Environmental Statistics and Computation
AATM315 (#7616) / AENV315 (#10561), Spring 2017

Credits: 4
Professor: Dr. Oliver ELISON TIMM (oelisontimm@albany.edu)
Class Location: ES232
Class Schedule: Tue/Thu/ 11:45am-1:05pm & Fri 9:20am-10:15am
Contact Information:
Office: ES 316A
Phone: 518-442-3584
E-mail: oelisontimm@albany.edu
Office hours: Mon 10:00am-11:00am, Wed 2:00pm-3:00pm, or by appointment
TA: Josh Alland (jalland@albany.edu, ES 325)
TA’s office hours: Tue/Thu 1:30pm-2:30pm

Course Description

General Goals

In this course you will work with environmental data and learn to apply basic concepts of statistical data analysis. You will use your computer/laptop and learn to solve statistical problems that are often found in atmospheric and environmental research. For example you will be working with computer programs and data examples that illustrate the following typical statistical problems: How can we summarize large data sets in visual graphs? Why is it important to have a large sample size when dealing with statistical data analysis methods? How can we detect climatic trends in global temperatures or local rainfall? When is a correlation significant? What is the idea behind statistical hypothesis tests? How do we make inferences from noisy and uncertain observational data? What signals can we detect in time series of atmospheric CO₂ concentrations? This course will cover standard concepts in probability theory, univariate and multivariate statistical analysis methods, statistical description of data, visualization of data and the concepts of hypothesis testing, and time series analysis.

Many research activities and jobs in the private sector operate with large volumes of data. In many cases, the analysis of big data sets can be streamlined by means of a few lines of computer code. In this course you will work with the general-purpose programming language Python. In this course you will practice the use of Python for your data analysis and visualization. The coding goes hand in hand with the statistical analysis techniques used during the class activities. After an introduction into the essential programming concepts (such as variables, loops, code-branching and functions) we will make use of packages numpy and pylab that allow us to handle our data analysis problems in an effective and illustrative programming format.

Computer and programming prerequisites

This course assumes no knowledge in programming. You will learn the basic principles in computer coding in Python (available for all common computer operating systems as free software). Bring your laptop computer to the classes! Please download the software and install it before the first class (Tue. Jan, 24th). Links and instructions can be found in the Blackboard course content.
**What students can expect**

The course is adopting the team-based learning (TBL) concept. We use the classroom for practicing the statistical and programming concepts through the application to real-life data examples. We will use the class-room to practice working in teams on specific tasks. You are expected to prepare for classes through reading assignments. The reading material will provide you with the background information and knowledge that you need for successfully solving the team-based activities. Note that reading about python programming is not effective unless you practice it on your own computer. Learning a programming language works best when you try things out and learn from mistakes!

Some class will be used to clarify encountered problems or to answer questions that came up during the team activities (this is done in lecture-centered format). This part will include short PPT presentations and use of the whiteboard. At the end of learning units, each student prepares as a homework assignment a written summary of the activities (with a discussion of the results and the statistical methods). Mid-term and final exam are based on mix of multiple choice questions and questions with free-format answers.

**The TBL procedure:**

The semester course is divided into several units. Before and during each unit, students are required to study the assigned readings. The readiness assurance progress tests the student’s individual preparation for the classroom activities (individual Readiness Assessment Test, iRAT).

The iRATs are immediately followed by team tests (tRATs), in which the teams discuss the same questions. As a team you have to form a consensus answer. The tRATs will be immediately evaluated in class and the team scores will be announced (individual scores will be handled discreetly). The RATs are scheduled in the beginning of each unit (or in middle of a unit).

**Team appeals:**

Teams are given the opportunity to submit a written appeal if they feel their answer was correct or if there were problems with the question. In any case, the written appeal (one-page limit) must be based on sound reasoning supported by references to the reading material or the calculations etc. Appeals must be submitted during or by the end of the class to the instructors, and the instructors will review and decide on the case either in class (or to the beginning of the next class).

Successful appeals will earn the associated points in the team’s tRAT score. Teams that did not appeal will not earn additional points in their tRAT score.

**Team activities:**

This class is about statistics and the interpretation of environmental data. The tasks will require a combination data analysis, and discussion of the underlying theoretical statistical principles. Participation in these group activities is a substantial part of the overall learning process, individually and through the discussion and communication with team members. From each unit one team activity report is used it for team’s grading.
Peer-evaluation process:

Since the team activity is part of the effective learning process, you will conduct two peer evaluation processes of your team members. The first is formative and not counted for the final grades; only the second peer evaluation at the end of the course will be considered in the final grade calculation. The peer review evaluates the attendance, preparedness, and constructive participation in the team work assignments.

Additional performance evaluations:

This course has a mid-term exam and a final exam that each student takes individually. Mid-term and final exam scores will contribute to the final grade.

Course Policies:

Attendance:

Your team needs you and you need your team! Therefore, attendance in class is crucially important. Attendance will be recorded, but not be graded. Missed RATs and team activities cannot be made up. Scores for missed iRAT and tRAT and team activities count as 0 points.

Safety-valve rule:

The lowest iRAT and lowest tRAT score will be dropped from the final grade calculations. The lowest team activity grade will be dropped (evaluated for each student individually).

Grading:

RATs (average of individual and team): 20%
Homework: 15%
Midterm: 25%
Final: 25%
Team Activities: ~10%
Peer Evaluation: ~5%

Note that about 75% of your grade is determined by your individual performance. Conversion from percentage points to grade is following the SUNY Albany grade conversion table.

Attachments: Course overview
AATM/AENV315 Spring 2017 course overview
Instructor: Oliver Elison Timm

The outlined timeline with its units and associated dates is a tentative timeline. RATs are scheduled on the first day of each learning unit, with a few exceptions (see below).

Learning Units:
(0) Introduction to the AATM/AENV315 course             01/24/17
  a. Syllabus
  b. Team-based learning approach
  c. Grading
  d. Team formation
  e. Setup of Python package Anaconda
  f. First steps using the software

(1) Introduction to Python                                  01/26/17-01/27/17
  a. Introduction to the IDE software “Spyder”
  b. Variables, values and data types
  c. Working with numbers: integer, float
  d. Working with strings
  e. Logical expressions

(2) Writing Python scripts                                  01/31/17-02/03/17
  RAT: 01/31/17
  a. Using functions
  b. Branching with if-then-else statements
  c. Iterations
  d. Tuples, lists and dictionaries
  e. Importing packages
  f. Defining and using your own functions
  g. Name spaces and scope rules

(3) Python for statistical data analysis                     02/07/17-02/10/17
  RAT: 02/07/17
  a. Package numpy and pylab
  b. Introduction to arrays
  c. Working with data arrays
  d. Plotting data with pylab
  e. Linear algebra with numpy arrays
(4) Basic probability concepts 02/14/17-02/24/17
   RAT: 02/14/17
   a. Randomness, events space, samples, population
   b. Independence, joint probability, conditional probability
   c. Descriptive statistical analysis (mean, median, standard deviation, histograms, box and whisker plots)

(5) Distributions 02/28/17-03/03/17
   RAT: 02/28/17
   a. Discrete distributions: Binomial Distribution, Poisson Distribution
   b. Continuous distributions: Uniform Distribution, Normal Distribution
   c. Central Limit Theorem
   d. Probability Density Functions, expected values

(6) Covariance and correlation 03/07/17-03/24/17
   Mid-term exam: 03/23/17
   a. Covariance between two random variables
   b. Pearson correlation coefficient
   c. Effects of sample size and outliers on correlation
   d. Linear regression
   e. Trend fitting and functional fitting

(7) Hypothesis testing 03/28/17-04/07/17
   RAT: 04/04/17
   a. Introducing T-test
   b. Rationale behind hypothesis testing
   c. Significance, Confidence level
   d. Error type 1 and Error type 2
   e. Monte Carlo testing

(8) Time series analysis 04/13/17-04/21/17
   RAT: 04/13/17
   a. Characteristics of time series: mean, variance, time stepping
   b. Autocorrelation and cross-correlation
   c. Low/High-pass filters
   d. Harmonic & spectral analysis
(9) Other statistical tools and python packages 04/25/17-05/05/17
   a. Spatial interpolation methods
   b. Multiple linear regression
   c. Principal component analysis
   d. Python packages scipy, pandas etc.

(10) Review session 05/09/17

Final exam* 05/18/17
8:00am-10:00am

(*www.albany.edu/registrar/registrar_assets/Spring_2017_Final_Examination_Schedule.pdf)

The final exam date and time is fixed and cannot be changed. Students are referred to the University’s Undergraduate Academic Regulations in the case of potential time conflicts.
(http://www.albany.edu/undergraduate_bulletin/regulations.html).