Syllabus for MAT 128/SOC 251: Foundations of Data Science
Spring 2018

Instructor: Prof. Megan Owen
E-mail: megan.owen@lehman.cuny.edu
Office hours: Tuesdays and Thursdays 12:00-12:50pm, Tuesdays 2:40-3:30pm, Gillet 137E
Course time: Tuesday and Thursday, 1:00-2:40pm, Gillet 231
Course website: http://comet.lehman.cuny.edu/owen/teaching/mat128/sp2018.html

Course Description MAT 128: *4 hours, 3 credit*s. Statistical and computational tools for analyzing data. Acquiring data from multiple sources, techniques for efficiently traversing, storing, and manipulating data. Emphasis on statistical analysis and visualization of real data. Prereq: Score of 65 or higher on College Math section of Accuplacer exam or department permission.

Textbook: This course will use parts of four free, online textbooks:
• *How to Think Like a Computer Scientist (Python)* by Jeffrey Elkner, Allen B. Downey, and Chris Meyers. Available at http://interactivepython.org/runestone/static/thinkcspy/index.html
• *Think Stats* by Allen B. Downey. Available at http://greenteapress.com/wp/think-stats-2e/
• *Online Statistics: An Interactive Multimedia Course of Study*. Project leader: David M. Lane. Available at http://onlinestatbook.com/2/index.html
• *R for Data Science* by Garrett Grolemund and Hadley Wickham. Available at http://r4ds.had.co.nz/

Grading: The grading for the course will be based on:

Assignments	30%
Classwork (in-class quizzes and group work)	35%
Final exam	35%

You must take and pass the final exam to pass the course.

Homework: All assignments should be submitted on Blackboard. **No late assignments are accepted.** To receive full credit for a program, the program must perform correctly, must include comments, and be written in good style. You can miss up to 5 programming assignments without affecting your grade (if you turn in all the programming assignments, we will drop the lowest 5 scores).

Classwork: At every class meeting, there will be an in-class quiz or group work based on the lecture notes, reading, submitted assignments, and laboratory exercises.
- As the semester progresses, quizzes will include review questions as well as short exercises based on the homework.
- All quizzes must be taken during class time, and **there are no make-up quizzes.** Instead, we will drop up to three missed quizzes (if you take all quizzes, we will drop the lowest three scores).

- The top 25 (out of 28 possible) quiz scores will count towards the quiz score in the final grade. If you do better on the final examination than on the top 25 quizzes, we will replace up to 5 quizzes with your final exam grade.
- All computer-based quizzes are open book, open notes, and you may use the Python shell and R programming environment
- Some quizzes will be on paper (in preparation for the final exam and many programming job interviews). For these, you may only use one 8"x11" sheet of paper with notes on both sides.

.

Final exam: The final exam is required and will be on Thursday May 24 from 1:30-3:30pm. You must pass the final exam to pass the course.

Use of Technology & Blackboard: Homework solutions and grades will be posted on Blackboard. The Blackboard system is provided by CUNY to all enrolled students. If you have not accessed Blackboard or are having difficulties, contact Blackboard Support in the Information Technology Division. You can also visit the Help Desk in the Computer Center (first floor, Carman Hall) in person. They can reset passwords and help with simple Blackboard issues.

Honor Code: You are encouraged to work together on solving the homework problems. However, you should write up the solutions on your own. You are responsible for knowing and following Lehman's academic integrity code (available from the Undergraduate Bulletin, Graduate Bulletin, or the Office of Academic Standards and Evaluations). All incidents of cheating will be reported to the Vice President of Student Affairs.

Accommodating Disabilities: Lehman College is committed to providing access to all programs and curricula to all students. Students with disabilities who may need classroom accommodations are encouraged to register with the Office of Student Disability Services. For more info, please contact the Office of Student Disability Services, Shuster Hall, Room 238, phone number, 718-960-8441.

Learning Objectives:
At the end of the course, students will be able to:
1. Interpret and draw appropriate inferences from quantitative representations, such as formulas, graphs, or tables.
   a. Graphs and tables will be used extensively to support inference.
2. Use algebraic, numerical, graphical, or statistical methods to draw accurate conclusions and solve mathematical problems.
   a. The emphasis is on inferring patterns and deducing properties using standard statistical techniques.
3. Represent quantitative problems expressed in natural language in a suitable mathematical format.
   a. The course focuses on translating quantitative problems about large data sets into suitable mathematical format that can be used to draw accurate conclusions (see

        #2).

4. Effectively communicate quantitative analysis or solutions to mathematical problems in written or oral form.
    a. In addition to written and oral communication, the course will also incorporate presenting information visually.
5. Evaluate solutions to problems for reasonableness using a variety of means, including informed estimation.
    a. Dealing with uncertainty creates natural informed estimation. The student will be encouraged to know when they are in the right ballpark.
6. Apply mathematical methods to problems in other fields of study.
    a. The underlying goal of this course is to give students the analytic reasoning skills and statistical tools to analyze data from other fields of study.