

## Course Description and XSEDE Justification

We'll be teaching a tutorial-style day-long session in neural data analysis in Python called **Data-driven Neuroimaging**. This short course has been developed jointly by the University of California's [Berkeley Institute for Data Science \(BIDS\)](#) and [Helen Wills Neuroscience Institute \(HWNI\)](#), as well as the [University of Washington \(UW\) eScience Institute](#). The course itself will be hosted at the University of California, San Francisco (UCSF) campus and will serve a broad range of faculty, postdocs, and graduate students from UC Berkeley and UC San Francisco. The course schedule is at: <http://grinberglab.ucsf.edu/python>, and course materials are available at: [https://github.com/choldgraf/UCSF-Data\\_Driven\\_Neuro](https://github.com/choldgraf/UCSF-Data_Driven_Neuro).

There will be ~40 participants attending and 3 instructors leading the sessions, along with several guest lecturers covering computational applications in neuroscience. A major objective of the course is to allow attendees to get hands-on with data analysis tools. Logging into a jupyterhub account that contains all the installed software and data will enable this objective.

For this we'll need:

- 50 Virtual Machines running simultaneously (40 students + 5 instructors + test/spare/debug VMs)
- Each VM will need to be a: [Jetstream m1.medium VMs](#) (6 vCPUs, 16GB RAM, 60GB Storage)
- Each VM will need an external IP address so students can connect remotely with a web browser to a Jupyter Notebook running on the machine
- We are requesting 10,000SUs in total which will cover:
  - 50 VMs (m1.medium) x 10 hours (full day of the course) =~ 3,000SUs
  - allow the instructors to test out materials ahead of the course itself
  - allow the ability to spin up some larger instances in case students want to experiment beyond the course material basics
  - allow the students to experiment on their own for a short time after the day of the course itself
  - No extra storage needed (the 60GB of instance storage itself is sufficient)
- The Campus Champion has helped determine this as the ideal configuration considering the requirements of our course materials:
  - <5GB per student to load datasets and perform analysis
  - 2-6GB of RAM per student
  - Access via a public facing IP address to interact with Jupyter notebooks securely.
- The ability for course instructors to access the services beforehand (and really as soon as possible) so we can develop for the course.

An existing proof-of-concept has been demonstrated via a Campus Champion allocation, and ongoing support for integrating our course software requirements with XSEDE CyberInfrastructure cloud resources is provided through Aaron Culich, our local Campus Champion and Consulting resource through Berkeley Research Computing (BRC).

Following this course, the Berkeley Campus Champion will work with CCs at UCSF and UW to help any researchers acquire their own Jetstream Startup and/or Research allocations if they wish to continue using XSEDE resources following the course.