Data 8, Lab 1

Hubert Luo

Spring 2020

24 January 2020

Berkeley
UNIVERSITY OF CALIFORNIA

I'm a senior from Toronto, Canada studying stats and applied math. This is my fourth semester on Data 8 staff, as well as my eighth and final semester teaching at Cal.

- **Email:** hubertluo@berkeley.edu
- **Lab Website:** https://hluo27.github.io/data8
- **Lab:** Friday 12-2pm in Evans 458
- **Office Hours:** Every Weekday from 12-5pm

Berkeley
UNIVERSITY OF CALIFORNIA

Lab Overview

Typical format of lab sections:

- Worksheet
  - Written practice problems we will go over in lab
  - Useful to review problems in a written format since the exams will be written on paper as well
- Lab Notebook
  - Review concepts by applying them to actual data
- Attendance for entire lab section needed to get checked off for lab notebook
- Usually we'll have 3 lab assistants – they'll join us next week!

Lab Policies

- Students can only attend the lab they are enrolled in
- Lab assignments released on Monday. Two ways to get credit:
  1. Attend **entire** lab section and make substantial progress
  2. Finish and pass all autograder tests by Wednesday at 8:59 AM
- Waitlisted Students
  - Enrolled once there is an open seat in lecture + lab section
  - Course staff do not know what position students are at on the waitlist and have no control over it

Course Policies

- Weekly homework and weekly labs can't be turned in late
- No alternate midterm or final exam
- DSP students must ask their DSP advisor to submit a letter through SCARAB for DSP accommodations
- See http://data8.org/sp20/policies.html for more details

Grading Logistics

- Each student should have Gradescope and Okpy accounts
  - Let me know if you don't have an account on either yet!

Okpy

- All labs, homework, and projects will be submitted to Okpy
- Code is auto-graded and marks posted on Okpy

Gradescope

- Written answers on homework/project assignments and exams are graded by course staff on Gradescope

Introductions

Form groups of 3-4 and find something all of you have in common!

Example: During a previous semester, a group of 4 all had a family relative with the same name!

Discussion Worksheet

- Introduce yourself to the people sitting around you and work as a table
- Don’t use online resources!

Discussion: Solutions

1. 50 (Actual average size is 48.2)
2. Cog Sci (343), Applied Math (314), Sociology (313), English (282), and Integrative Biology (255)
3. Hawaii (81.3) and California (80.9)
4. America: 78, World: 72

Post-Discussion Review

- Numerical data needed to answer the four questions earlier
- Some questions can be answered without error
  - Example: the two Berkeley questions are based on complete information
- Many questions cannot be answered without error because we have incomplete information
  - Example: the two LEB questions since we don't know exactly how long everyone is going to live but we need estimates based on assumptions

Post-Discussion Review

- There is always uncertainty around our estimates. How can we quantify it?
  - Example: Article about life expectancy in California only gives one number (80.9) but actual estimate is an interval (71.9 to 81.9)
  - Uncertainty means we need some wiggle room for the estimates
  - Other states could potentially have a higher estimate! Range for Connecticut is 79.7-81.8
- The world is fuzzy so we have to be able to quantify our uncertainties – we have to interpret data carefully.

Class Survey

Please fill out this form before starting the lab notebook:

http://tinyurl.com/data8-hubert