Tutorial adaptation overview:

Biology majors intuitively understand that the assembly and analysis of very large datasets can support exploring and solidifying ecological concepts. Students also value opportunities to conduct their own investigations but are seldom given occasion to combine personally collected data with much larger datasets assembled by other researchers. This NEON tutorial adaptation provides a framework for familiarizing upper-level, Biology students with large, standardized datasets focused on easily observed phenomena that students can replicated on a more limited scale by collect and evaluating their own data and comparing results.

In this adaptation of the three-part *Work with NEON OS & IS Data – Plant Phenology & Temperature* tutorial, students are first exposed to R, RStudio, and work through a freely available, DataCamp course that familiarizes them with some basic aspects of R. During a second session, students then complete the NEON tutorial, with some very basic modifications, as described further in this Faculty Notes document. Each session makes use of a three-hour laboratory meeting, and a to-be-developed extension that may be used to provide students with their own investigative, field experience.

Original materials:

This adaptation makes use of the following resources:

DataCamp. n.d. *Introduction to R*. Freely available from <https://www.datacamp.com/home>

DataCamp offers for-free and subscription-based data sciences courses in R, Python, and SQL that facilitate learners’ ability to structure and visualize data. [Introduction to R](https://campus.datacamp.com/courses/free-introduction-to-r/) acquaints students with the basics of R syntax and functions.

Jones, M.A., Stanish, L.F., Robinson, N., Jones, K.D., Flagg, K. 2018. [*Large Datasets in R – Plant Phenology & Temperature Data from NEON*](https://qubeshub.org/qubesresources/publications/515/1). [NEON – National Ecological Observatory Network](https://qubeshub.org/community/groups/neon), QUBES [doi:10.25334/Q4DQ3F](Jones,%20M.%20A.,%20Stanish,%20L.%20F.,%20Robinson,%20N.,%20Jones,%20K.%20D.,%20Flagg,%20C.%20(2018).%20Large%20Datasets%20in%20R%20-%20Plant%20Phenology%20&%20Temperature%20Data%20from%20NEON.%20NEON%20-%20National%20Ecological%20Observatory%20Network,%20QUBES.%20doi:10.25334/Q4DQ3F)

The National Ecological Research Network (NEON), funded by the National Science Foundation and operated by Battelle Memorial Institute, collects and provides continent-wide ecological data for investigators to evaluate and document ecosystems’ change. Readily accessed data from monitoring stations provide numerous opportunities for exploring biotic and abiotic interactions; data used in the [Work with NEON OS & IS Data – Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) tutorial are a subset data collected from [NEON’s Domain 2 field sites](https://www.neonscience.org/field-sites/field-sites-map).

Machlis, S. 2015. *Learn to use R Your Hands-on Guide*. Computerworld.com March 2018. Available as r4beginners\_ver3.pdf from <https://www.computerworld.com/article/2497143/business-intelligence/business-intelligence-beginner-s-guide-to-r-introduction.html>

*Computerworld* and Computerworld.com publish articles that inform readers on emerging IT technologies and software. The *r4beginners\_ver3* pdf provides readers a basic, working knowledge of R and how to implement it. Although the article’s commentary addresses using R in non-Biology applications, students easily identify with the utility message, and a short discussion can quickly make the point that any kind of properly formatted datasets can be investigated with R.

Datasets:

All datasets come directly from the [Work with NEON OS & IS Data – Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) tutorial, without modification. The data product, **NEON-Pheno-temp-timeseries**, downloads as a folder containing two subfolders, one labeled **pheno** that contains:

**NEON.DP1.10055.001\_readme** – a text file about the data product and its contents

**NEON.D02.SERC.DP1.10055.001.variables.20170727T182615Z.csv** – an Excel .csv file that provides field names, term descriptions, data types, and units for tree data

**phe\_perindividual.csv** – an Excel .csv file with geolocation and taxonomic identity for each plant observed

**phe\_perindividualperyear** – an Excel .csv file reporting once-per-year plant growth status data

**phe\_statusintensity** – an Excel .csv file with plant phenophase and intensity data

The second folder, labeled **temp**, includes:

**NEON.DP1.00002.001\_readme** – a text file detailing naming conventions for sensor data files so that the sensor location and associated sampling information can be determined

**NEON.DP1.00002.001\_variables.csv** – an Excel .csv file that provides field names, term descriptions, data types, and units for temperature data

**SAAT\_1min.csv** - an Excel .csv file containing one-minute mean single aspirated air temperature obtained at a specified sensor location

**SAAT\_30min. csv** - an Excel .csv file containing thirty- minute mean single aspirated air temperature obtained at a specified sensor location

Instructor Notes on Student Activities:

This section provides guidance on proceeding through version 1.0.0 of Gaining Familiarity with R and Work with NEON OS & IS Data – Plant Phenology & Temperature tutorial adaptation. Preliminary notes on extending the activities to possible field investigations are provided, with the expectation that more formal notes will be forthcoming in a revision of this adaptation.

Before the first laboratory meeting working with R, students are provided access to the r4beginners\_ver3.pdf and read the document. Students are also instructed to set up a [DataCamp](https://www.datacamp.com/home) account prior to coming to the first lab session.

Adaptation – laboratory 1st meeting (3 hour session):

1. Students install R, RStudio, and set their working directory on individual computers according to instructions given in r4beginners\_ver3.pdf. The instructor then indicates that further use of R, RStudio, and working directory files will occur in the second laboratory meeting.
2. Once students have completed the first task, each individual navigates to [DataCamp](https://www.datacamp.com/home), logs in to his/her account, selects the [Introduction to R](https://campus.datacamp.com/courses/free-introduction-to-r/) course and progresses through it. It is worth pointing out that the DataCamp course activities occur through a web portal rather than on the student’s computer, a situation that changes in week two of the tutorial adaptation.

Although each student sets up an individual account, I found that pairing students and encouraging them to talk through the course components bolsters collaborative problem solving and collegiality. Instructors may find it helpful to periodically comment on particular aspects of the DataCamp course as students progress. For example I pointed out the value of practicing using “#” to annotate code and provide reminders of what a particular objective was. What makes sense to comment on will likely vary depending on students’ experiences while working through the course.

Completing the DataCamp [Introduction to R](https://campus.datacamp.com/courses/free-introduction-to-r/) consumes the balance of a three hour lab session and, admittedly, leaves everyone ready to turn attentions elsewhere. I wrap up the first lab session by providing my students with a word or two of encouragement and giving a very brief preview of the [Work with NEON OS & IS Data – Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) as they depart.

Adaptation – laboratory 2nd meeting (3 hour session):

1. Students access their individual computers and navigate to the [Work with NEON OS & IS Data - Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) tutorial.
2. Progress through the three series of the tutorial follows instructions as provided, with minor adaptations noted directly below:

* After downloading the **NEON-Pheno-temp-timeseries** data folder, students can move it to the working directory set during the first lab session.
* Although the *Work with NEON’s Plant Phenology* series of the Work with NEON OS & IS Data – Plant Phenology & Temperature tutorial discusses using the [**neonUtilities Package**](https://www.neonscience.org/neonDataStackR) (formerly, the neonDataStackR package) to unzip and join individual files, the downloaded tutorial files are already configured for use. In other words, the phenology and temperature data files do not need the application of neonUtilities functions before continuing the *Work with NEON’s Plant Phenology* series of the complete tutorial.
* Students were encouraged to select one among the possible observatory sites (BLAN, SCBI, or SERC)in the tutoral dataset to focus on. Likewise, students could select one plant species among those represented in the phenology data to examine. We restricted the possible selection to tree species (JUNI or LITU), reasoning that the same species could be more reliably located in our region than some of the other taxa represented in the tutorial’s phenology data, thus enabling extending investigations to our local area. Students were also encouraged to select a phenophase of interest, but the instructor indicated that any field sampling for comparative purposes would obviously be based on the phenophase available during our class’ observation period.
* For the *Work with NEON’s Single-Aspirated Air Temperature Data* series, part two of the overall tutorial, students had already established the working directory for their data files and determined that the data were already configured for exploration. So, as with the first series of the tutorial, it was useful to know about the neonUtilities Package, but none of the functions were required to complete the second portion of the tutorial. Students were reminded to select the same observatory site to plot temperature data as was chosen for plotting phenology data. Prior to selecting the temperature values (**TempSingleMean**, **TempSingleMinimum**, or **TempSingleMaximum**) to plot, we briefly revisited some plant physiology basics to help select a temperature regime most associated with the phenophase of interest. For example **colored leaves**, an indication of leaf senescence progress, is partially influenced by the difference between daily high and low temperatures so selecting **TempSingleMean** makes better sense than **TempSingleMaximum**.
* The third series in the [Work with NEON OS & IS Data - Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) tutorial, *Plot Continuous & Discrete Data Together*, was followed as laid out in the original tutorial, with the exception that students had already downloaded the data, set a working directory, and had data files already configured for manipulation and visualization. This series in the tutorial provides good opportunity to review data presentation best practices with students. Although I did not develop a set of questions/prompts prior to assisting students with this series in the tutorial, an update of this adaptation will include some suggested, guiding questions.

1. Upon completing the entire tutorial, students exported their figures saved as a .svg file. Although there are multiple options for saved file formats, I indicated to my students that vector-based graphic images are preferred for most publications so establishing proper practices make sense.

* My students were able to complete the entire [Work with NEON OS & IS Data - Plant Phenology & Temperature](https://www.neonscience.org/osis-pheno-temp-series) tutorial in sufficient time that we were able to compare students’ final figures and very briefly consider a number of possible, testable hypotheses that could serve as the basis of local field investigations.

Adaptation – preliminary notes on local field-based extensions (additional laboratory sessions as needed):

1. After completing the [NEON.adaptation\_Large\_Datasets\_in\_R\_Plant\_Phenology\_&\_Temperature\_Data](https://qubeshub.org/community/projects/neonfmnsp2018/publications/621) sessions, students and their instructor consider what species and phenophase represented in regionally relevant NEON observatories’ posted data could be comparably explored in a local setting.

* Decisions on a taxon or taxa for local investigation should be matched as closely as possible based on sites comparable in topography, land use history, and current vegetation as possible to regional NEON observatories. (My students are fortunate in that multiple NEON sites, with good comparability exist.) Accessibility to comparable, ideally same-site, temperature data are required as well. Investigating suitable field sites and available temperature data may necessitate some out-of-lab time to evaluate and select for use.
* Based on ability to generate and evaluate comparable phenology and temperature data in a much smaller volume, students could be assigned a short write-up or longer, formal lab report.