



# Welcome to the Workshop!

## **Workshop Website**

<http://sarahlrstevens.info/2017-08-02-chicago-frb/>

## **Etherpad (for collaborative notes)**

<http://pad.software-carpentry.org/2017-08-02-chicago-frb>

## **Pre-Workshop Survey**

[https://www.surveymonkey.com/r/swc\\_pre\\_workshop\\_v1](https://www.surveymonkey.com/r/swc_pre_workshop_v1)



# Check your setup!

Open GitBash (Windows), and type the following:

**python --version**

Raise your hand if you DON'T have version **3.5.x** (where x can vary).

**python -c "import pandas"**

Raise your hand if you get an error.

**git --version**

Raise your hand if you get an error.



# Welcome to Software Carpentry!

## August 2-3, 2017

### Instructors:

Prateek Mehta  
Sarah Stevens

### Hosts:

Deng Pan  
Ken Housinger



# What is 'Software Carpentry'??

- Non-profit, international organization
- Teaches workshops to help researchers adopt reproducible computational practices
- Instructors are all volunteers
- Materials developed by open science community
- Code-along learning model



# Code of Conduct

- <http://software-carpentry.org/conduct/>
- Harassment includes offensive verbal comments related to gender, sexual orientation, disability, physical appearance, body size, race, religion, sexual images in public spaces, deliberate intimidation, stalking, following, harassing photography or recording, sustained disruption of talks or other events, inappropriate physical contact, and unwelcome sexual attention.
- All communication should be appropriate for a professional audience including people of many different backgrounds. Sexual language and imagery is not appropriate for any event.
- Be kind to others. Do not insult or put down other attendees.
- Behave professionally. Remember that harassment and sexist, racist, or exclusionary jokes are not appropriate.



# Workshop Logistics: Where Stuff Is

## **Restrooms**

Across the hall

## **Beverages**

Drinking fountain: across the hall

Coffee/tea: front left corner of room, all day

## **Lunch**

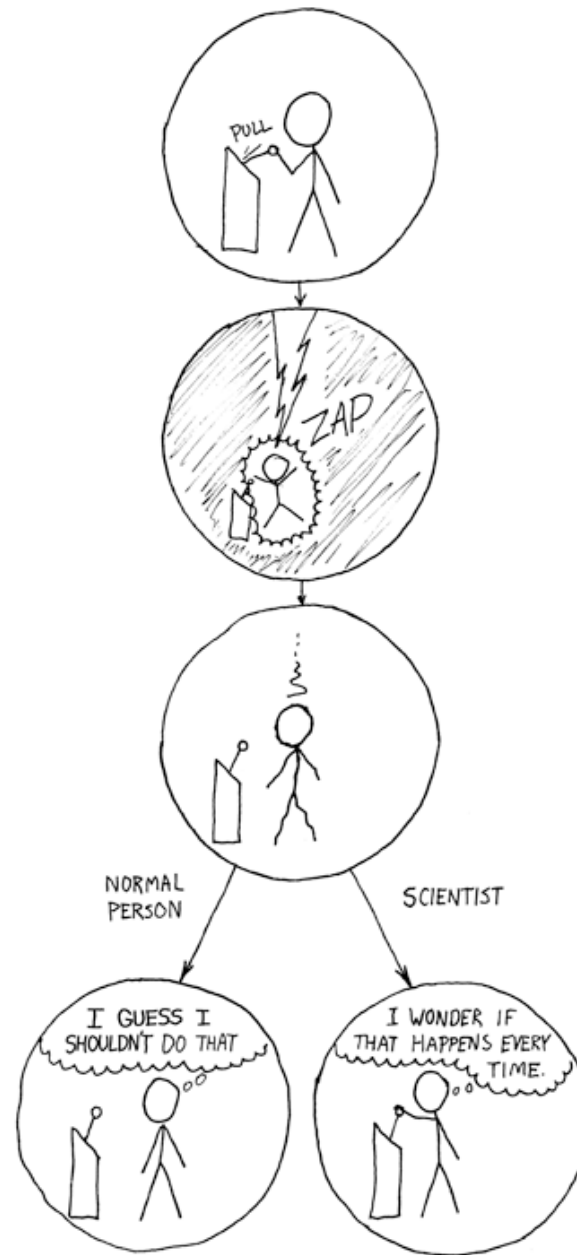
On your own.

Need a fridge? Let us know.



If You Can't Reproduce It,  
Is It Still Science?  
And how long will it take?

Inspired by Greg Wilson,  
Software Carpentry







# Reality of Research Computing

- Many scientists/researchers spend most of their time developing, maintaining, or running software
  - Most don't consider themselves software engineers
  - Few have ever been taught how



# So what?

- Most results take longer to produce than they need to
- Difficult to assess quality



# Software Carpentry to the Rescue

- Best practices used by the best software engineers whose business is development of quality software
  - They don't always have formal training
  - They don't always follow all the practices
  - Growing evidence supported by empirical studies



# Software Carpentry Practices

- Write software for people, not computers
- Automate repetitive tasks
- Use the computer to record history
- Make incremental changes
- Use version control
- Don't repeat yourself
- Plan for mistakes
- First make it correct, then make it fast
- Document design & purpose, not just mechanics
- Conduct code reviews

Wilson et al. (2014) Best practices for scientific computing. PLoS Biology 12: e1001745



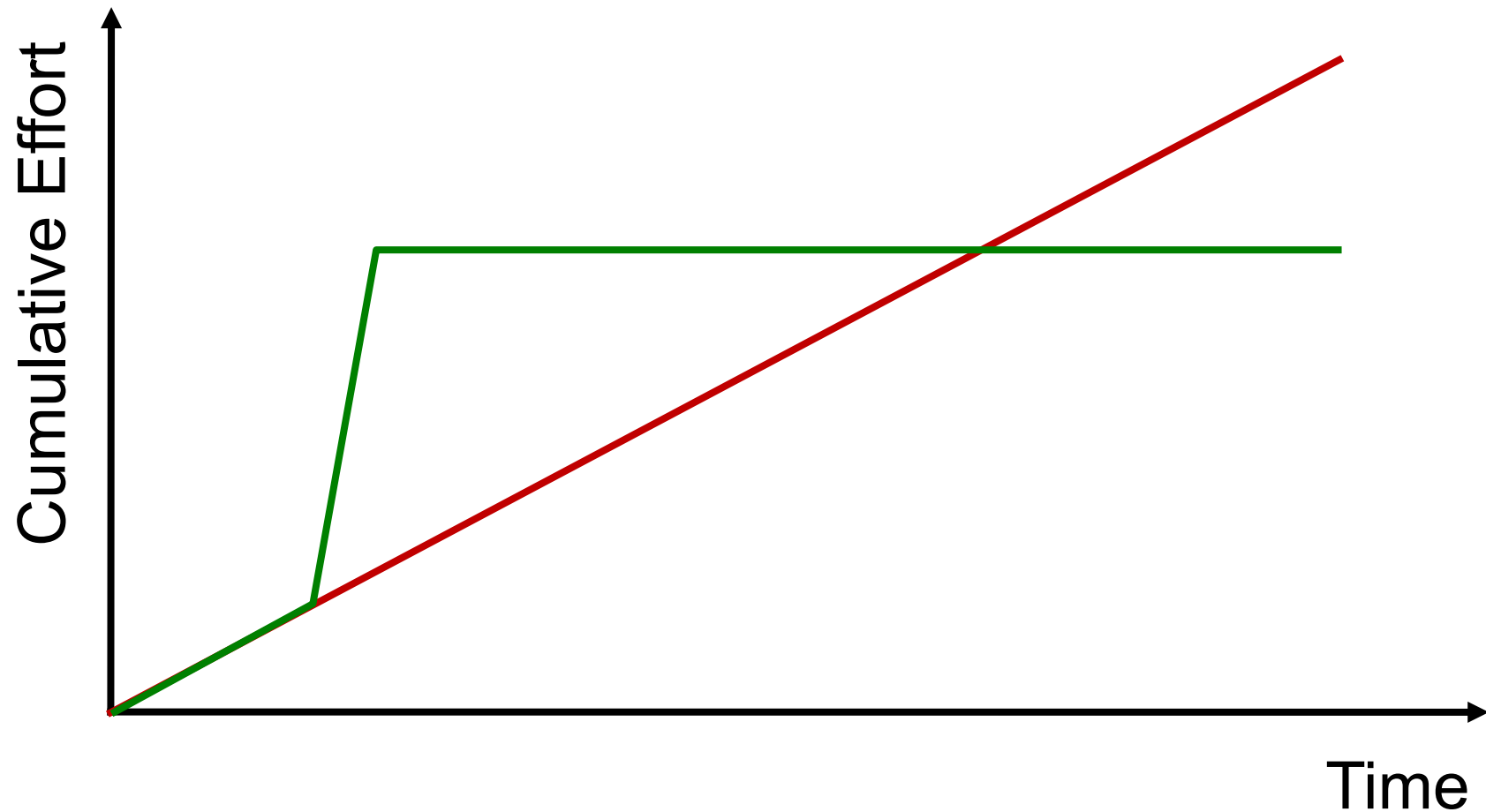
# Good Enough Practices

- **Data management** - saving both raw and intermediate forms, documenting all steps, creating tidy data amenable to analysis
- **Software** - writing, organizing, and sharing scripts and programs used in an analysis
- **Collaboration** - making it easy for existing and new collaborators to understand and contribute to a project
- **Project organization** - organizing the digital artifacts of a project to ease discovery and understanding
- **Tracking changes** - recording how various components of your project change over time
- **Manuscripts** - writing manuscripts in a way that leaves an audit trail and minimizes manual merging of conflicts

Wilson et al. (2017) Good enough practices in scientific computing. PLoS Comput Biol 13(6): e1005510



# Thoughts on Productivity and Automation





# Thoughts on Productivity and Automation

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE?  
(ACROSS FIVE YEARS)

		HOW OFTEN YOU DO THE TASK					
		50/DAY	5/DAY	DAILY	WEEKLY	MONTHLY	YEARLY
HOW MUCH TIME YOU SHAVE OFF	1 SECOND	<div><div>1</div> DAY</div>	2 HOURS	30 MINUTES	4 MINUTES	1 MINUTE	5 SECONDS
	5 SECONDS	<div><div>5</div> DAYS</div>	12 HOURS	2 HOURS	21 MINUTES	5 MINUTES	25 SECONDS
	30 SECONDS	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>4 WEEKS</div>	<div><div>3</div> DAYS</div>	12 HOURS	2 HOURS	30 MINUTES	2 MINUTES
	1 MINUTE	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>8 WEEKS</div>	<div><div>6</div> DAYS</div>	<div><div>1</div> DAY</div>	4 HOURS	1 HOUR	5 MINUTES
	5 MINUTES	9 MONTHS	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>4 WEEKS</div>	<div><div>6</div> DAYS</div>	21 HOURS	5 HOURS	25 MINUTES
	30 MINUTES		6 MONTHS	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>5 WEEKS</div>	<div><div>5</div> DAYS</div>	<div><div>1</div> DAY</div>	2 HOURS
	1 HOUR		10 MONTHS	2 MONTHS	<div><div>10</div> DAYS</div>	<div><div>2</div> DAYS</div>	5 HOURS
	6 HOURS				2 MONTHS	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>2 WEEKS</div>	<div><div>1</div> DAY</div>
	<div><div>1</div> DAY</div>					<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> <div>8 WEEKS</div>	<div><div>5</div> DAYS</div>



**Karen Cranston**

@kcranstn

 Follow

@mtholder motivating git: You mostly collaborate with yourself, and me-from-two-months-ago never responds to email. @swcarpentry

RETWEETS

25

LIKES

15



7:23 AM - 23 Aug 2013



25



15



[http://bit.ly/motivate\\_git](http://bit.ly/motivate_git)





# Make Incremental Changes Redux

- Choose one practice
  - Implement it in your work
  - Share it with your lab group
  - Allow it to sink in
- Repeat



# Where to Start?

- Depends on the nature of your work
- Consider the purpose:
  - Improve productivity
  - Improve quality



# Workshop Logistics: Continuous Feedback

## Post-It Usage #1

green sticky up – exercise completed

red sticky up – exercise in progress; HELP!

nametag down – all is okay

## Post-It Usage #2

At each break, indicate something that was good and something that could be better.

## Post-Workshop Survey

You'll receive a link from us after the workshop.



# Welcome to the Workshop!

## **Workshop Website**

<http://sarahlrstevens.info/2017-08-02-chicago-frb/>

## **Etherpad (for collaborative notes)**

<http://pad.software-carpentry.org/2017-08-02-chicago-frb>

## **Pre-Workshop Survey**

[https://www.surveymonkey.com/r/swc\\_pre\\_workshop\\_v1](https://www.surveymonkey.com/r/swc_pre_workshop_v1)



# Why Learn the Shell?

- Installing software
- Configuring your computer
- Controlling remote machines
- Running scripts from others
- Writing your own scripts/pipelines