BSTA101: Population Health Data Science - II

About the Course

Instructor tom mcandrew
Email: mcandrew@lehigh.edu
Office Coordinates: STEPS 109
Office Hours: To be voted on by students | by Appt.

Class times Tuesday, Thursday 07:55 - 09:10 in Maginnes Hall, Room 110

Course Website I will update course website at http://thomasmcandrew.com/classes/2021F_PHDS2/public/ regularly with lecture notes and materials used in class. Homework assignments will be distributed on Course Site and the Midterm and Final Exam will be distributed in person.

Description Students will expand their statistics and machine learning toolkit by building traditional regression models for continuous and binary data, explore supervised learning methods such as: Tree-based learning, KNN/Collaborative filtering, and Feed forward Neural networks, and understand how to manipulate, ask, and answer questions from big datasets. May be taken in conjunction with BSTA 103.
We will plan to cover the following topics:

- Matrix Algebra
- Random variables and common distributions
- The Maximum Likelihood Approach and learning from data
- Simple Linear Regression
- Multiple Linear Regression
- Logistic Regression
- Poisson Regression
- Generalized Linear Models
- Quasi-likelihood
- Tree-based learning and KNN regression

Textbook We will mostly follow notes provided on the website but the following open source (free) materials for class may also be helpful:

- Introductory Statistics for the Life and Biomedical Sciences

A pdf of the book can be downloaded for free from the author's website at https://www.openintro.org/go/?id=biostat0&referrer=/book/biostat/index.php.

- Computational and Inferential Thinking

I will occasionally assign reading from Computational and Inferential Thinking. This is a free textbook available at https://www.inferentialthinking.com/chapters/intro.html

Time commitment I recommend budgeting approximately three out-of-class hours for every in-class hour to complete the reading, assignments, and homework. Twelve hours per week spent on class should be enough time to complete class requirements. If you spend more than 12 hours per week on a regular basis, I encourage you to check in with me.

Scheduling an appointment Students can schedule times for us to meet.I am happy to meet with a student one on one or as a group of 2-10 students. Appointments should be scheduled in advance and not last minute.

Policies

Attendance Your attendance in class is crucial. If you are sick or otherwise cannot attend class, please let me know and stay home and rest.

Collaboration Much of this course will operate on a collaborative basis, and you are expected and encouraged to work together with a partner or in small groups to study, complete homework assignments, and prepare for exams. However, every word that you write must be your own. Copying and pasting sentences, paragraphs from another student is not acceptable and will receive no credit or a penalty. No interaction with anyone but the instructor is allowed on any exams or quizzes. All students, staff, and faculty are bound by the Lehigh University Honor Code.

To sum up: On homeworks I want you to work together, but you must write up your answers yourself.. Dishonesty, plagiarism, etc., will be reported.

Technology

Computing We will use Python 3 throughout this course to illustrate statistical concepts applied to datasets. However, students are **not** required to code on assignments. Coding is reserved for BSTA103—Algorithms Lab II.

Assignments

Your grade for this course will be a weighted average of scores from several components:

Item	Weight
Participation	7%
Homework	50%
Midterm	22.5%
Final	22.5%

Homework The best way to learn statistics is to do it. There is an expected seven homework assignments that will be assigned throughout the semester. Homework assignments will be due two weeks from the date that they are assigned.
A late homework assignment will receive a reduced grade according to the following formula:

$$f(\text{grade}) = e^{-0.35 * \text{number of days late}}$$

Exams There is one midterm exam and one final exam. Both the midterm and final will be in-class. There are no quizzes. No communication with anyone besides the instructor is allowed on these assessments.

Discussion about grades Students are welcome to discuss how their assignments were graded. However, students have one week to discuss grades with the instructor after which grades are final.

Extra Credit Extra credit is available in several ways: attending an out-of-class lecture (as will be announced) and writing a short review of it; pointing out a substantial mistake in the book, a homework exercise or exam solution, If you read closely and to this point in the syllabus you can receive five percent extra credit on your first quiz by emailing me the name of your favorite icecream flavor and if you don't have a favorite, make one up; drawing my attention to an interesting data set or news article; etc. The extra credit is typically applied when a student is near the boundary of a letter grade.

Grading When grading your written work, I am looking for solutions that are technically correct and reasoning that is clearly explained. *Numerically correct answers, alone, are not sufficient* on homework, tests or quizzes. Neatness and organization are valued, with brief, clear answers that explain your thinking. If I cannot read or follow your work, I cannot give you full credit for it.

Accommodations for Students with Disabilities Lehigh University is committed to maintaining an equitable and inclusive community and welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact Disability Support Services (DSS), provide documentation, and participate in an interactive review process. If the documentation supports a request for reasonable accommodations, DSS will provide students with a Letter of Accommodations. Students who are approved for accommodations at Lehigh should share this letter and discuss their accommodations and learning needs with instructors as early in the semester as possible. For more information or to request services, please contact Disability Support Services in

person in Williams Hall, Suite 301, via phone at 610-758-4152, via email at indss@lehigh.edu, or online at https://studentaffairs.lehigh.edu/disabilities.

The Principles of Our Equitable Community: Lehigh University endorses The Principles of Our Equitable Community. We expect each member of this class to acknowledge and practice these Principles. Respect for each other and for differing viewpoints is a vital component of the learning environment inside and outside the classroom.