

**Syllabus for MAT 128/SOC 251: Foundations of Data Science  
Spring 2019**

**Instructor:** Prof. Megan Owen (she/her)

**E-mail:** [megan.owen@lehman.cuny.edu](mailto:megan.owen@lehman.cuny.edu)

**Office hours:** Tuesdays and Thursdays 2:40 - 4pm, Gillet 231 or 137E; or by appointment

**Course time:** Tuesday and Thursday, 1:00-2:40pm, Gillet 231

**Course website:** <http://comet.lehman.cuny.edu/owen/teaching/mat128/sp2019.html>

**Course Description MAT 128:** *4 hours, 3 credits.* 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.

**Course Goal:** Learn the fundamentals of data science by analyzing real world datasets with Python programming language.

**Textbooks:**

• *Computational and Inferential Thinking: The Foundations of Data Science* by Ani Adhikari and John DeNero. <https://www.inferentialthinking.com/chapters/intro.html>

Supplemental:

• *Online Statistics: An Interactive Multimedia Course of Study.* Project leader: David M. Lane. <http://onlinestatbook.com/2/index.html>

• *How to Think Like a Computer Scientist (Python)* by Jeffrey Elkner, Allen B. Downey, and Chris Meyers. <http://interactivepython.org/runestone/static/thinkcspy/index.html>

**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.** You can miss up to 5 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). We will also replace any classwork grade that is lower than your final exam grade, with your final exam grade.

- All computer-based quizzes are open book, open notes, and you may use the Jupyter notebook.
- Some quizzes will be on paper (in preparation for the final exam and many technical 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 16 from 1:30-3:30pm. You must pass the final exam to pass the course.

**Jupyter Hub:** Labs and assignments will be done in a Jupyter Notebook, which can be accessed through Jupyter Hub (under Lehman Apps on [Lehman One Access](#)). To access Jupyter Hub off campus (ie. at home), you will have to connect to the Cisco AnyConnect VPN first (instructions at: <http://www.lehman.cuny.edu/itr/reference-guides.php#cisco>).

**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 and plagiarism will receive a 0 on the test or assignment, and will be reported to the Office 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.