Iabibi Documentation

Release 1.0

C. Titus Brown

Nov 03, 2019

Contents

This is the overview Web site for the Data Intensive Biology training program at UC Davis, run by Dr. C. Titus Brown. This is a series of workshops primarily focused on sequence analysis, using a short-course format taken from the Next-Gen Sequence Analysis Workshops and inspired by Software Carpentry and Data Carpentry.

- All workshops and community events are on UC Davis campus unless otherwise specified.

- Workshops are open to everyone whether or not they are part of the UC Davis community, subject to space constraints.

- Workshops are free unless otherwise specified.

If you want to be notified of new workshops or events, you can sign up for the training mailing list; workshops are generally announced there first.

You can subscribe to our Google Calendar containing all of our events here.

Or, you can simply view the calendar here.

Meet and Analyze Data (MAD)

1.1 In brief:

MAD is a weekly meetup of data scientists and scientists with data at UC Davis, usually in the Bennett Conference room on the Veterinary Medicine campus from 3-5 pm on Wednesdays (updated Fall quarter 2019). This year we are asking all participants to share their questions, ideally in advance, on our Discourse forum.

The schedule (Fall 2019):

Date	Topic
2-Oct	MAD 2.0: how to help us help you **Location: Valley Hall 1043**
9-Oct	What is the shell?
16-Oct	Practical software management with conda, Pt 1
23-Oct	no MAD
30-Oct	Practical software management with conda, Pt 2
6-Nov	Project organization and management
13-Nov	How to backup data on a remote computer
20-Nov	Quick sequencing data insights with sourmash
27-Nov	Thanksgiving – no MAD
4-Dec	Automation with Snakemake, Pt 1
11-Dec	Automation with Snakemake, Pt 2

1.2 The whole story:

MAD is intended to function as a community of practice. Everyone has a role to play to support this community for the benefit of all participants. Don't worry, though – novice or expert, bold or shy, you can both benefit from and contribute to this community!

The goal of this program is to share and distribute skills. We help researchers with solving computational problems, and in this context we ask participants to actively learn and develop computational skills. As those skills advance,

researchers are in a better position to share those skills with newer community-members. This is not only fun and rewarding: it keeps the community going and, more broadly, it helps to resolve the scarcity of expertise in computational biology.

Help us help you by actively participating in this community! Participation is fun and open to everyone. To keep it fun, please be sure to check out our DIB-Lab Code of Conduct.

How to get the most out of your visit: There are many ways to benefit from participating in this community. Asking a question is only one of them!

- Ask a question on Discourse. It can be hard to formulate a question sometimes, and it's ok if you can't! Trying to formulate your problem as a question will help you learn, and will also help you find the right person to assist you. Please post these before MAD if you can! Answers might appear outside of MAD hours, but coming to MAD is the best way to get an answer.

- Try things out on your own. Once your first question has been answered, try moving forward with some next steps while you are here. The person assisting you may move on, but they can always come back if you get stuck! Working independently in a supported environment is a great way to learn.

- Listen to other questions. It might be relevant to your work! Feel free to join in on other conversations in the room.

Talk to your neighbors about their computing problems. Not getting help at the moment? Explaining your work and listening to other problems will advance your understanding more broadly. It's also more fun. * See if you can help! Helping solve other people's problems is a great way to learn. How do you think our experts got to be so... expert? It's common to feel underqualified for this role. We're here to support your attempts!

1.3 How to pay it forward:

This community gets a lot of questions! How can we sustainably answer them all?

- Support our documentation. If you got a question answered, please post the answer to Discourse. This will both secure your understanding and support others who have the same question. Not sure you got the explanation right? Don't worry – our mods will catch it (and you might learn more in that process!)

- Respect helper time limits. The volunteers who are most actively engaged in helping roles often get a lot of questions in a day. Working independently after getting a preliminary answer is not only a great way to learn, it really helps to let these folks move on to help others for a while. If a helper has been working with you for 10 minutes or more, consider whether there is anything you can try on your own and have them check back later.

- Come! Even if you don't have a question. Attending group lessons, doing your work here, making yourself available to chat about problems (even if you can't help), getting to know othe people in the community – all of these things support the community.

1.4 How to skip the guilt:

Yes, we want you to come back and give back! But academics have enough to feel guilty about. Some days, you just have to come for the help to get you through, and other days, it just isn't happening. We get it. Come when you can, enjoy the snacks and the company, and do what you can when you can to pay it forward! Then, feel good about the work that you do to keep this community afloat.

1.5 If you need a lot of help:

Another thing we get is that some projects are really complicated and require extensive expert support. If you need a lot of help from one person on a project, or if it's turned out that way despite your best efforts, it is appropriate to add your helper as a co-author on publications that result from that work.

1.6 Come for the lessons, stay for the snacks, get help along the way!

We look forward to seeing you at MAD!

YouTube Videos from Streamed Workshops

- October 20, 2016 - Open Refine for working with large spreadsheets - half day workshop

- October 6, 2016 - UNIX Shell - half-day workshop

- May 18, 2016 - Vega-Lite for visualization (Dominik Moritz and Ham Wongsuphasawat) - half day workshop (broadcast)

- May 11, 2016 - RMarkdown (Marian Schmidt) - half day workshop (broadcast)

- April 27, 2016 - Differential Gene Expression in R (Meeta Mistry)

- April 13, 2016 - Intermediate git (Daniel Chen)

- April 6, 2016 - Introductory git (Daniel Chen)

- April 4, 2016 - Advanced Beginner Python (Titus Brown) Part One and Part Two

- March 28, 2016 - Short-read trimming and quality evaluation - half day workshop (Titus Brown)

- March 7, 2016 - Amazon Web Services (Titus Brown)

- February 29, 2016 - Sphinx and Bitbucket (Adelaide Rhodes)

- February 19, 2016 - scipy.optimize (Ariel Rokem)

- February 17, 2016 - Regular Expressions & Python (Tiffany Timbers)

- January 27, 2016 - Advanced git (Raniere Silva)

- January 20, 2016 - pydoit for workflow automation (Camille Scott)

- December 4, 2015 - d3.js for Interactive Data Visualization (Emily Dolson)

Workshop materials

All workshop materials are available under an open license, and they are available indefinitely. Please feel free to send them on to others.

- mRNAseq on non-model systems (last updated: May 2015)
- mRNAseq assembly with a reference genome (last updated: March 2015)

Past Workshops

- December 15th, 2016 - R, RStudio, and RMarkdown for exploring your data.

- December 8th, 2016 - Introductory version control with git.

- November 29th, 2016 - Introduction to UNIX shell

- November 17th, 2016 - Amazon Web Services for cloud computing.

- October 20th, 2016 - Open Refine for working with large spreadsheets.

- October 6, 2016 - UNIX Shell - half-day workshop

- May 18, 2016 - Vega-Lite for visualization (Dominik Moritz and Ham Wongsuphasawat) - half day workshop (broadcast)

- May 11, 2016 - RMarkdown (Marian Schmidt) - half day workshop (broadcast)

- April 27, 2016 - Differential Gene Expression in R (Meeta Mistry) - half day workshop (broadcast)

- April 13, 2016 - Intermediate git (Daniel Chen) - half day workshop (broadcast)

- April 6, 2016 - Introductory git (Daniel Chen) - half day workshop (broadcast)

- April 4, 2016 - Python, advanced beginner - half day workshop (broadcast)

- March 30, 2016 - De novo RNAseq Assembly - half day workshop

- March 28, 2016 - Short-read trimming and quality evaluation - half day workshop (broadcast)

- March 9, 2016 - Jupyter Notebook - half day workshop

- March 7, 2016 - Amazon Web Services - half day workshop (broadcast)

- February 29, 2016 - Sphinx, Webhooks, and Bitbucket (Adelaide Rhodes) - half day workshop (broadcast)

- February 19, 2016 - scipy.optimize (Ariel Rokem) - half day workshop (broadcast)

- February 17, 2016 - Regular expressions and Python (Tiffany Timbers) - half day workshop (broadcast)

- February 5, 2016 - Introductory git - half day workshop

- January 27, 2016 - Advanced git (Raniere Silva) - half day workshop (broadcast).

- January 20, 2016 - pydoit for workflow automation (Camille Scott) - half day workshop (broadcast).
- January 13, 2016 - Advanced beginner UNIX shell - half day workshop.
- December 7-8, 2015: UCSC Genome Browser workshop
- December 4, 2015: d3.js for Interactive Data Visualization
- December 3, 2015: Half Day Shell Workshop
- December 3, 2015: Half Day Python Workshop
- November 9-10, 2015: Docker hands-on
- November 2-3, 2015: mRNA Sequencing Intermediate workshop (Invitation Only)
- September 24-25, 2015: Microbial genomics and transcriptomics hands-on.
- September 21-23, 2015: iPlant Collaborative multiworkshop - introductory, intermediate, advanced, and Data Carpentry.
- August 17-18, Software Carpentry for Women in Science and Engineering.
- July 6-7, 2015: Software Carpentry
- May 4-5, 2015: 2015-05-04-mRNAseq-nonmodel (mRNAseq, de novo, intro)
- April 21, 2015:2015-04-21-mothur (intro, microbial)
- March 4-5, 2015: 2015-03-04-mRNAseq-semimodel (mRNAseq, ab initio, intro)
- February 9-10, 2015: Data Carpentry workshop (data carpentry, intro)
- January 6-8, 2015: Software and Data Carpentry Instructor Training (instructor, t3)
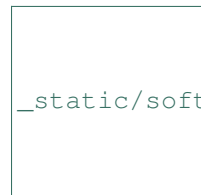
Partners and Affiliates

UC Davis Data Sciences Initiative - we are part of the fledgling Data Sciences Initiative at UC Davis.

—

UC Davis Genome Center - Dr. Brown is a member of the Genome Center at UC Davis. Note that the Genome Center hosts a Bioinformatics Training Program!

—

_static/software-carpentry-banner.png

We are an Affiliate Member of the Software Carpentry Foundation, a non-profit organization whose members teach researchers basic software skills.

—

_static/DC1_logo_small.png

We are friends of Data Carpentry, an organization designed to teach basic concepts, skills and tools for working more effectively with data.

CHAPTER 6

Contact information

You can contact Titus Brown at ctbrown@ucdavis.edu.