

# CSE 160: Introduction to Data Science

Fall 2017 – 10:10am-11:00am MWF – STEPS 101

## 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.

## 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.

## Prerequisites

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

## Contact Information

Prof. Brian D. Davison

E-mail: [davison@cse.lehigh.edu](mailto:davison@cse.lehigh.edu)

Office Hours: Tuesdays and Wednesdays 11-noon (or by appointment) in Packard Lab 514C

Graders: Hannah Lambert and Dana Cunningham

## Online Resources

Homework, Projects, Grades: <http://coursesite.lehigh.edu/>

Announcements, Discussions, Notes: <http://piazza.com/lehigh/fall2017/cse160>

## Textbook (required)

**Data Science for Business**. By F. Provost and T. Fawcett. O'Reilly, 2013. ISBN 978-1-449-36132-7.

## Supplements

**Beginning R: An Introduction to Statistical Programming, 2<sup>nd</sup> Ed.** By J.F. Wiley and L.A. Pace, Apress, 2015. ISBN 978-1-4842-0374-3.

**An Introduction to Data Science**. By J. Stanton, 2013. Available free online at <https://drive.google.com/file/d/0B6iefdnF22XQeVZDSkxjZ0Z5VUE/edit>

## 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 lecturers throughout the semester 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.

## Grading

Grades will be a function of homework and quizzes (collectively worth 30%), class participation (10%) and exams. There may be one or two guest talks outside of class. Attendance is required (classes and guest presentations) and pop quizzes may occur at any time. 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.

There will be two in-class exams (20% each) and a final project (20%). Missed exams without a legitimate excuse will result in a score of 0 on that exam.

## Accommodations for Students with Disabilities

If you have a disability for which you are or may be requesting accommodations, please contact both your instructor and the Office of Academic Support Services, Williams Hall, Suite 301 (610-758-4152) as early as possible in the semester. You must have documentation from the Academic Support Services office before accommodations can be granted.

## 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.

Last revised: 15 August 2017