Data Analysis with Python

Eric Firing
Spring, 2020, Tu-Th 9-10:15
OCN 760-4, CRN 86201
https://currents.soest.hawaii.edu/ocn_data_analysis/

You have a data set—it might be a time series of ocean current or wind measurements, or a sequence of maps of sea-surface temperature, or a series of profiles of ocean nutrients, or gridded output from a numerical model. How do you extract useful information from this data set—that is, use it to address questions of scientific interest? The goal of this course is to help you build the tools and the understanding that you will need for this purpose. We will use the Python language for all stages, from reading data in various formats, to quick plotting, the application of standard time series analysis methods (correlation, filtering, spectral analysis, EOFs, complex demodulation), and ending with a document containing code, explanatory text, and publication-ready plots. Statistical concepts and data analysis methods will be explored using numerical experimentation to complement their conventional mathematical expressions. We will work with data sets from the web chosen based on class interests, and their suitability for particular analysis methods. Each student will also work with an individual data set, typically one relevant to the student's thesis.

Although the course is centered on data and methods common in physical oceanography and meteorology, students from other disciplines are welcome. You will need a laptop computer running OS X or Linux, either directly or in a virtual machine. We will use Python and its libraries installed using miniconda; assistance in setup, including installing a virtual machine (via VirtualBox), can be provided. If you are interested but uncertain as to whether the course would be suitable for you, peruse the course web site and contact efiring@hawaii.edu.