

Data 8 – Foundations of Data Science

- Computational Thinking + Inferential Thinking in the context of working with real world data
- Introduce you to several computational concepts in a simple data-centered setting
 - Authoring computational documents
 - Tables
 - Within Python3 and “SciPy”

8/26/16 UCB CS88 Sp16 L1 7

CS88 – Computational Structures in Data Science

- Deeper understanding of the computing concepts introduced in c8
 - Hands-on experience => Foundational Concept
 - How would you create what you use in c8 ?
- Extend your understanding of the structure of computation
 - What is involved in interpreting the code you write ?
 - Deeper CS Concepts: Recursion, Objects, Classes, Higher-order Functions, Declarative programming, ...
 - Managing complexity in creating larger software systems through composition
- Create complete (and fun) applications
- In a data-centric approach

1/25/16 UCB CS88 Sp16 L1 8

Pathways

8/26/16 UCB CS88 Sp16 L1 9

How does CS88 relate to CS61A ?

8/26/16 UCB CS88 Sp16 L1 10

Course Structure

- 1 Lecture + 1 Lab/Discussion on Monday (!!!)
- Lecture introduces concepts (quickly)
- Lab provides concrete detail hands-on
- Homework (10) cements your understanding
 - Out Monday, Due Sunday
- Projects (3) put your understanding to work in building complete applications
 - Maps
 - Hangman
 - Open Projects!
- Readings: <http://composingprograms.com>
 - Same as cs61a

8/26/16 UCB CS88 Sp16 L1 11

CS88 Team - uGSIs

Dr. Gerald Friedland
fractor@eecs.berkeley.edu



Gunjan Baid
cunjan_baid@berkeley.edu

Lab Assistants (hopefully):
 Anthony Xian, Rana Zee Maneri, Dashiell Brennan Stander, Pransu Dash, Niharika Jain, David Sang-chul Nahm, Minsu Kim, Caleb Casimir Chuck, Daniel Bernard Ricciardelli, Rena Chen, Kenneth Kao, Andrew Tan, Peter Yuan, Arman Madani, Calvin Dong, Erik Sanders Cheng

8/26/16 UCB CS88 Sp16 L1 12

CS88 Team - me

- Dr. Gerald Friedland (fractor@berkeley.edu)
 - 424 Saturday Daj Hall (CITRIS)
 - <http://www.gerald-friedland.org>
 - Office hours: Fr 1-2 @ 424 SDH
 - Before/after class



- Adjunct Assistant Professor, EECS UC Berkeley
- Principal Data Scientist, Lawrence Livermore National Labs

8/26/16 UCB CS88 Sp16 L1 13

CS88 Team - me

Projects you might want to check out:



- <http://mmcommons.org>
 - Work with 100M images, 1M videos in your own Amazon instance.
- <http://www.teachingprivacy.org>
 - Creating teaching materials informing about data over sharing.

1/25/16 UCB CS88 Sp16 L1 14

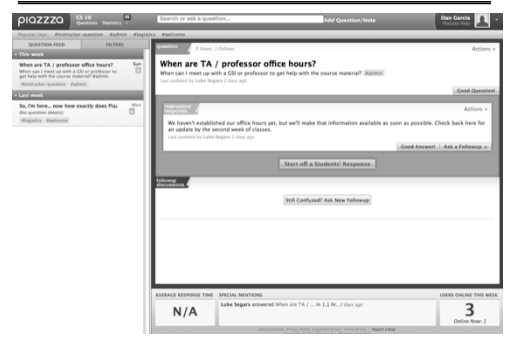
Course Culture

- Learning
- Community
- Respect
- Collaboration
- Peer Instruction

1/25/16 UCB CS88 Sp16 L1 15

Piazza for {ask,answer}ing questions



8/26/16 UCB CS88 Fa16 L1 16


Pro-student Grading Policies

- EPA
 - Rewards good behavior
 - Effort
 - » E.g., Office hours, doing every single lab, hw, reading Piazza pages
 - Participation
 - » E.g., Raising hand in lec or discussion, asking questions on Piazza
 - Altruism
 - » E.g., helping other students in lab, answering questions on Piazza
- You have 2 "Slip Days"
 - You use them to extend due date, 1 slip day for 1 day extension
 - You can use them one at a time or all at once or in any combination
 - They follow you around when you pair up (you are counted individually)
 - » E.g., A has 2, B has 0. Project is late by 1 day. A uses 1, B is 1 day late

UCB CS88 Fa16 L1 17

Abstraction

- Detail removal
 - "The act or process of leaving out of consideration one or more properties of a complex object so as to attend to others."
- Generalization
 - "The process of formulating general concepts by abstracting common properties of instances"



Henri Matisse "Naked Blue IV"

8/26/16 UCB CS88 Fa16 L1 18

Experiment

8/26/16 UCB CS88 Fa16 L1 19

Where are you from?

Possible Answers:

- China
- California
- The Bay Area
- San Mateo
- 1947 Center Street, Berkeley, CA
- 37.8693° N, 122.2696° W

All correct but different levels of abstraction!

8/26/16 UCB CS88 Fa16 L1 20

Abstraction gone wrong!

8/26/16 UCB CS88 Fa16 L1 21

Detail Removal (in Data Science)

- You'll want to look at only the interesting data, leave out the details, zoom in/out...
- Abstraction is the idea that you focus on the essence, the cleanest way to map the messy real world to one you can build
- Experts are often brought in to know what to remove and what to keep!

The London Underground 1928 Map & the 1933 map by Harry Beck.

8/26/16 UCB CS88 Fa16 L1 22

The Power of Abstraction, Everywhere!

- **Examples:**
 - Functions (e.g., $\sin x$)
 - Hiring contractors
 - Application Programming Interfaces (APIs)
 - Technology (e.g., cars)
- **Amazing things are built when these layer**
 - And the abstraction layers are getting deeper by the day!

We only need to worry about the interface, or specification, or contract NOT how (or by whom) it's built

Above the abstraction line

Abstraction Barrier (Interface)
(the interface, or specification, or contract)

Below the abstraction line

This is where / how / when / by whom it is actually built, which is done according to the interface, specification, or contract.

8/26/16 UCB CS88 Fa16 L1 23

Abstraction in CS: Data Type

- What's this?

8/26/16 UCB CS88 Fa16 L1 24

Data Types and Operations



- **Set of elements**
 - with some internal representation
 - E.g. Integers, Floats, Booleans, Strings, ...
- **Set of operations on elements of the type**
 - e.g. +, *, -, /, %, //, **
 - ==, <, >, <=, >=
- **Properties**
 - Commutative, Associative, ... , Closure (???)
- **Expressions are valid well-defined sets of operations on elements that produce a value of a type**

8/26/16

UCB CS88 Fa16 L1

25

Questions



- What's the difference between '==' and '=' ?

8/26/16

UCB CS88 Fa16 L1

26

Lab and HW this week



- **Lab will get you to where you have a *program development environment***
 - Even on your computer
- **HW will give practice and explain subtleties of types, operators, and expressions**
 - In a program development environment

8/26/16

UCB CS88 Fa16 L1

27