CSE 160: Introduction to Data Science

Fall 2020 – Online and Asynchronous
Except for occasional W1:35-2:25 synchronous labs

Catalog Description

Data Science is an interdisciplinary field focusing on the computational analysis of data to extract knowledge and insight. This course introduces the student to the collection, preparation, analysis, modeling and visualization of data, covering both conceptual and practical issues. Examples from diverse fields will be presented, and hands-on use of statistical and data manipulation software will be included.

Prerequisites: CSE 2, 4, 7 or 12 or BIS 335 or other programming experience (with permission of instructor)

Teaching Mode

The primary lectures will be asynchronous. Mondays during class time will be Prof. Davison's zoom office hours (open to all). Wednesdays during class time will be either synchronous online zoom Labs or Prof. Davison's zoom office hours (just for this course). Fridays during class time will be zoom office hours with TA Dan Luo.

Student Learning Outcomes

After taking Introduction to Data Science, you will:

   i.   Recognize the various disciplines that contribute to a successful data science effort.
   ii.   Understand the processes of data science: identifying the problem to be solved, data collection, preparation, modeling, evaluation and visualization.
   iii. Be aware of validity challenges as well as ethical issues that arise in data science tasks.
   iv.   Develop an appreciation of the many techniques for data modeling and mining.
   v.   Be comfortable using computational tools such as the R language and its associated libraries for data analytics and visualization.

Contact Information

Prof. Brian D. Davison – http://www.cse.lehigh.edu/~brian/
E-mail: davison@cse.lehigh.edu or bdd3@lehigh.edu

Teaching Assistant: Dan Luo
Graders: Tori Dorn and Felix Quintana
Tutoring/Help/Office Hours: to be posted on Piazza

For fastest response, use the Piazza site to post your questions. You should get an answer quickly from the instructor, the TA, a grader, or a fellow student. Posts can be anonymous or private to the staff.

Online Resources

Schedule, Lectures, Notes, Homework, Projects, Grades: http://coursesite.lehigh.edu/
Announcements, Discussions: The Piazza plug-in within CourseSite.

Textbooks (required)

Data Science for Business. By F. Provost and T. Fawcett. O'Reilly, 2013. ISBN 978-1-449-36132-7. Available for reading online: https://asa.lib.lehigh.edu/Record/10757826

An Introduction to Data Science. By J. Saltz and J. Stanton. Sage, 2018. ISBN 978-1-5063-7753-7.

(Subject to change)

Topics Expected to be Covered
Introduction to the field of data science; data collection; experimental design; data attributes; data cleaning; data characterization and analysis; data modeling and mining techniques; model evaluation; visualization; applications of data science; R scripting. Along the way we will also discuss aspects of privacy, security and social impacts. The course will also include a number of guest speakers to introduce students to the variety of applications of data science.

Academic Honesty
Unless specifically permitted otherwise, **the work you submit must be entirely your own**. While we encourage you to discuss basic concepts and strategies with friends and classmates, the copying or sharing of solutions, in whole or in part, is never acceptable. Both the person receiving the copied work and the person providing the copied work are equally responsible. Such cases will be referred to the University Committee on Discipline and, if found guilty, you may be given a failing grade in the course (or worse). If you have questions about this policy at any point throughout the semester, ask. It is far better to be safe than sorry when your academic career may be on the line.

COVID-19 / Online Course
Please contact the instructor if you are experiencing a problem with respect to the ongoing pandemic or to your ability to succeed in an online course (computer issues, internet problems).

Grading
Grades will be a function of homework (25%) and quizzes (10%), collaborative classwork (15%) and exams. You are responsible for everything that occurs in class as well as assigned readings. **Late assignments will be penalized 10% per day, up to three days late; after that, no credit is possible.** There will be three (cumulative) midterm exams (10% each) and a final group project (20%). Exam dates are announced at the beginning of the semester. Missed exams without a legitimate excuse will result in a score of 0. If at least 75% of the class completes the end-of-semester course evaluations, I will drop the lowest quiz grade. Final letter grades are assigned at the discretion of the instructor (i.e., not always using a fixed metric).

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 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.

Principles of 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.

Course Application
This course is one of many that count toward the Certificate in Business Analytics (for students in CBE) and is required for the Minor in Data Science (available to all undergraduates). It also counts as a Science and Technology course for CS majors (BS degree), and as a CS elective for the CS minor.

Last revised: 19 August 2020