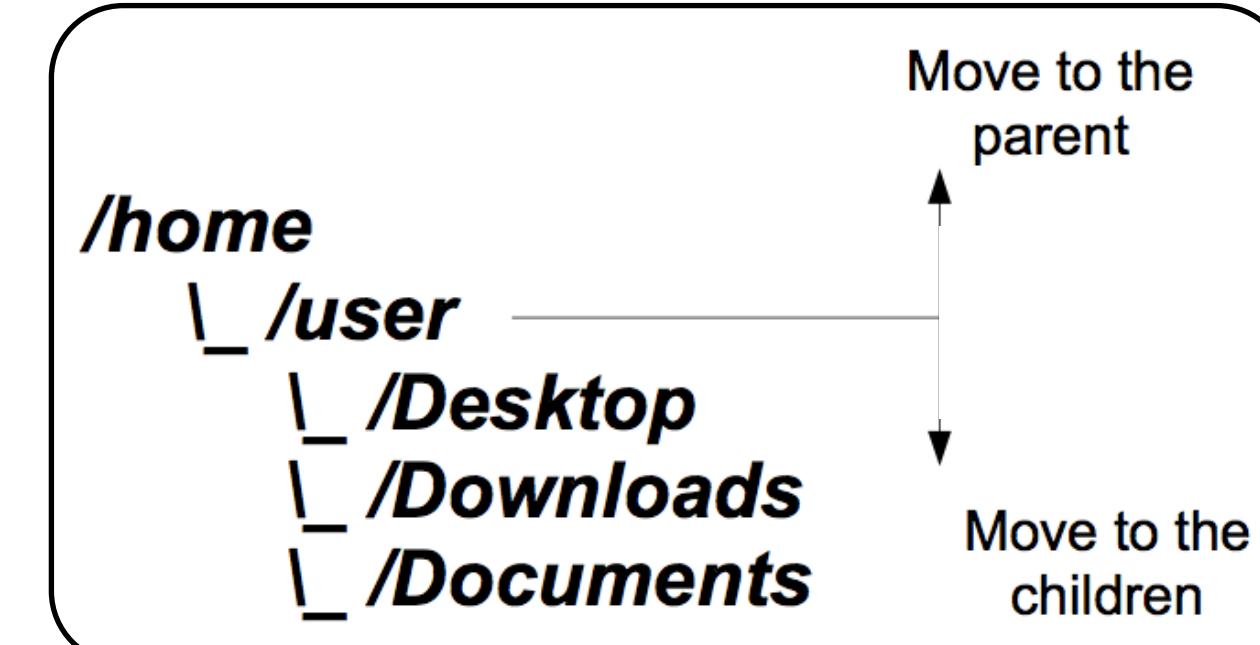


Use up and down
arrows to navigate
the command
history

cd changes directory



Use tab key to
autocomplete names



Absolute and relative paths

list files in Desktop using an absolute path

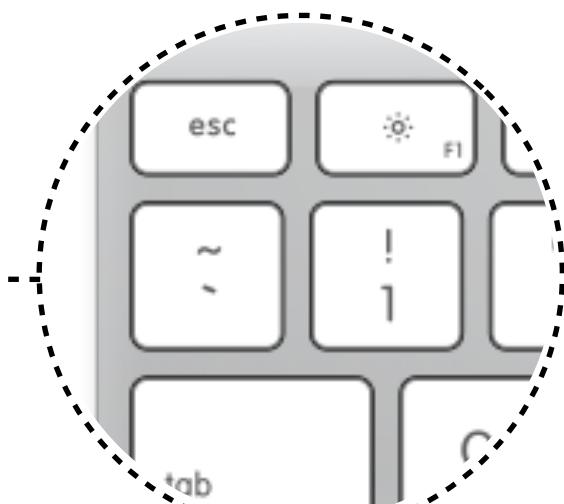
```
ls /home/user/Desktop
```

```
ls Desktop/
```

```
ls ~/Desktop
```

list files in Desktop using your home as a reference

list files in Documents using a relative path (from your home: /home/bioinfo)



Absolute and relative paths

Absolute paths do not depend on where you are

```
ls /home/bioinfo/Desktop
```

```
ls ~/Desktop
```

~/ is equivalent to /home/bioinfo/

Absolute and relative paths

```
goes to Desktop from when you are in your home (/home/bioinfo)
```

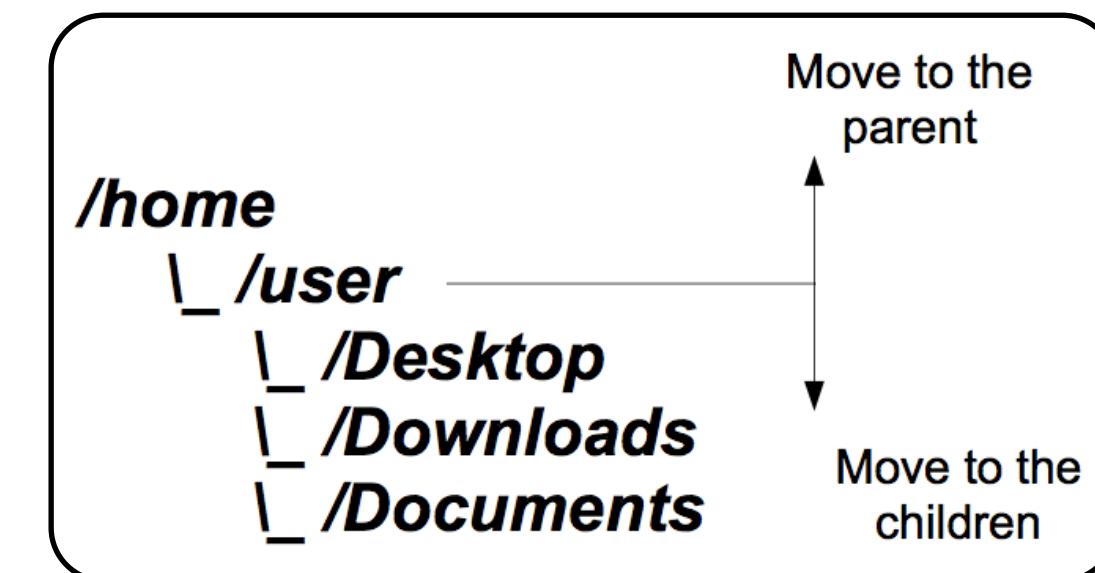
```
cd Desktop/
```

```
ls .. /Documents
```

```
list files from Documents when you are in Desktop
```

Exercise 6:

- List ALL files in the “/bin” directory that start
- with “ntfs” (do this once *without* changing your working directory, and once *with* changing your working directory)



Escaping special characters

! @ \$ ^ & * ~ ? . | / [] < > \ ` " ; # ()

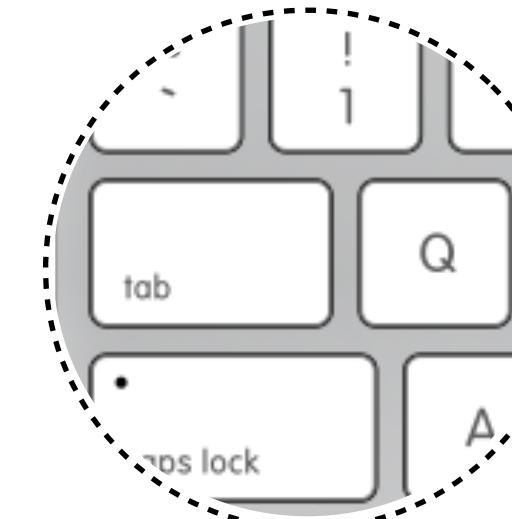
```
ls my_folder
```

list a folder

```
ls my\ folder
```

list a folder containing a space

Tip: file names in lower case
and with underscores
instead of spaces



Use tab key to
autocomplete names

Create, copy, move and delete files

creates an empty file called tmp_file.txt

copies tmp_file.txt in file_copy.txt

```
touch tmp_file.txt
```

```
cp tmp_file.txt file_copy.txt
```

```
mv file1.txt file2.txt
```

```
rm file.txt
```

Tip: name files in lower case
and with underscores
instead of using spaces

moves or rename a file

deletes file.txt

Locate a file

Locate the path for the file *unix_class_file_samples.zip*

```
locate unix_class_file_samples.zip
```

```
locate unix_class
```

Locate the path for all the files containing *unix_class*

Create, copy and delete directories

creates an empty directory called *dir_name*

deletes *dir_name* directory if it is empty

```
mkdir dir_name
```

```
rmdir dir_name
```

```
rm -r dir_name
```

```
cp -r dir_name dir_copy
```

copy *dir_name* and its files in a new folder



Music



Pictures



programs

Compression commands

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

groups and compress files

```
tar -zcvf file.tar.gz f1 f2
```

```
tar -zxvf file.tar.gz
```

decompress and ungroup a tar.gz file

BTI UNIX Command-Line Cheat Sheet
BTI-SGN Bioinformatics Course 2014

The diagram consists of several boxes connected by arrows. A red arrow points from the 'groups and compress files' box to the first tar command example. Another red arrow points from the same box to the 'Compression commands' table. A green arrow points from the 'decompress and ungroup a tar.gz file' box to the second tar command example. A final green arrow points from the same box to the 'files, directories or wildcards' table.

File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd dir	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch file_name	creates an empty file
cp file file_copy	copy a file
cp -r	copy files contained in directories
rm file	deletes a file
rm -r dir	deletes a directory and its files
mv file1 file2	moves or renames a file
mkdir dir_name	creates a directory
rmdir dir_name	deletes a directory
locate file_name	searches a file
man command	shows commands manual
top	shows process activity
df -h	shows disk space info

Text handling commands	
command > file	saves STDOUT in a file
command >> file	appends STDOUT in a file
cat file	concatenate and print files
cat file1 file2 > file3	merges files 1 and 2 into file3
cat *fasta > all.fasta	concatenates all fasta files in the current directory
head file	prints first lines from a file
head -n 5 file	prints first five lines from a file
tail file	prints last lines from a file
tail -n 5 file	prints last five lines from a file
less file	view a file
less -N file	includes line numbers
less -S file	wraps long lines
grep 'pattern' file	Prints lines matching a pattern
grep -c 'pattern' file	counts lines matching a pattern
cut -f 1,3 file	retrieves data from selected columns in a tab-delimited file
sort file	sorts lines from a file
sort -u file	sorts and return unique lines
uniq -c file	filters adjacent repeated lines
wc file	counts lines, words and bytes
paste file1 file2	concatenates the lines of input files
paste -d ","	concatenates the lines of input files by commas
sed	transforms text

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

Networking Commands	
wget URL	download a file from an URL
ssh user@server	connects to a server
scp	copy files between computers
apt-get install	installs applications in linux

files, directories or wildcards

Compression commands

compress file f1.txt in f1.txt.gz

compress files f1 and f2 in file.zip

gzip f1.txt

zip file.zip f1 f2

unzip file.zip

gunzip file.gz

decompress file.zip

decompress file.gz

Practice Exercises

- a) Go to your Desktop directory
- b) Create a file called:
Do not Use “special characters” in file names!.txt
- c) Delete the file you created in exercise b)
- d) Create a folder called unix_data in your desktop
- e) Find the file called unix_class_file_samples.zip
- f) Copy unix_class_fild_samples.zip to the folder unix_data, in your desktop
- g) Uncompress the file unix_class_file_samples.zip in /home/bioinfo/Desktop/unix_data
- h) Remove the _MACOSX folder

Solutions

- a) cd Desktop (**from your home: /home/bioinfo/**) **or** cd /home/bioinfo/Desktop **or** cd ~/Desktop
- b) touch Do\ not\ Use\ \"special\ characters\"\ in\ file\ names\!.txt
- c) rm Do\ not\ Use\ \"special\ characters\"\ in\ file\ names\!.txt **(use the tab key)**
- d) mkdir unix_data (**from Desktop: /home/bioinfo/Desktop**) **or** mkdir /home/bioinfo/Desktop/unix_data **or** mkdir ~/Desktop/unix_data
- e) locate unix_class_file_samples.zip
- f) cp /home/bioinfo/Data/unix_class_file_samples.zip /home/bioinfo/Desktop/unix_data
- g) cd /home/bioinfo/Desktop/unix_data && unzip unix_class_file_samples.zip
- h) rm -r _MACOSX

Networking Commands

● Networking commands



UNIX Command-Line Cheat Sheet
BTI-SGN Bioinformatics Course 2014

File system Commands	
ls	lists directories and files
ls -a	lists all files including hidden files
ls -lh	formatted list including more data
ls -t	lists sorted by date
pwd	returns path to working directory
cd dir	changes directory
cd ..	goes to parent directory
cd /	goes to root directory
cd	goes to home directory
touch file_name	creates an empty file
cp file file_copy	copy a file
cp -r	copy files contained in directories
rm file	deletes a file
rm -r dir	deletes a directory and its files
mv file1 file2	moves or renames a file
mkdir dir_name	creates a directory
rmdir dir_name	deletes a directory
locate file_name	searches a file
man command	shows commands manual
top	shows process activity
df -h	shows disk space info

Text handling commands	
command > file	saves STDOUT in a file
command >> file	appends STDOUT in a file
cat file	concatenate and print files
cat file1 file2 > file3	merges files 1 and 2 into file3
cat *.fasta > all.fasta	concatenates all fasta files in the current directory
head file	prints first lines from a file
head -n 5 file	prints first five lines from a file
tail file	prints last lines from a file
tail -n 5 file	prints last five lines from a file
less file	view a file
less -N file	includes line numbers
less -S file	wraps long lines
grep 'pattern' file	Prints lines matching a pattern
grep -c 'pattern' file	counts lines matching a pattern
cut -f 1,3 file	retrieves data from selected columns in a tab-delimited file
sort file	sorts lines from a file
sort -u file	sorts and return unique lines
uniq -c file	filters adjacent repeated lines
wc file	counts lines, words and bytes
paste file1 file2	concatenates the lines of input files
paste -d ","	concatenates the lines of input files by commas
sed	transforms text

Compression commands	
gzip/zip	compress a file
gunzip/unzip	decompress a file
tar -cvf	groups files
tar -xvf	ungroups files
tar -zcvf	groups and gzip files
tar -zxvf	gunzip and ungroups files

Networking Commands	
wget URL	download a file from an URL
ssh user@server	connects to a server
scp	copy files between computers
apt-get install	installs applications in linux

Networking Commands

connects your terminal to your account in a server

`ssh user_name@server_adress`

Downloads the BCBC logo!

`wget https://btiscience.org/wp-content/uploads/BCBClogo.png`

`scp noe@boyce.sgn.cornell.edu:/home/noe/file.txt .`

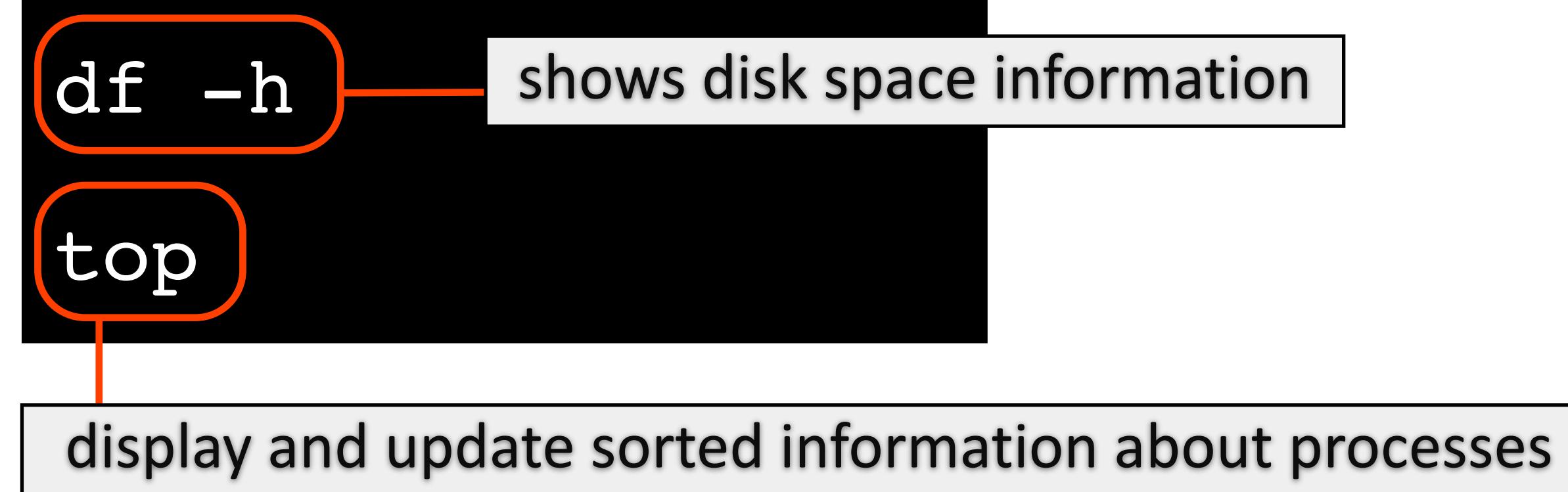
copy `file.txt` from your home in the server to the current directory in your computer

Tip: use the command `pwd` to get the path for `cp` and `scp`

Exercise 7:

- Download the BCBC logo (<https://btiscience.org/wp-content/uploads/BCBClogo.png>) to your Desktop.

Useful commands in the server



Processes

- Every running program is treated as a process
- Every process has a process ID and an “environment”
- Processes are created only from other processes through a *fork*. (parent ID)
- First process is init, with process ID 1
- Viewing processes: ps , jobs , top, pstree
- Terminating processes: kill

“Man” pages

Man pages are the documentation for UNIX commands

- \$ man <command>
- \$ man ls

Searching man pages

- Use the apropos command
- \$ apropos “text editor”

Controlling processes

- Interrupting, terminating execution
- control-Z , control-C
- Viewing running jobs (jobs)
- Background/foreground jobs (bg , fg, &)
- Use sleep 100 to test

Exercise 8:

Practice using bg and fg with the `sleep 100` command.

Top displays and update sorted information about processes

```
bioinfo@biodebian: ~
File Edit View Terminal Help
top - 15:07:10 up 3:50, 2 users, load average: 0.00, 0.00, 0.00
Tasks: 116 total, 1 running, 115 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.3%us, 0.0%sy, 0.0%ni, 99.7%id, 0.0%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 1026940k total, 518232k used, 508708k free, 51872k buffers
Swap: 1134584k total, 0k used, 1134584k free, 221056k cached

PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
132 root 20 0 0 0 0 S 0.3 0.0 0:07.05 ata/0
1861 bioinfo 20 0 40744 1700 1088 S 0.3 0.2 0:28.12 VBoxClient
2234 bioinfo 20 0 220m 12m 9792 S 0.3 1.3 0:00.61 gnome-terminal
2304 bioinfo 20 0 19072 1352 1012 R 0.3 0.1 0:00.03 top
1 root 20 0 8356 804 672 S 0.0 0.1 0:00.93 init
2 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kthreadd
3 root RT 0 0 0 0 S 0.0 0.0 0:00.00 migration/0
4 root 20 0 0 0 0 S 0.0 0.0 0:00.04 ksoftirqd/0
5 root RT 0 0 0 0 S 0.0 0.0 0:00.00 watchdog/0
6 root 20 0 0 0 0 S 0.0 0.0 0:00.15 events/0
7 root 20 0 0 0 0 S 0.0 0.0 0:00.00 cpuset
8 root 20 0 0 0 0 S 0.0 0.0 0:00.00 khelper
9 root 20 0 0 0 0 S 0.0 0.0 0:00.00 netns
10 root 20 0 0 0 0 S 0.0 0.0 0:00.00 async/mgr
11 root 20 0 0 0 0 S 0.0 0.0 0:00.00 pm
12 root 20 0 0 0 0 S 0.0 0.0 0:00.02 sync_supers
13 root 20 0 0 0 0 S 0.0 0.0 0:00.02 bdi-default
14 root 20 0 0 0 0 S 0.0 0.0 0:00.00 integrityd/0
15 root 20 0 0 0 0 S 0.0 0.0 0:00.06 kblockd/0
16 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kacpid
17 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kacpi_notify
18 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kacpi_hotplug
19 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kseriod
21 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kondemand/0
22 root 20 0 0 0 0 S 0.0 0.0 0:00.00 khungtaskd
23 root 20 0 0 0 0 S 0.0 0.0 0:00.00 kswapd0
24 root 25 5 0 0 0 S 0.0 0.0 0:00.00 ksmd
25 root 20 0 0 0 0 S 0.0 0.0 0:00.00 aio/0
26 root 20 0 0 0 0 S 0.0 0.0 0:00.00 crypto/0
130 root 20 0 0 0 0 S 0.0 0.0 0:00.00 ksuspend_usbd
131 root 20 0 0 0 0 S 0.0 0.0 0:00.00 khubd
```

q quit
u user (top -u user)
M sort by memory usage

Commands to install software

```
aptitude search blast
```

```
sudo aptitude install blast2
```

```
sudo apt-get install pbzip2
```

call the command with super user permissions

installs *pbzip2* in your computer

Exercise 9:
Install the package `fortune`

Text Editors for Programming

- Don't write code in Word!
- Vi, Vim, Emacs
- **Atom**
- Sublime
- Visual Studio

Homework

- Log on to the “william” server at BTI. User names and passwords will be supplied later today.
- Make a text file called class1_answers.txt. You can either make the file on the server or copy it to the server. Please use a text editor (ie. Not Word).
- Please record your answers in this file before the next class.
- Feel free to work with your team mate if you get stuck.