Mathematical and Statistical Computing (MATH-510)

Syllabus for Fall 2015

The goal of this course is to provide students with sufficient programming background for graduate level study in mathematics and statistics. The course is focused on building basic skills using Python, R, Spark (distributed computing) and SAS, which are four tools used in professional workplaces. However, there are many other tools such as IBM SPSS, Statsoft Statistica, NETICA (for Bayesian Networks), WEKA (another open source suite of tools), RapidMiner and many others that a student might encounter in his/her professional experience. Students should leave the course with the ability and confidence to apply basic programming concepts across a variety of tools and programming languages.

Instructor: Gerhard Pilcher, Office 313.10 Pilcher (Cubicle in the Math Lounge Space)

Email: gp411@georgetown.edu

Office hours: Thursdays 5:00 – 6:15PM (on campus), and by appointment.

Room & Time: STM 343, Thursday 6:30 – 9:00PM.

Textbooks: None. Internet resources will be utilized during the course.

Software:
GitHub: https://github.com/Math510/Math510.github.io.git

Python: We will be using iPython distributed with Anaconda.
https://store.continuum.io/cshop/anaconda/

R and **RStudio** are freely distributed under the terms of the GNU General Public License. http://www.r-project.org/ (for latest version of R) and http://www.rstudio.com/ for the latest version of RStudio. RStudio is a handy gui wrapped around R that will save you lots of time. There is documentation all over the web for R. I will post a couple of documents on Blackboard.

Spark/Scala: <add links soon!>

SAS is available in all Georgetown University Information Services (UIS) computer labs. SAS software is free to students and compatible with both windows based PCs and Macs. I will provide the course notes.

Internet: The course will use Blackboard, http://campus.georgetown.edu/. Announcements, homework assignments and solutions, course material such as data sets, Python files, R scripts and SAS information will all be posted there. **You will also submit all homework assignments by 11:59pm on the due date.** You can look up your grades, and online surveys will also be conducted here. **I highly recommend that students in this course visit Blackboard once a day after 6pm.** Announcements will usually only appear on the web page.

Prerequisites: Linear algebra, multivariable calculus, familiarity with computers and regular access to a computer.

Topics to be covered:

- **Python:** Basic programming; working with matrices and arrays; scripts and functions; loops; debugging; graphics; toolboxes.
- **R:** Basic interface and workspace management; data types; working with arrays and data frames; importing and exporting data; probability distributions; simple statistical procedures in R; loops; scripts; graphics; downloading data from the web; using R in the cloud.
- **Distributed Computing:** Introduction to distributed computing concepts; working with Spark and Scala; exposure to Amazon Web Services (AWS) cloud computing
- **SAS:** Basic interface and workspace management; importing, manipulating, and exporting data; descriptive statistics, plots and SAS output; SAS macros; basic SAS procedures.

Grading:
Programming is a discipline that requires patience and practice. I encourage discovery of a solution on your own before working with your colleagues. Struggling with the problem for a while creates more permanent learning.

> **Class Work and Homework – 50% of your final grade.**
> Weekly homework assignments and class participation. For class participation, expect to be called up individually to demonstrate practice problems, explain concepts, and formulate questions concerning new material. Class participation will be taken into account when assigning grades (up to 10% of your grade)
>
> **Tests – 50% of your final grade. (see Honor Code Section below)**
> There will be two tests (15% each for a total of 30%) and a Final Exam (20%).

Late Submission:
Five (5) points will be deducted for every day or portion of a day that a weekly homework assignment is turned in late.

Grading scheme:

A	95+	A-	90-94	B+	85-89
B	80-84	B-	75-79	C	70-74
D	Below 70				

Rounding is at the discretion of the instructor.

Honor Code: Please be aware of the academic integrity rules for graduate students, see ch. VI of the Graduate Bulletin: http://grad.georgetown.edu/pages/bulletin.cfm

- **You may _not_ cooperate with other students or seek any human, except the professor, help on any of the tests or final exam.**
- **Computers will be used to take the tests and exam.**