



CE 88: Fall 2017 Data Science Connector Course

---

# Data Science for Smart Cities

## Introduction to smart cities

---

Instructor: Milad Memarzadeh

GSI: Sangjae Bae

Mondays 12-2pm, 406 Davis Hall



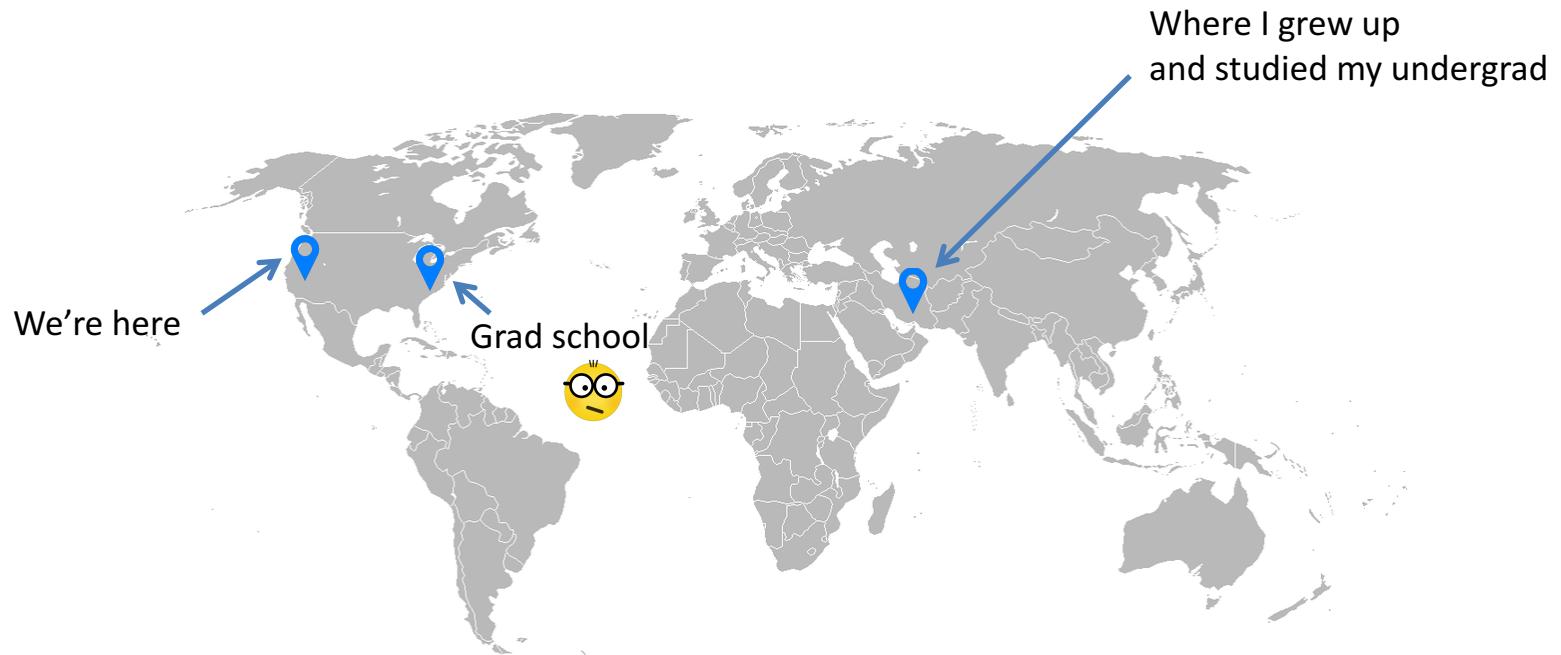
---

 Milad Memarzadeh

 @miladmemarzadeh

 [www.miladmemarzadeh.com/ce88](http://www.miladmemarzadeh.com/ce88)

 miladm@berkeley.edu



# Topics I have work/ed on...

---

<http://www.miladmemarzadeh.com/research>

# What do I do if I'm not working!

I play soccer and I'm pretty good ☺



and hiking once in a while.



# About GSI

---

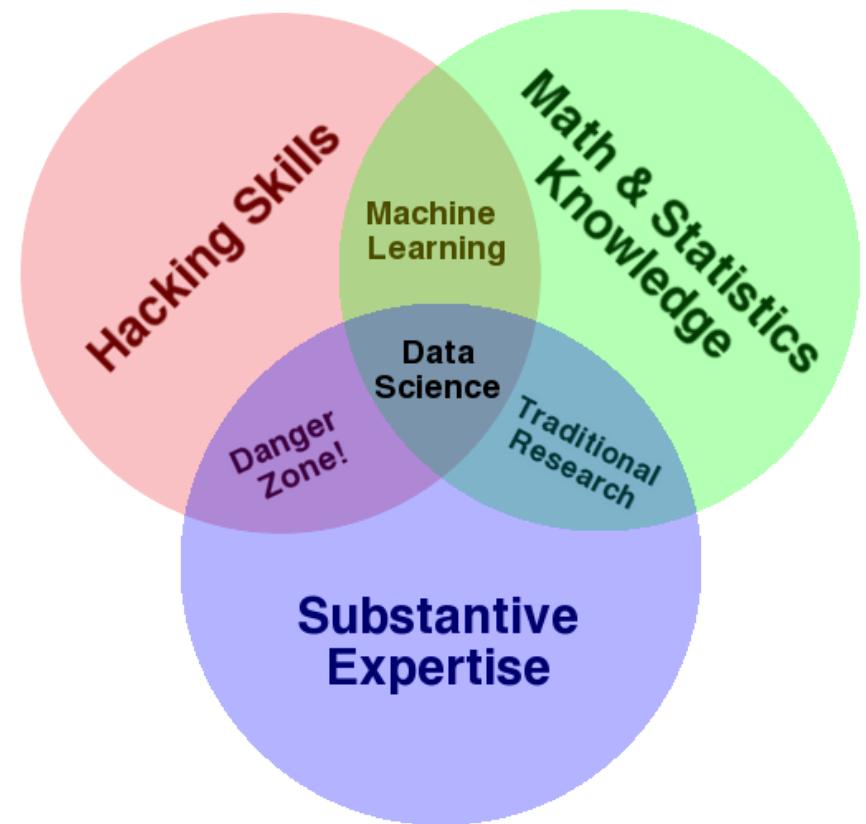
- Email: [sangjae.bae@berkeley.edu](mailto:sangjae.bae@berkeley.edu)
- Office Hours: Friday 10-12, 537 Davis
- About me:
  - I am PhD student in Systems Engineering
  - I am working with Energy, Controls, & Applications Lab
  - I am from Korea (oh, yeah, South)
  - I have EECS background
  - I am a terrible cook so please do not ask.
  - I like to play sports & outdoor activities.  
(I am currently into Squash!)
  - Learn more here: <https://sangjaebae.com>



# What is Data Science

---

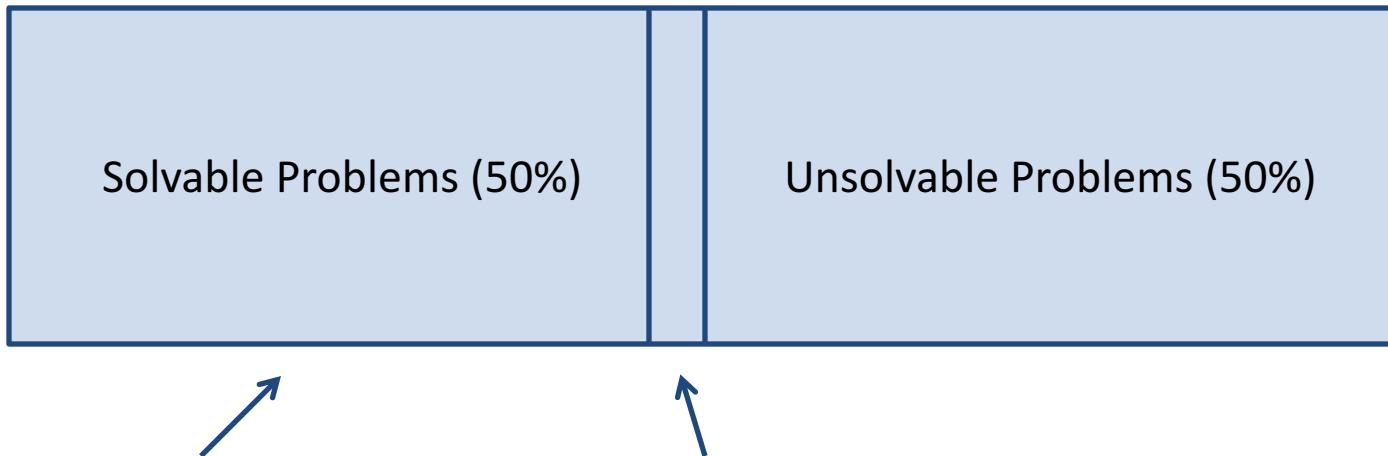
Data science is the application of **computational** and **statistical** techniques to address or gain insight into some problem in the **real world**



# Data Science is not machine learning

---

Data problems we would like to solve



---

Adapted from Zico Kolter (CMU)

# Data Science is not Big Data

---

- Sometimes, in order to truly understand and answer your question, you need massive amounts of data.
- But most of the time you don't, because most of the data is small.

# What is Big Data?

---

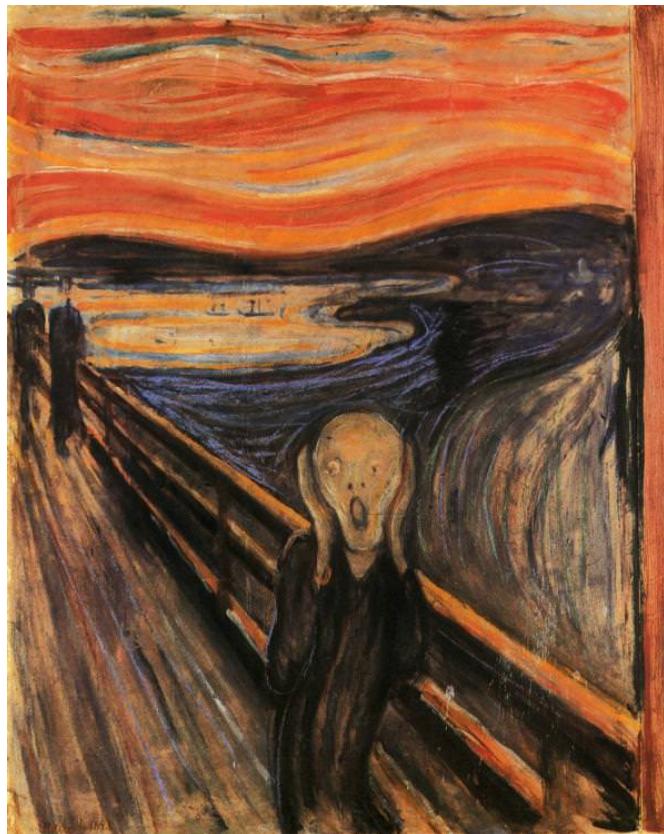
"Big data is like teenage sex. Everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it." – *Dan Ariely, Professor at Duke.*

Big data is blue...



# Why do we need smart cities?

---



GPA: D+



2017

INFRASTRUCTURE  
REPORT CARD



C+

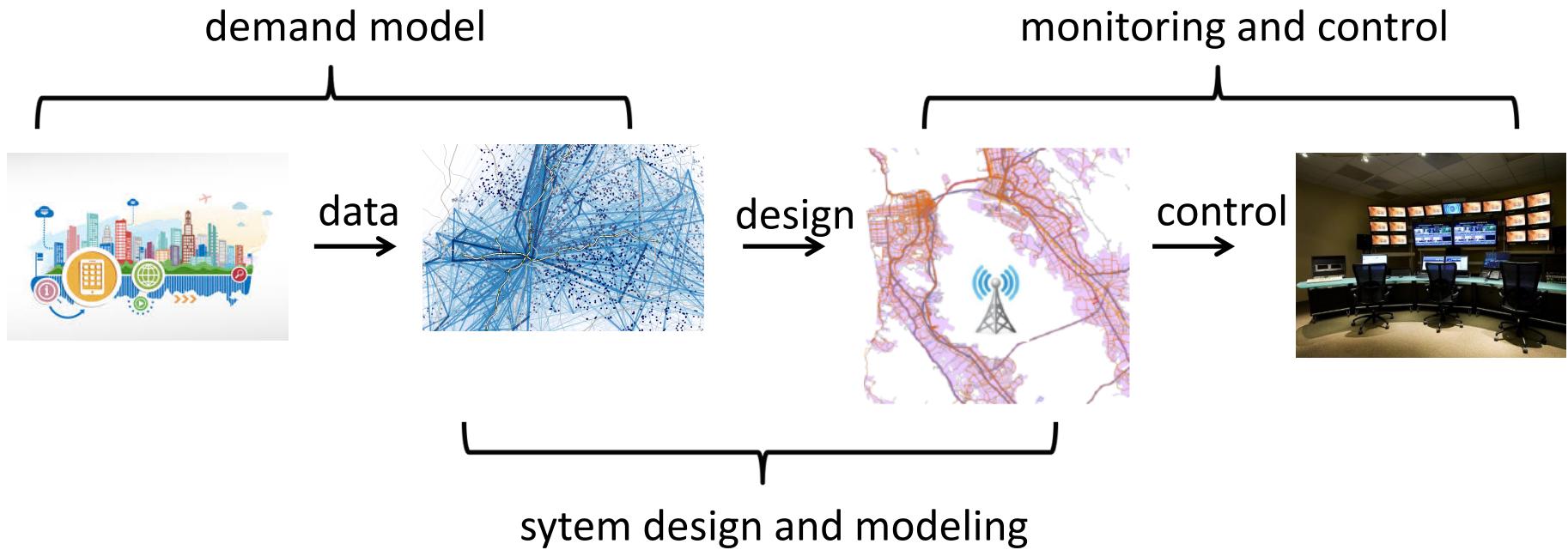


D+



D

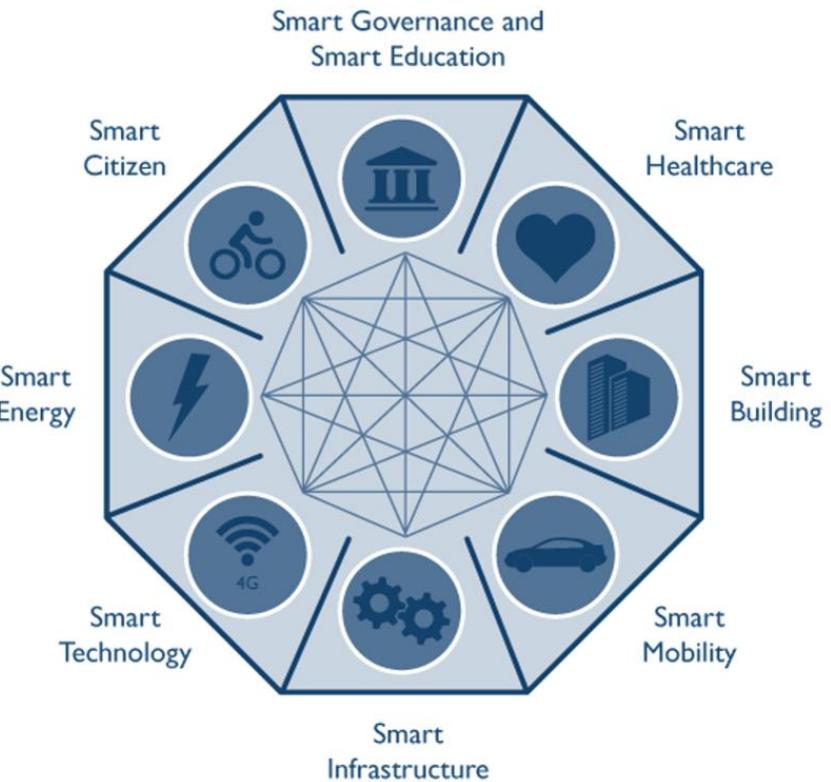
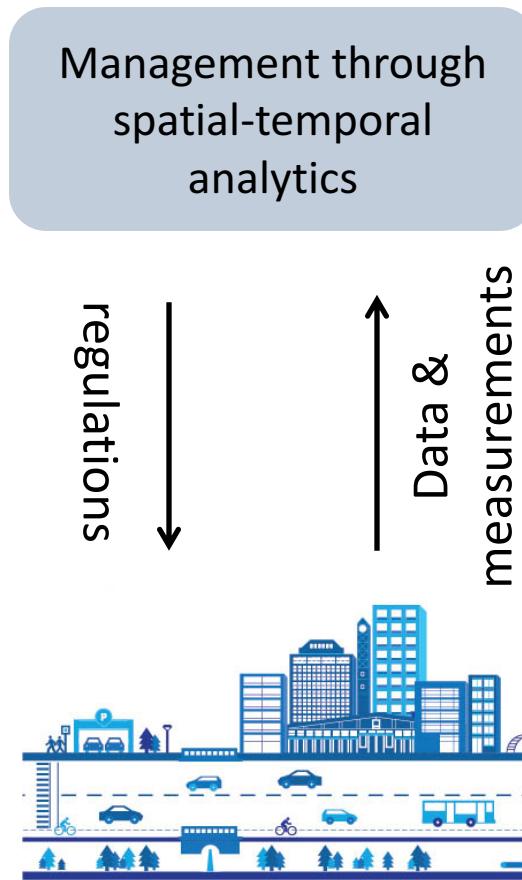
# The role of data



“If you are looking for a career where your services will be in **high demand**, you should find something where you provide **a scarce, complementary service** to something that is getting **ubiquitous and cheap**. So what’s getting ubiquitous and cheap? **Data**. And what is complementary to data? **Analysis**.”

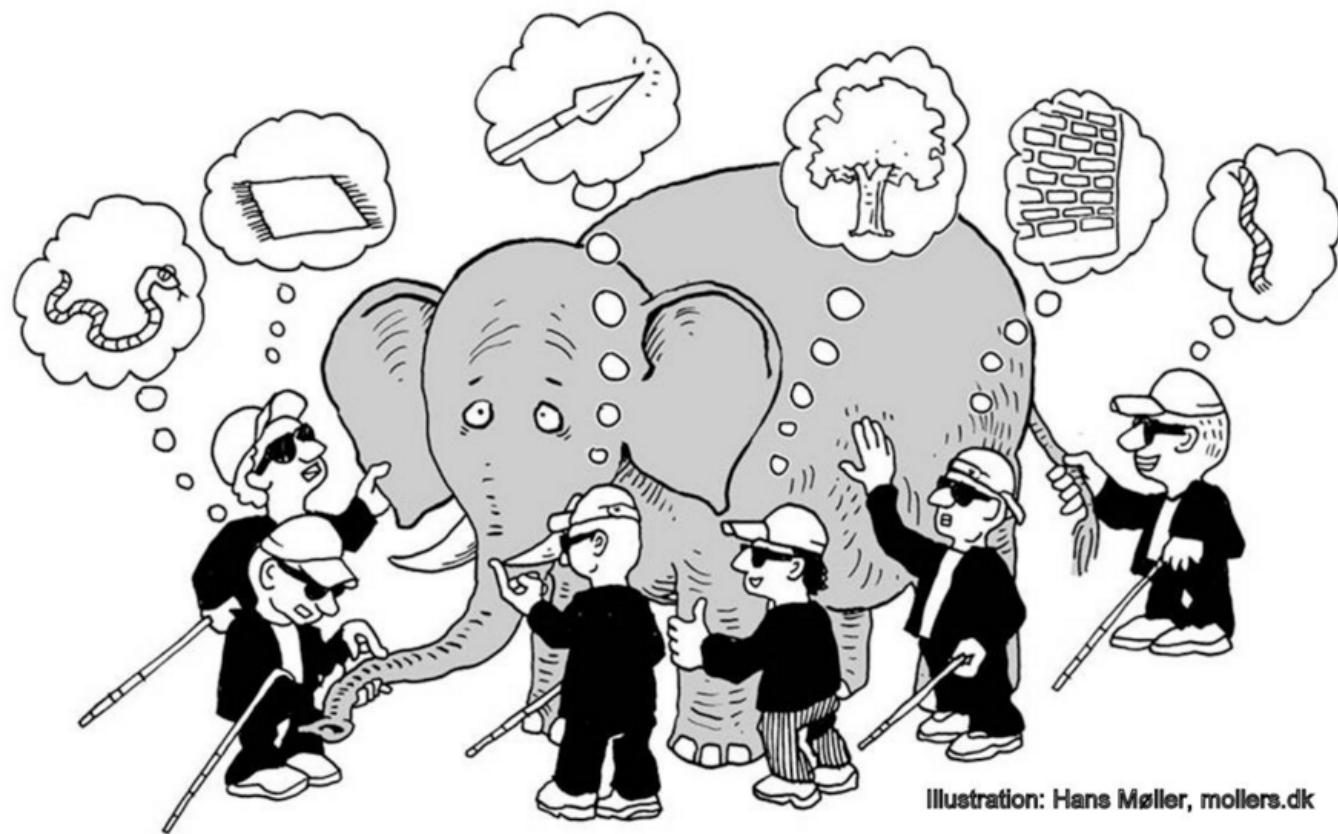
– Prof. Hal Varian, UC Berkeley, Chief Economist at Google

# Smart city – Interconnected Infrastructure



# Why Interconnected?

---



# About the course

---

Course will be taught in **four modules**:

- Mobility & Transportation
- Energy Systems
- Extreme Events & Urban Resilience
- Climate Change & Environmental Variability

We will learn **programming**

We use simple **statistical analysis**

We work with **real-data**

# Jupyter Notebook

---



- Scripting language
- Widely used in data science and beyond
- `datascience` module
  - Good:
    - Developed uniquely for DS-8 by core faculty, staff and students
    - Provides `Table` objects: a data structure for data science
    - Power of abstraction
  - Challenges
    - Still under active development
    - little-known beyond this course (read: hard to Google)
    - What about `pandas`?

# Goals of the class

---

- Become familiar with several key concepts/areas in **civil and environmental engineering**
- Learn how to work with **real databases**
- Become comfortable in **python environment**
- Learn **statistical analysis** and apply them to data

# This is my first time... *teaching*

---

- We are going to learn together and have fun
- How to teach data science is an open question world-wide
- Pascal -> Fortran -> Java -> Matlab -> R -> Python

# Assessment

---

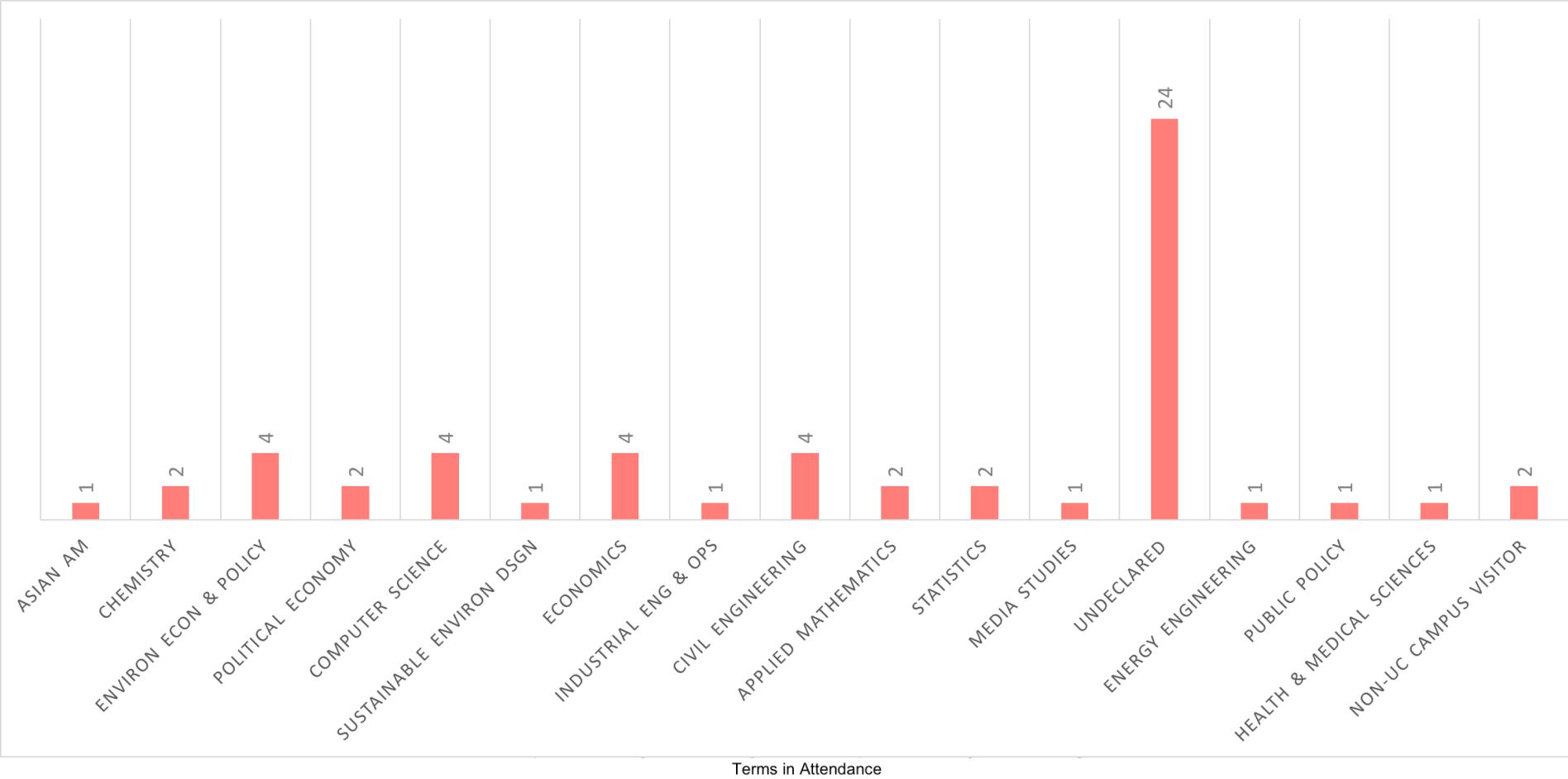
Participation (& mini quiz) 20%

Homework 40%

Final project 40%

# Class statistics

---



# Python Installation

---



ANACONDA®

<https://www.continuum.io/downloads>

# Piazza: our discussion (Q/A) platform

---



<https://piazza.com/berkeley/fall2017/ce88/home>

# OK: our submission platform

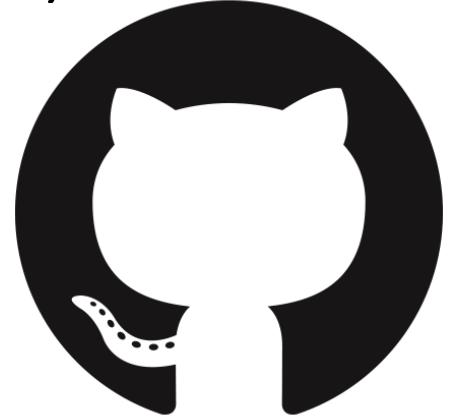
---

ok

<https://okpy.org/admin/course/55/>

# Github: our repository (don't worry about it)

---



<https://github.com/data-8/smart-cities-connector>

# Let's here from you

---

- Name
- Intended Major
- Class/year
- Previous programming experience
- Previous statistics training
- What led you to enroll in this course?
- What do you hope to get out of it?