YData: An Introduction to Data Science

Lecture 01: Introduction

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

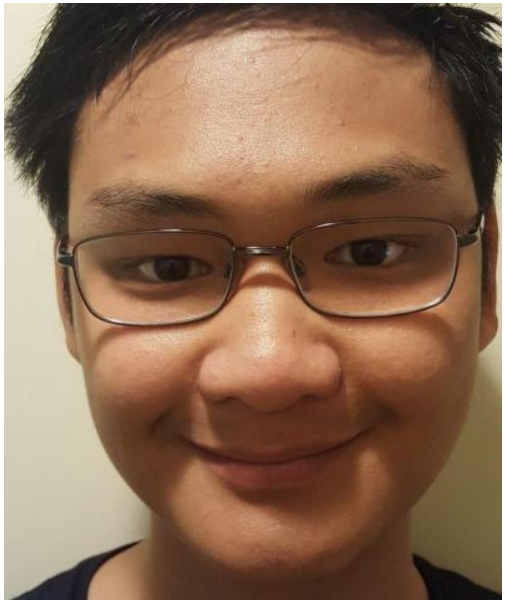
Credit: data8.org

YData: Instructors

Jessi Cisewski-Kehe

John Lafferty

YData: Teaching fellows



Brandon Chow

Parker Holzer

Prateek Malik

Course website

http://ydata123.org/sp19/

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 Labs are small, independent courses taught by Yale faculty who are excited to share their expertise

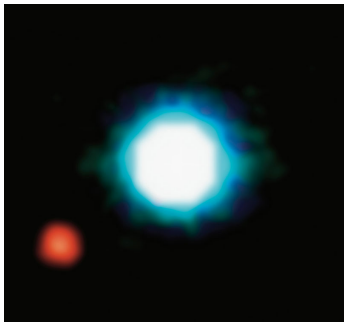
AddLabURL

Currently available YData Seminar Courses

YData: ExoStatistics:

Exploring Extrasolar Planets with Data Science

(S&DS 170) Jessi Cisewski-Kehe

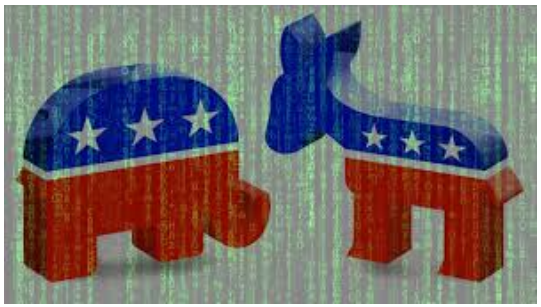


YData: Text Data Science: An Introduction

(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
- Three projects
- **List office hours**
- Midterm during lecture hour on Wednesday, March 6, 2019
- Final exam scheduled on Sunday, May 5, 2019 at 2PM

Details can be found at AddLinkToCoursePolicies

Course Textbook

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
- Come to office hours

Course Policies

Learning

Community

Course Staff

AddLinkToCoursePolicies

Collaboration

Asking questions is encouraged

- Discuss questions with each other (except on exams)
- Submit homework individually, but discuss with others
- Submit projects individually or 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.