

User Training

Fall 2019 Registrations Open!

Notice: In addition to our training programs, there are many other computational groups at HMS that provide training, including at least:

- * <http://bioinformatics.sph.harvard.edu/training/>
- * <https://datamanagement.hms.harvard.edu/news-events>
- * <https://sbgrid.org/news/sbgrid-webinar-series>
- * <https://www.biogrids.org/news/biogrids-workshop-series>

Please directly contact them if you are interested to attend their trainings.

You must have a Harvard University ID (HUID) to be able to register for a class. This is required to access the Harvard Training Portal.

Those members of the HMS community who do not have HUIDs - such as employees at affiliate hospitals, or collaborators from other institutions - may self-register for one as a "Person of Interest" with their faculty member's sponsorship. The form may take several days to process, so we encourage non-HUID users to fill out this form ASAP in preparation for upcoming training:

<http://iam.harvard.edu/files/iam/files/authorized-identity-request-form.pdf>
(please read through the form for submission details)

Additional Matlab Trainings Offered by MathWorks

| Class | Date | Time | Room | Seats | Page | Registration |
|--|------------|------|----------|-------|------|---|
| Hands-On Workshop: Data Analysis and Visualization with MATLAB | 10/17/2019 | 1-3p | TMEC 328 | 48 | | Please email: Alyssa Silverman <asilverm@mathworks.com> |
| Deep Learning for Signal Data in MATLAB | 10/28/2019 | 1-3p | TMEC 306 | 48 | | Please email: Alyssa Silverman <asilverm@mathworks.com> |

Current Schedule

| Class | Date | Time | Room | Seats | Page | Registration |
|---------------------|------------|------|-------------------------|-------|--------------------------------------|----------------------|
| Intro to O2 | 09/04/2019 | 3-5p | Countway 506 Minot Room | 30 | User Training Github | link |
| Intro to Matlab | 09/11/2019 | 2-4p | Countway 506 Minot Room | 30 | User Training Github | link |
| Intro to Python | 09/18/2019 | 3-5p | TMEC 106 | 48 | User Training Github | link |
| Intro to Perl | 09/25/2019 | 3-5p | Countway 403 | 32 | User Training Github | link |
| Intro to R | 10/02/2019 | 3-5p | Countway 403 | 32 | R Class Files | link |
| Intro to O2 | 10/09/2019 | 3-5p | Countway 403 | 32 | User Training Github | link |
| Intermediate O2 | 10/16/2019 | 3-5p | Countway 403 | 32 | User Training Github | link |
| Intro to Git/Github | 10/30/2019 | 3-5p | TMEC 128 | 48 | User Training Github | link |

| | | | | | | |
|-----------------------------|------------|------|-------------------------|----|-------------------------|------|
| Intro to Parallel Computing | 11/06/2019 | 2-4p | Countway 506 Minot Room | 30 | User Training Github | link |
| Intro to Python | 11/13/2019 | 3-5p | Countway 403 | 32 | User Training Github | link |
| Intro to O2 | 12/04/2019 | 3-5p | Countway 403 | 32 | User Training Github | link |

Additional Matlab Trainings Offered by MathWorks

October 17th

Hands-On Workshop: Data Analysis and Visualization with MATLAB

MATLAB helps you obtain deeper scientific and clinical insights by giving you the ability to analyze, visualize and parallelize ever increasing amounts of data. Whether you're analyzing data, developing algorithms, or creating models, MATLAB is designed for the way you think and the work you do.

In this hands-on workshop, we help you get started quickly with the basics of MATLAB through a series of practice exercises. We introduce you to the MATLAB desktop environment and help you quickly gain necessary skills and experience in MATLAB. Our hands-on approach allows you to practice, apply, and evaluate your knowledge in a classroom setting guided by MathWorks engineers.

Learning objectives:

- Enter commands, perform calculations and create variables
- Import data into MATLAB from external files using the import tool
- Use MATLAB documentation to find help and discover built-in functions
- Visualize data using MATLAB's plotting functions and customize the visualizations
- Use logical expressions to extract elements of interest from MATLAB arrays
- Bring together concepts learned and test your knowledge with a final project

October 28th

Deep Learning for Signal Data in MATLAB

Deep learning is a type of machine learning that can achieve state-of-the-art accuracy in many humanlike tasks such as voice recognition and time series forecasting.

The main tasks in a deep learning workflow include assembling large data sets, creating a neural network, to train, visualizing and evaluating different models, and using specialized hardware - often requiring unique programming knowledge. These tasks are frequently even more challenging because of the complex theory behind them.

In this seminar, we'll demonstrate new MATLAB features that simplify these tasks and eliminate the low-level programming. In doing so, we'll decipher practical knowledge of the domain of deep learning. Specifically, we'll show how to build and train neural networks that perform classification and regression on signal and time-series data.

Learning Objectives:

- • Manage large sets of data
- • Create and analyze deep learning networks
- • Build networks from scratch with a drag-and-drop interface

- • Perform classification and regression tasks on signal and time-series data
- • Tweak pre-trained models to apply them to our specific data
- • Speed up network training with parallel computing
- • Deploy models by generating code or interfacing with other programming languages

Classes offered:

Intro to O2

O2 for New Users addresses the needs of users who have very little linux experience, and are just getting started with HPC. More time will be devoted to covering linux basics, and the concepts of schedulers and jobs, and data management best practices. The lecture portion of this class is one hour, the second hour will be spent clinic-style with HMS RC staff to address workflow-specific questions and help convert commands to O2 SLURM syntax.

Intermediate O2

If you have taken HMS RC's Intro to O2 class or have equivalent experience, Intermediate O2 will help you leverage Linux and Slurm more effectively on the O2 cluster. This is a hands-on class with topics that include: handling Linux command output, using "for" loops in the bash shell, customizing your O2 account environment, a more in-depth look at Slurm commands for managing cluster jobs, and scheduling automated jobs with cron.

Intro to Python

Python is a popular scripting language for scientific computing and available across all computer platforms. The course will introduce you to some of the basics of the Python language as well as some of the nuances involved with its use specific to the O2 environment. The goal is to provide users with a foundational level of familiarity. Topics covered include basic data types and declaration, flow control (if/else), loops, a brief introduction to constructing a script, and a briefer introduction to modules. The course will be taught on O2, but general concepts are easily translatable to desktop and local installations.

Intro to R

Intro to using R and Bioconductor. R is a powerful, open-source, highly adaptable statistical language useful for crunching numbers to datasets like those produced by next-gen sequencing. This class covers R basics and learning to think like/understand R. Users will learn how to set up personal R libraries on O2, and use O2 R for its high memory allocations and parallelization. Topics include how to install packages, learn about variables, data types. data manipulation, flow control, and functions, perform simple statistical tests, and create a variety of plots. Laptops are encouraged.

Intro to MATLAB

Matlab has become the "language of science" in the past few decades. It is simple to use, yet powerful enough to be productive on large computing infrastructures. If you need: 1) Fast prototyping of research ideas; or 2) avoid spending too much time in coding instead of doing real science by taking advantage of Matlab's built-in functions; 3) User friendly graphical interface and educational documentation; 4) Simplicity of code; 5) Easy access to GPU computing power; 6) Easy plotting and presentation of data; you will find this introduction course useful. This course will introduce the basics of the MATLAB coding language with O2-scalability and data presentation.

Intro to Parallel Computing

This is a short introduction to Parallel Computing that will include an overview of the basic concepts of parallel programming: from running your job in an embarrassingly parallel way to writing simple shared and distributed memory parallelization codes in different languages. The seminar will cover several examples of actual parallel codes however it will not have any "hands on" components. A basic programming experience (of any language, no parallelization) is preferred in order to better follow the topics presented during the seminar.

Intro to Perl

Perl is an open source programming language that's flexible, available for nearly all platforms, and is easy to learn. It is well suited for data munging, and processing biological data. Topics covered include variables (scalars, arrays and hashes), numerical and string functions, loops and conditions, regular expressions, and reading and writing files. A brief introduction will be given for subroutines and BioPerl. The class will be focused on running Perl on the O2 cluster, though the fundamentals are applicable for using Perl installations elsewhere.

Intro to Git and GitHub

This course introduces Git and GitHub and covers topics including: Getting Started with Git for version control, Using GitHub Desktop effectively, Collaborating with others on GitHub, and Utilizing GitHub Flow for better workflow. No previous exposure is assumed. We hope attendees will leave the class with the knowledge and tools necessary to start integrating Git into their workflows and excited to begin collaborating on GitHub.