Computational Methods for Analytics

Class Time: MR 12-1:15 P.M. *Class Room*: RLC 102

Instructor: Angel R. Pineda, Ph.D. *Office*: RLC 200A
Email: angel.pineda@manhattan.edu *Phone*: 718-862-7730
Web Page: http://turing.manhattan.edu/~apineda01/

Office Hours: Monday 2-2:50 P.M., Thursday 11-11:50 A.M. and 2-2:50 P.M., or by appointment.

Suggested Textbooks:
Amos Gilat, *MATLAB An Introduction with Applications*, John Wiley & Sons, Fifth Edition, 2015.
Peter Dalgaard, *Introductory Statistics with R*, Springer, Second Edition, 2008.
Joel Grus, *Data Science from Scratch: First Principles with Python*, O'Reilly, First Edition, 2015.

Free Online Resources:
Cleve Moler, *Numerical Computing with MATLAB*
http://www.mathworks.com/moler/chapters.html
Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, *Introduction to Statistical Learning with applications in R*
http://www-bcf.usc.edu/~gareth/ISL/

Course Description:
This course is a survey of commonly used tools in mathematical computing and data science. The emphasis will be in programming for applied mathematics and data science. For engineering applications, we will use MATLAB, for statistical problems we will use R and for more general data science problems we will use Python.

Course Pre-requisites:
Acceptance into the master's program in applied mathematics – data analytics or consent by instructor.

Course Objectives:
After completing this course, the students should be able to:
- Program basic programs in MATLAB, R and Python
- Understand the types of problems where each of these programs might be most appropriate
- Use computation for exploring problems in data science including descriptive and predictive analytics as well as data visualization

Course Homepage (Moodle):
Here you will find four features that will be used in this course:
- *Email:* make sure that your email on Moodle is one that you check regularly. Homework assignments, announcements and other class related information will be sent via email.
- *Course Information and Documents:* material covered each week, assignments and solution keys.
- *Student Discussion Board:* this online forum allows for students and faculty to communicate.
- *Grades:* students will be able to keep track of their grades online.

Grading:
Homework (20%)

Midterm Exams: (20 % each)

Exam I	Exam II	Exam III
Thursday September 22	Thursday October 20	Thursday November 17

Final Project (20 %), due Saturday Dec. 17, 11:00 A.M.-1:00 P.M.
In the course project, the student will write a paper and give a presentation on a topic of their choice with the instructor's approval. The project will use computation to solve a data science problem. Details for the course project will be given after Exam I.

Tentative Grading Scale

Percent	93-100	90-92	87-89	83-86	80-82	77-79	70-76	67-69	60-66	0-59
Grade	A	A-	B+	B	B-	C+	C	D+	D	F

The exact grading scale will be determined after the final exam. The numerical scores in the tentative grading scale guarantee the associated letter grade but the instructor may change the scale to the student's benefit.

Dates to Remember

September 2: Late Registration & Add/Drop Ends
September 5: Labor Day Holiday (No Classes)
October 10: Columbus Day Holiday (No Classes)
October 11: Monday Schedule
October 17: Midterm Grades Due
November 18: Last Day to Withdraw from Courses
November 23 – 25: Thanksgiving Holiday (No Classes)
December 9: Last Day of Classes

Class Policies

- Late homework will not be accepted after the solutions are distributed. In case the homework is handed in before the solutions are posted it will be marked 20% off for every day (or part thereof) it is late.
- The lowest HW grade will be dropped.
- No make-up exams will be given, unless you have a medical or family emergency. These emergencies require valid documentation. The grade for a missed exam is zero.
- Cell phones (or other technology not related to the class) in the classroom is only allowed with express permission of the instructor for special circumstances. In general cell phone or other potentially disruptive technology use is not allowed in class.

Suggestions

- The course requires a time commitment of about 9 hours outside of class time. Make sure to make enough space in your schedule to spend the time needed.
- I suggest you work in groups on your homework but hand in individual solutions, not copied from each other. Doing the homework is when most of the learning occurs. Make sure that you write your own computer code and that it is understandable to you.
- I encourage you to come to office hours regularly. I will do my best to help you.

Sources of Support

- The Center for Academic Success provides tutoring to support students with their courses. It also provides writing assistance. See their website for a list of their services and hours: https://manhattan.edu/academics/academic-support/center-academic-success.php

Academic Integrity:
Recall that as students of Manhattan College, you have each signed The Manhattan College Honor Pledge:

As a Manhattan College student, I will not lie, cheat, or steal in my academic endeavors, nor will I accept the actions of those who do. I will conduct myself responsibly and honorably in all my activities as a Manhattan College student. I am accountable to the Manhattan College community and dedicate myself to a life of honor.

Whenever you put your name on work to be handed in for grading in this class, you are reaffirming the above pledge. Violations of the Honor Code include, but are not limited to, cheating, plagiarism, fabrication, and other forms of academic misconduct. Students should familiarize themselves with the Manhattan College Student Code of Conduct and Academic Policies which can be found at:
http://www-archive.manhattan.edu/community-standards-and-student-code-conduct

Special Accommodations:

- Students with special needs should bring appropriate documentation to the Specialized Resource Center, Miguel 300, https://manhattan.edu/academics/academic-support/specialized-resource-center.php, to obtain an Academic Adjustment/Auxiliary Aid form. Bring the completed form to me as soon as possible, and together we will decide on how best to fulfill the adjustments and/or aids listed on the form.
- Student athletes should bring their event schedules to me as soon as possible.

Course Outline:

(MATLAB, Gilat Text)
MATLAB will be used to review the principles of programming and for applied mathematics.

Chapter 1 Starting With MATLAB
Chapter 2 Creating Arrays
Chapter 3 Mathematical Operations with Arrays
Chapter 4 Using Script Files and Managing Data
Chapter 5 Two-Dimensional Plots
Chapter 6 Programming in MATLAB
Chapter 7 User-Defined Functions and Function Files
Chapter 8 Polynomials, Curve Fitting, and Interpolation
Chapter 9 Applications in Numerical Analysis
Chapter 10 Three-Dimensional Plots

(R, Dalgaard Text using RStudio)
R syntax and structure will be presented and used primarily for statistical applications and data cleaning.

Chapter 1 Basics: First Steps and R Language Essentials
Chapter 2 R Environment: Session Management, Graphics, Programming, Data Entry
Chapter 10 Advanced Data Handling: Data frames
Using SQL in R

(PYTHON, Grus Text)
Python syntax will be and used for more general data science problems including Kaggle competitions.

Chapter 2 A Crash Course in Python: Programming, Packages, Visualizing Data, Data Structures, Objects
Chapter 14 Linear Regression in Python
Using Logistic Regression in Python for a Kaggle Competition

The material in this syllabus may be changed at the instructor's discretion. Any changes will be communicated to the students.