BGGN 213
Introduction to UNIX
Barry Grant
UC San Diego
http://thegrantlab.org/bggn213
Let's get started...

Mac
Terminal

PC
Git Bash
**SideNote: Terminal vs Shell**

- **Shell**: A command-line interface that allows a user to interact with the operating system by typing commands.

- **Terminal [emulator]**: A graphical interface to the shell (*i.e.* the window you get when you launch Git Bash/iTerm/etc.).

---

- **Setting Up**
  - **Mac users**: open a Terminal
  - **Windows users**: install MobaXterm and then open a terminal Shell prompt
Introduction To UNIX
Introduction To UNIX Shell
Print Working Directory: a.k.a. where the hell am I?
> pwd

List out the files and directories where you are
> ls

Q. What do you see after each command?

Q. Does it make sense if you compare to your Mac: Finder or Windows: File Explorer?

On Mac or PC type
> open .

Note the [SPACE] is important
# Download any file to your current directory/folder

```bash
> curl -O https://bioboot.github.io/bggn213_S18/class-material/bggn213_01_unix.zip
```

Q. Does what you see at each step make sense if you compare to your Mac: *Finder* or Windows: *File Explorer*?
# Download any file to your current directory/folder
> curl -O https://bioboot.github.io/bggn213_S18/class-material/bggn213_01_unix.zip

# List out the files and directories where you are (NB: Use TAB for auto-complete)
> ls bggn213_01_unix.zip

# Un-zip your downloaded file
> unzip bggn213_01_unix.zip

# Change directory (i.e. move to the folder named bggn213_01_unix)
> cd bggn213_01_unix

> ls
> pwd

Q. Does what you see at each step make sense if you compare to your Mac: Finder or Windows: File Explorer?

# Practice moving around the file system...
> cd projects
> ls
> pwd
> cd ..
Motivation

Why do we use Unix?
<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modularity</td>
<td>Core programs are modular and work well with others</td>
</tr>
<tr>
<td>Programmability</td>
<td>Best software development environment</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Access to existing tools and cutting-edge methods</td>
</tr>
<tr>
<td>Reliability</td>
<td>Unparalleled uptime and stability</td>
</tr>
<tr>
<td>Unix Philosophy</td>
<td>Encourages open standards</td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Modularity</td>
<td>Core programs are modular and work well with others</td>
</tr>
<tr>
<td>Programmability</td>
<td>Best software development environment</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Access to existing tools and cutting-edge methods</td>
</tr>
<tr>
<td>Reliability</td>
<td>Unparalleled uptime and stability</td>
</tr>
<tr>
<td>Unix Philosophy</td>
<td>Encourages open standards</td>
</tr>
</tbody>
</table>
Modularity

The Unix shell was designed to allow users to easily build complex workflows by interfacing smaller modular programs together.

An alternative approach is to write a single complex program that takes raw data as input, and after hours of data processing, outputs publication figures and a final table of results.
Advantages/Disadvantages

The ‘monster approach’ is customized to a particular project but results in massive, fragile and difficult to modify (therefore inflexible, untransferable, and error prone) code.

With modular workflows, it’s easier to:

• **Spot errors** and figure out where they’re occurring by inspecting intermediate results.

• **Experiment** with alternative methods by swapping out components.

• **Tackle novel problems** by remixing existing modular tools.
Unix ‘Philosophy’

“Write programs that do one thing and do it well. Write programs to work together and that encourage open standards. Write programs to handle text streams, because that is a universal interface.”

— Doug McIlroy
Unix family tree [1969-2010]

Source: https://commons.wikimedia.org/wiki/File:Unix_history-simple.svg
<table>
<thead>
<tr>
<th>Basics</th>
<th>File Control</th>
<th>Viewing &amp; Editing Files</th>
<th>Misc. useful</th>
<th>Power commands</th>
<th>Process related</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ls</code></td>
<td><code>mv</code></td>
<td><code>less</code></td>
<td><code>curl</code></td>
<td><code>grep</code></td>
<td><code>top</code></td>
</tr>
<tr>
<td><code>cd</code></td>
<td><code>cp</code></td>
<td><code>head</code></td>
<td><code>chmod</code></td>
<td><code>find</code></td>
<td><code>ps</code></td>
</tr>
<tr>
<td><code>pwd</code></td>
<td><code>mkdir</code></td>
<td><code>tail</code></td>
<td><code>wc</code></td>
<td><code>sed</code></td>
<td><code>kill</code></td>
</tr>
<tr>
<td><code>man</code></td>
<td><code>rm</code></td>
<td><code>nano</code></td>
<td><code>echo</code></td>
<td><code>sudo</code></td>
<td><code>Crl-c</code></td>
</tr>
<tr>
<td><code>ssh</code></td>
<td><code>I</code> (pipe)</td>
<td><code>touch</code></td>
<td><code>source</code></td>
<td><code>git</code></td>
<td><code>Crl-z</code></td>
</tr>
<tr>
<td><code>scp</code></td>
<td><code>&gt;</code> (write to file)</td>
<td></td>
<td><code>cat</code></td>
<td><code>R</code></td>
<td><code>bg</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;</code> (read from file)</td>
<td></td>
<td><code>tmux</code></td>
<td><code>python</code></td>
<td><code>fg</code></td>
</tr>
<tr>
<td>Basics</td>
<td>File Control</td>
<td>Viewing &amp; Editing Files</td>
<td>Misc. useful</td>
<td>Power commands</td>
<td>Process related</td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>ls</td>
<td>mv</td>
<td>less</td>
<td>curl</td>
<td>grep</td>
<td>top</td>
</tr>
<tr>
<td>cd</td>
<td>cp</td>
<td>head</td>
<td>chmod</td>
<td>find</td>
<td>ps</td>
</tr>
<tr>
<td>pwd</td>
<td>mkdir</td>
<td>tail</td>
<td>wc</td>
<td>sed</td>
<td>kill</td>
</tr>
<tr>
<td>man</td>
<td>rm</td>
<td>nano</td>
<td>echo</td>
<td>sudo</td>
<td>Crl-c</td>
</tr>
<tr>
<td>ssh</td>
<td>(pipe)</td>
<td>touch</td>
<td>source</td>
<td>git</td>
<td>Crl-z</td>
</tr>
<tr>
<td>scp</td>
<td>(write to file)</td>
<td>cat</td>
<td>R</td>
<td>bg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(read from file)</td>
<td>tmux</td>
<td>python</td>
<td>fg</td>
<td></td>
</tr>
</tbody>
</table>
Beginning Unix

Getting started with basic Unix commands
Information in the file system is stored in files, which are stored in directories (folders). Directories can also store other directories, which forms a directory tree.

The forward slash character ‘/’ is used to represent the root directory of the whole file system, and is also used to separate directory names. E.g. /home/jono/work/bggn213_notes.txt
## Basics: Using the filesystem

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ls</strong></td>
<td>List files and directories</td>
</tr>
<tr>
<td><strong>cd</strong></td>
<td>Change directory (i.e. move to a different ‘folder’)</td>
</tr>
<tr>
<td><strong>pwd</strong></td>
<td>Print working directory (which folder are you in)</td>
</tr>
<tr>
<td><strong>mkdir</strong></td>
<td>Make a new Directory</td>
</tr>
<tr>
<td><strong>cp</strong></td>
<td>Copy a file or directory to somewhere else</td>
</tr>
<tr>
<td><strong>mv</strong></td>
<td>Move a file or directory (basically rename)</td>
</tr>
<tr>
<td><strong>rm</strong></td>
<td>Remove a file or directory</td>
</tr>
</tbody>
</table>
**Side Note: File Paths**

- **An absolute path** specifies a location from the root of the file system. E.g. `/home/jono/work/bggn213_notes.txt`

- **A relative path** specifies a location starting from the current location. E.g. `./bggn213_notes.txt`

**Some useful shortcuts:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>Single dot ‘.’ (for current directory)</td>
</tr>
<tr>
<td>..</td>
<td>Double dot ‘..’ (for parent directory)</td>
</tr>
<tr>
<td>~</td>
<td>Tilda ‘~’ (for your home directory)</td>
</tr>
<tr>
<td>[Tab]</td>
<td>Pressing the tab key can autocomplete names</td>
</tr>
</tbody>
</table>
Finding the Right Hammer (man and apropos)

- You can access the manual (i.e. user documentation) on a command with `man`, e.g:
  ```
  > man pwd
  ```
- The man page is only helpful if you know the name of the command you’re looking for. `apropos` will search the man pages for keywords.
  ```
  > apropos "working directory"
  ```

Windows only: `pwd --help`
Inspecting text files

- **less** - visualize a text file:
  - use arrow keys
  - page down/page up with “space”/“b” keys
  - search by typing "/"
  - **quit** by typing "q"

- Also see: **head, tail, cat, more**
Creating text files

Creating files can be done in a few ways:
- With a **text editor** (such as **nano**, **emacs**, or **vi**)
- With the **touch** command (**> touch a_file**)
- From the command line with **cat** or **echo** and **redirection** (more on this later)

- **nano** is a simple text editor that is recommended for first-time users. Other text editors have more powerful features but also steep learning curves
Creating and editing text files with **nano**

In the terminal type:

```bash
> nano yourfilename.txt
```

- There are many other text file editors (e.g. **vim**, **emacs** and **sublime text**, **atom** etc.)
Connecting to remote machines (with **ssh**)

Most high-performance computing (HPC) resources can only be accessed by ssh (**Secure SHell**)

> ssh [user@host.address]

For example:

> ssh barry@bio3d.ucsd.edu

<table>
<thead>
<tr>
<th>User</th>
<th>Host address</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>barry</strong></td>
<td><strong>bio3d.ucsd.edu</strong></td>
</tr>
</tbody>
</table>

> ssh -i ~/bggn213_private_key tb170077@IP_ADDRESS

<table>
<thead>
<tr>
<th>Optional key file</th>
<th>User</th>
<th>Host address</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>~/.ssh/bgg213</strong></td>
<td><strong>tb170077</strong></td>
<td><strong>IP_ADDRESS</strong></td>
</tr>
</tbody>
</table>
Connecting to jetstream (with **ssh**)

• We will go through this process together in the last hands-on sections today!

✅ Downloaded our class specific **jetstream keyfile**
(See web link on board: This is required for connecting to jetstream virtual machines later.)

✅ Use your assigned **IP_ADDRESS**

> **ssh** -i ~/bggn213_private_key  **tb170077@IP_ADDRESS**

  **Optional key file**  **User**  **Host address**
Copying to and from remote machines (**scp**)

- The **scp** (**S**ecure **C**o**Py**) command can be used to copy files and directories from one computer to another.

```
> scp [file] [user@host]:[destination]
> scp localfile.txt barry@bigcomputer.net:/remotedir/.
```
<table>
<thead>
<tr>
<th>Basics</th>
<th>File Control</th>
<th>Viewing &amp; Editing Files</th>
<th>Misc. useful</th>
<th>Power commands</th>
<th>Process related</th>
</tr>
</thead>
<tbody>
<tr>
<td>ls</td>
<td>mv</td>
<td>less</td>
<td>chmod</td>
<td>grep</td>
<td>top</td>
</tr>
<tr>
<td>cd</td>
<td>cp</td>
<td>head</td>
<td>echo</td>
<td>find</td>
<td>ps</td>
</tr>
<tr>
<td>pwd</td>
<td>mkdir</td>
<td>tail</td>
<td>wc</td>
<td>sed</td>
<td>kill</td>
</tr>
<tr>
<td>man</td>
<td>rm</td>
<td>nano</td>
<td>curl</td>
<td>sudo</td>
<td>Crl-c</td>
</tr>
<tr>
<td>ssh</td>
<td>(pipe)</td>
<td>touch</td>
<td>source</td>
<td>git</td>
<td>Crl-z</td>
</tr>
<tr>
<td>scp</td>
<td>&gt; (write to file)</td>
<td></td>
<td>cat</td>
<td>R</td>
<td>bg</td>
</tr>
<tr>
<td>scp</td>
<td>&lt; (read from file)</td>
<td></td>
<td>tmux</td>
<td>python</td>
<td>fg</td>
</tr>
</tbody>
</table>
**Process** refers to a running instance of a program

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>top</strong></td>
<td>Provides a real-time view of all running processes</td>
</tr>
<tr>
<td><strong>ps</strong></td>
<td>Report a snapshot of the current processes</td>
</tr>
<tr>
<td><strong>kill</strong></td>
<td>Terminate a process (the “force quit” of the unix world)</td>
</tr>
<tr>
<td>Crl-c</td>
<td>Stop a job</td>
</tr>
<tr>
<td>Crl-z</td>
<td>Suspend a job</td>
</tr>
<tr>
<td><strong>bg</strong></td>
<td>Resume a suspended job in the background</td>
</tr>
<tr>
<td><strong>fg</strong></td>
<td>Resume a suspended job in the foreground</td>
</tr>
<tr>
<td>&amp;</td>
<td>Start a job in the background</td>
</tr>
</tbody>
</table>
Hands-on time

Sections 1 to 3 of software carpentry UNIX lesson
https://swcarpentry.github.io/shell-novice/

https://explainshell.com ~20 mins
• Many bioinformatics tools can only be used through a command line interface, or have extra capabilities in the command line version that are not available in the GUI.

• The shell makes your work less error-prone, more reproducible and less boring allowing you to automate repetitive tasks and concentrate on more exciting things.

• Many bioinformatic tasks require large amounts of computing power and can’t realistically be run on your own machine. These tasks are best performed using remote computers or cloud computing, which can only be accessed through a shell.
Working with Unix

How do we actually use Unix?
Combining Utilities with Redirection (>, <) and Pipes (|)

• The power of the shell lies in the ability to combine simple utilities (i.e. commands) into more complex algorithms very quickly.

• A key element of this is the ability to send the output from one command into a file or to pass it directly to another program.

• This is the job of >, < and |
Side-Note: Standard Input and Standard Output streams

Two very important concepts that unpin Unix workflows:

- **Standard Output (stdout)** - default destination of a program's output. It is generally the terminal screen.

- **Standard Input (stdin)** - default source of a program's input. It is generally the command line.
Output redirection and piping

> ls ~/Desktop  # stdin is "~/Desktop"; stdout to screen
Output redirection and piping

- `ls ~/Desktop` # stdin is "~/Desktop"; stdout to screen
- `ls ~/Desktop > mylist.txt` # stdout redirected to file
- `ls ~/Desktop | less` # stdout piped to less (no file created)
Output redirection and piping

> ls ~/Desktop # stdin is "~/Desktop"; stdout to screen

> ls ~/Desktop > mylist.txt # stdout redirected to file

> ls ~/Desktop | less # stdout piped to less (no file created)

> ls -l ~/Desktop # extra optional input argument "-l"
Output redirection and piping

- `ls ~/Desktop`  # stdin is “~/Desktop”; stdout to screen
- `ls ~/Desktop > mylist.txt`  # stdout redirected to file
- `ls ~/Desktop | less`  # stdout piped to less (no file created)
- `ls /nodirexists/ 2> binlist.txt`  # stderr to file
Output redirection and piping

> `ls ~/Desktop`  # stdin is “~/Desktop”; stdout to screen

> `ls ~/Desktop > mylist.txt`  # stdout redirected to file

> `ls ~/Desktop | less`  # stdout piped to less (no file created)

> `ls /nodirexists/ 2> binlist.txt`  # stderr to file
# Output redirection and piping

> `ls ~/Desktop`  
  # stdin is “~/Desktop”; stdout to screen

> `ls ~/Desktop > mylist.txt`  
  # stdout redirected to file

> `ls ~/Desktop | less`  
  # stdout piped to less (no file created)

> `ls /nodirexists/ 2> binlist.txt`  
  # stderr to file

Do it Yourself!
Output redirection summary

- `stdin`: Input from file descriptors
- `stdout`: Output to file descriptors
- `stderr`: Error output to file descriptors
- `args`: Command line arguments
- `disk`: Disk storage
- `network`: Network communication
- `I/O`: Input/output operations
- `X11`: Display window

Flow diagram showing input and output redirection in a program.
ls -l
ls -l > list_of_files
ls -l | grep partial_name > list_of_files

We have piped ( | ) the stdout of one command into the stdin of another command!
ls -l /usr/bin/ | grep "tree" > list_of_files

grep: prints lines containing a string. Also searches for strings in text files.
<table>
<thead>
<tr>
<th>Basics</th>
<th>File Control</th>
<th>Viewing &amp; Editing Files</th>
<th>Misc. useful</th>
<th>Power commands</th>
<th>Process related</th>
</tr>
</thead>
<tbody>
<tr>
<td>ls</td>
<td>mv</td>
<td>less</td>
<td>chmod</td>
<td>grep</td>
<td>top</td>
</tr>
<tr>
<td>cd</td>
<td>cp</td>
<td>head</td>
<td>echo</td>
<td>find</td>
<td>ps</td>
</tr>
<tr>
<td>pwd</td>
<td>mkdir</td>
<td>tail</td>
<td>wc</td>
<td>sed</td>
<td>kill</td>
</tr>
<tr>
<td>man</td>
<td>rm</td>
<td>nano</td>
<td>curl</td>
<td>sudo</td>
<td>Crl-c</td>
</tr>
<tr>
<td>ssh</td>
<td>l (pipe)</td>
<td>touch</td>
<td>source</td>
<td>git</td>
<td>Crl-z</td>
</tr>
<tr>
<td>scp</td>
<td>&gt; (write to file)</td>
<td>cat</td>
<td>R</td>
<td>bg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; (read from file)</td>
<td>tmux</td>
<td>python</td>
<td>fg</td>
<td></td>
</tr>
</tbody>
</table>
Side-Note: grep ‘power command’

• grep - prints lines containing a string pattern. Also searches for strings in text files, e.g.

```
> cd ~/bggn213_01_unix/projects/transducin/sequences/data/
> grep --color "GESGKS" seqdump.fasta
```

```
REVKLLLLLGA<hi>GESGKS</hi>TIVKQMKIIHEAGYSEEECKQYK
```

• grep is a ‘power tool’ that is often used with pipes as it accepts regular expressions as input (e.g. “G..GK[ST]”) and has lots of useful options - see the man page for details.
grep example using regular expressions

• Suppose a program that you are working with complains that your input sequence file contains non-nucleotide characters. You can eye-ball your file or …

```bash
$ grep -v "^>" seqdump.fasta | grep --color "[^ATGC]"
```

**Exercises:**

1. Use “man grep” (or on windows “grep --help”) to find out what the -v argument option is doing!
2. How could we also show line number for each match along with the output? (tip you can grep the output of “man grep” for ‘line number’)
Suppose a program that you are working with complains that your input sequence file contains non-nucleotide characters. You can eye-ball your file or …

First we remove (with -v option) lines that start with a “>” character (these are sequence identifiers).

Next we find characters that are not A, T, C or G. To do this we use ^ symbols second meaning: match anything but the pattern in square brackets. We also print line number (with -n option) and color output (with --color option).

```
> grep -v "^>" seqdump.fasta | grep --color -n "[^ATGC]"
```
Key Point: Pipes and redirects avoid unnecessary i/o

- Disc i/o is often a bottleneck in data processing!
- Pipes prevent unnecessary disc i/o operations by connecting the stdout of one process to the stdin of another (these are frequently called "streams")

```bash
> program1 input.txt 2> program1.stderr | \ program2 2> program2.stderr > results.txt
```
- Pipes and redirects allow us to build solutions from modular parts that work with stdin and stdout streams.
Unix ‘Philosophy’ Revisited

“Write programs that do one thing and do it well. Write programs to work together and that encourage open standards. Write programs to handle text streams, because that is a universal interface.”

— Doug McIlory
Pipes provide speed, flexibility and sometimes simplicity...

- In 1986 “Communications of the ACM magazine” asked famous computer scientist Donald Knuth to write a simple program to count and print the $k$ most common words in a file alongside their counts, in descending order.

- Kunth wrote a literate programming solution that was 7 pages long, and also highly customized to this problem (e.g. Kunth implemented a custom data structure for counting English words).

- Doug McIlroy replied with one line:

  ```bash
  cat input.txt | tr A-Z a-z | sort | uniq -c | sort -rn | sed 10q
  ```
Key Point:

You can chain any number of programs together to achieve your goal!

This allows you to build up fairly complex workflows within one command-line.
<table>
<thead>
<tr>
<th>Basics</th>
<th>File Control</th>
<th>Viewing &amp; Editing Files</th>
<th>Misc. useful</th>
<th>Power commands</th>
<th>Process related</th>
</tr>
</thead>
<tbody>
<tr>
<td>ls</td>
<td>mv</td>
<td>less</td>
<td>chmod</td>
<td>grep</td>
<td>top</td>
</tr>
<tr>
<td>cd</td>
<td>cp</td>
<td>head</td>
<td>echo</td>
<td>find</td>
<td>ps</td>
</tr>
<tr>
<td>pwd</td>
<td>mkdir</td>
<td>tail</td>
<td>wc</td>
<td>sed</td>
<td>kill</td>
</tr>
<tr>
<td>man</td>
<td>rm</td>
<td>nano</td>
<td>curl</td>
<td>sudo</td>
<td>Crl-c</td>
</tr>
<tr>
<td>ssh</td>
<td>l (pipe)</td>
<td>touch</td>
<td>source</td>
<td>git</td>
<td>Crl-z</td>
</tr>
<tr>
<td></td>
<td>&gt; (write to file)</td>
<td>cat</td>
<td>R</td>
<td>bg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt; (read from file)</td>
<td>python</td>
<td></td>
<td></td>
<td>fg</td>
</tr>
</tbody>
</table>
Hands-on time

Using Jetstream for Bioinformatics

• Using remote UNIX machines (Part I)
• Using remote UNIX machines (Part II)
• Using remote UNIX machines (Part III)

~35 mins
# New commands

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sudo</strong></td>
<td>Execute a command with root permissions</td>
</tr>
<tr>
<td><strong>apt-get</strong></td>
<td>Package handling utility for updating &amp; installing software</td>
</tr>
<tr>
<td><strong>curl</strong></td>
<td>Download data</td>
</tr>
<tr>
<td><strong>gunzip</strong></td>
<td>File compression and decompression</td>
</tr>
<tr>
<td><strong>blastp</strong></td>
<td>Command line BLAST</td>
</tr>
<tr>
<td><strong>shmlast</strong></td>
<td>Mapping orthologs from RNA-seq data</td>
</tr>
<tr>
<td><strong>?</strong></td>
<td>Use <em>man</em> to find out about other new commands</td>
</tr>
</tbody>
</table>
How to Get Working

Best practices for organizing your computational biology projects
All files and directories used in your project should live in a single project directory.

- Use sub-directories to divide your project into sub-projects.
- Do not use spaces in file and directory names!

Document your methods and workflows with plain text README files

- Also document the origin of all data in your project directory
- Also document the versions of the software that you ran and the options you used.
- Consider using Markdown for your documentation.

Use version control and backup to multiple destinations!

Be reproducible:
http://ropensci.github.io/reproducibility-guide/sections/introduction/
Geeks and repetitive tasks

time spent

<table>
<thead>
<tr>
<th>Time Spent</th>
<th>Task Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>geek</td>
<td>wins</td>
</tr>
<tr>
<td>non-geek</td>
<td>loses</td>
</tr>
</tbody>
</table>

- runs script
- gets annoyed
- does it manually
- writes script to automate
- makes fun of geek's complicated method

Task size increases along the horizontal axis, and time spent increases along the vertical axis.
Summary

• Built-in unix shell commands allow for easy data manipulation (e.g. sort, grep, etc.)

• Commands can be easily combined to generate flexible solutions to data manipulation tasks.

• The unix shell allows users to automate repetitive tasks through the use of shell scripts that promote reproducibility and easy troubleshooting

• Introduced the 21 key unix commands that you will use during ~95% of your future unix work…