Syllabus for GEL160: Geological Data Analysis **Fall 2019**

Instructor: Max Rudolph maxrudolph@ucdavis.edu
Office: Earth and Physical Sciences 1133
Office Hours: Tuesdays 11-12, Fridays 3-4 or by appointment, in the iMac lab (2231)
TA: Gigja Hollyday gohollyday@ucdavis.edu
Office Hours: Wednesdays 10-12 in the iMac lab (2231)

Classroom: Earth and Physical Sciences 1316
Lab: On Wednesdays and Fridays, we may have class in the iMac lab, E&PS 2231
Lab system: https://hood.geology.ucdavis.edu/ (UC Davis sign in)
Canvas: All course communication and distribution of materials will be done using Canvas. ***Please use Canvas to send messages to the TA and instructor.*** Please make sure that your Canvas messages are forwarded to an email account that you check regularly.

Grading:

Homework	40%
Midterm exam	20%
Final project	30%
In-class Participation/Quizzes	10%

Reading: There is a textbook for the course, plus additional readings (to be distributed on Canvas) from other sources.

Guttag, J. Introduction to Computation and Programming Using Python, with Application to Understanding Data. *2nd edition*. MIT Press (available as e-book or hardcopy)
Note: 2nd edition uses Python3 – 1st does not.

Course overview:
This course is about understanding data in the geosciences. We encounter data in many forms. For instance, a time-series of acceleration measurements from a seismometer, a set of strike-and-dip measurements obtained in the field, chemical measurements such as the major- and trace-element composition of a basalt, an aerial photograph of a landscape, a deep-time record of the variations of isotopes in carbonate rocks. There is no one-size fits all solution to analyzing and understanding these disparate types of data. We will focus on a few key, fundamental approaches to understanding data, including spatial data, time series data, and large datasets of, e.g. rock composition. We will explore 'big data' problems and touch upon some applications of machine learning in the geosciences.

The course is organized into modules, and a detailed schedule follows. The modules are:
Module 0: Introduction, python 'crash course'
Module 1: Data statistics – mean, median, mode, curve fitting, parameter estimation
Module 2: Time series data
Module 3: 2D data (such as topography) & images
Module 4: Data mining & 'Big Data' – finding patterns in higher dimensional data
Principal Component Analysis, k-means clustering, machine learning

Academic conduct

The UC Davis Code of Academic Conduct applies to this course in the usual manner (see http://sja.ucdavis.edu/cac.htm). This course involves programming, and there is a temptation to use posted example code from the internet. If you copy and paste code, you must attribute the source (e.g. provide a url as a comment and clearly indicate which code is not yours). If you work with other student(s) on homework assignments, acknowledge your collaborators. The standard for 'intellectual ownership' of code in this course is that you should be able to explain to the instructor or TA line-by-line what your code does.

Note: This schedule is guaranteed to change. Changes to the schedule will be posted on Canvas.

Week		Dates	Topic	Readings
0	W	25-Sep	Introduction and overview	
	F	27-Sep	Introduction to Python and Jupyter (iMac Lab)	Guttag Ch. 1-2
1	M	30-Sep	Matrix and vector algebra	
	W	2-Oct	Basic numerical programs in python	Guttag Ch. 3
	F	4-Oct	Uncertainty and Error	
2	M	7-Oct	Probability distributions	Guttag Chapter 15
	W	9-Oct	Curve fitting I - linear least squares	Guttag Chapter 18
	F	11-Oct	Curve fitting lab	
3	M	14-Oct	Curve fitting II - nonlinear least squares	
	W	16-Oct	Parameter estimation	
	F	18-Oct	Monte-Carlo methods and model-space search	Guttag Chapter 16
4	M	21-Oct	Time series data I - Fourier series	Canvas
	W	23-Oct	The fourier transform and filtering	
	F	25-Oct	Signal processing in Python	
5	M	28-Oct	Phase lags and cross-correlation	
	W	30-Oct	Review for midterm	
	F	1-Nov	**Midterm Exam (in lab)**	
6	M	4-Nov	2D representations of data	Canvas
	W	6-Nov	Black and white (binary) images	
	F	8-Nov	Greyscale Images and DEMs	
7		No class Monday		
	W	13-Nov	Image segmentation	Canvas
	F	15-Nov	Quantitative information from images	
8	M	18-Nov	Quantitative information from images	
	W	20-Nov	Georeferenced images and map projections	
	F	22-Nov	Principal component analysis	
9	M	25-Nov	Machine Learning	Guttag 22
	W	27-Nov	k-means Clustering	
	F	No class Friday		
10	M	2-Dec	Hierarchical clustering	
	W	4-Dec		
	F	6-Dec	**Final Projects due**	
Finals	**Monday**	**8:00 AM**	**Final Project presentations**	