

YData: An Introduction to Data Science

Lecture 01: Introduction

Jessi Cisewski-Kehe and John Lafferty
Statistics & Data Science, Yale University
Spring 2019

Credit: data8.org



YData: Instructors



Jessi Cisewski-Kehe

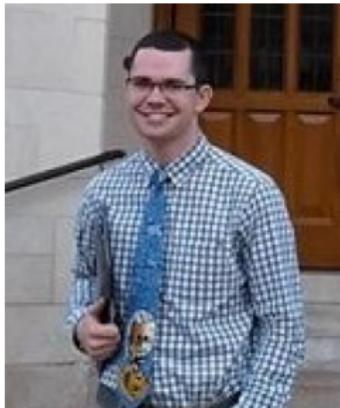


John Lafferty

YData: Teaching fellows



Brandon Chow



Parker Holzer



Prateek Malik

Schedule and files:
<http://ydata123.org/sp19/>

Course info, grades, etc:
<https://canvas.yale.edu>

YData website: <http://ydata123.org>

What is Data Science?

Drawing useful conclusions from data using computation

- **Exploration**

- Identifying patterns in information
- Uses visualizations

- **Inference**

- Quantifying whether those patterns are reliable
- Uses randomization

- **Prediction**

- Making informed guesses
- Uses machine learning

YData Seminar Courses

- Data science is driven by applications
- Every data-driven subject brings new challenges
- YData seminars are small, independent courses taught by Yale faculty who are excited to share their expertise
- We encourage you to consider enrolling in one of the three YData seminars offered this semester

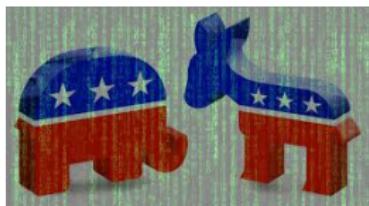
Currently available YData Seminar Courses

YData: ExoStatistics:
Exploring Extrasolar Planets with Data Science
(S&DS 170) Jessi Cisewski-Kehe



YData: Text Data Science: An Introduction
(S&DS 171) John Lafferty

YData: Data Science for Political Campaigns
(S&DS 172) Joshua Kalla



Course Structure

- Three lectures per week
- Weekly homework assignments
(lowest homework grade will be dropped at end of semester)
- Weekly practice exercises (ungraded)
- Three projects
- Drop-in office hours (see Canvas)
- Midterm during lecture hour on Wednesday, March 6, 2019
- Final exam scheduled on Sunday, May 5, 2019 at 2PM

Details can be found at <https://canvas.yale.edu>

Computational and Inferential Thinking: The Foundations of Data Science

By Ani Adhikari and John DeNero ([Adhikari and DeNero, 2018](#))

Freely available at <https://www.inferentialthinking.com>

Getting Help

- Ask a friend
- Ask on Piazza
<https://piazza.com/yale/spring2019/sds123/home>
Participation on Piazza through asking and answering questions is *strongly* encouraged
- Come to office hours

Collaboration

Asking questions is encouraged

- Discuss questions with each other (except on exams)
- Submit homework individually, but discuss with others (don't share written solutions or code)
- Submit projects individually or with one partner (only undergraduates may work with a partner)

The limits of collaboration

- Don't share solutions with each other (except project partners)
- Copying or other dishonesty will result in failing the course

Example

(DEMO)

References

Adhikari, A. and DeNero, J. (2018), "Computational and Inferential Thinking," Gitbook.