CS 5630: PYTHON FOR COMPUTATIONAL AND DATA SCIENCES

Semester Hours: 3.0   *Contact Hours:* 3

Coordinator: Ray Kresman

Text: Various

Author(s):   VARIOUS

Year: Various

SPECIFIC COURSE INFORMATION

Catalog Description:

Accelerated introduction to Python. Sample problems in STEM domains and computational approaches to solving them. Generic, and domain-specific libraries and tools. Introduction to data variety, analysis, and visualization. Prerequisite: Admission to MS in Computer Science, MS/Ph.D. in Data Science, or permission of instructor. Cannot earn credit for both CS 4630 and CS 5630.

Course type: **ELECTIVE**

SPECIFIC COURSE GOALS

- I can use language libraries to solve basic computational problems in STEM domain [examples: a) sequence alignment and use of STEM datasets; b) scripting in STEM applications; c) hypothesis testing and optimization].

- I can explain language mechanisms for handling missing data, and cite sample STEM applications where missing data is prevalent.

- I can use basic visualization and data classification on STEM datasets.

- I can explain certain data formats in STEM fields.

- I can use the primitives in certain libraries, for example: Numpy, Scipy, BiopythonSympy, Pyomo, Mathplotlib, Pandas.

- I can analyze relevant research and communicate my findings.

LIST OF TOPICS COVERED

- Accelerated introduction to Python (~ 15%)

- Datasets in the sciences (~ 10%)

- Data formats in STEM fields, examples: atmospheric science, biology
- Missing data - for example, radar measurements
- Data wrangling and analysis

- Applications - Math & Physics (~ 15%)
  - Matrix operations & ODE
  - Projectile motion and simple harmonic motion
  - Optimization

- Applications - Geology/Hydrology/GIS (~ 15%)
  - Raster & vector data
  - Line and contour plots
  - Basics of filtering and noise reduction
  - Process map layers and time series data

- Applications - Psychology and Statistics (~ 15%)
  - Descriptive and inferential statistics,
  - Models & hypothesis
  - Significance and hypothesis testing

- Applications - Chemistry/Biology/CS (~ 25%)
  - Chemical equations, stoichiometry
  - Bioinformatics and sequence alignment
  - Dynamic programming
  - Data and spatial visualization
  - Data science programming