Lecture 1, August 24

DATA 8

Data Science

Slides created by Ani Adhikari and John DeNero

Welcome!

http://data8.org

- Collaboratively developed by faculty across campus
- Taught by a team of instructors
- And by a team of (U)GSIs, tutors, and lab assistants

- Ani Adhikari (413 Evans)
- Office hours: 11-12:30 Monday; 10-11:30 Tuesday
- And you'll find me around the labs, in the (U)GSI/tutor office hour rooms, etc etc. Generally everywhere.
- Email: adhikari@berkeley.edu

Data Science

What is Data Science?

Drawing useful conclusions from data using computation

- **Exploration**
  - Identifying patterns in information
  - Uses visualizations
- **Prediction**
  - Making informed guesses
  - Uses machine learning and optimization
- **Inference**
  - Quantifying our degree of certainty
  - Uses stochastics and statistical decision theory

Connector Courses

- Data science is driven by applications

- Every data-driven subject brings new challenges

- Connectors are small, independent courses taught by Berkeley faculty who are excited to share their expertise

http://data8.org/connector

Course Structure

Parts of Data 8

- Lecture has a screencast, but it's better to attend

- Weekly lab

- Weekly homework assignments & a few projects

- (U)GSI office hours tentatively Monday-Friday 2-5, starting next week

- Midterm exam in class (Fri 10/14)
- Final exam (Mon Dec 12, 8 a.m. to 11 a.m.)

Collaboration

Asking questions is highly encouraged

- Discuss all questions with each other (except exams)
- Submit lab assignments individually (graded on completeness)
  - If you come to lab, you can collaborate liberally
  - If you choose not to come to lab, you must work alone
- Submit homework individually and list collaborators
- Submit projects in pairs; find a partner in your lab

The limits of collaboration

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

Finally … some data

(Demo)