

---

**INTRO TO DATA SCIENCE**

---

# **INTRO to DATA SCIENCE**

## **LECTURE 5.3: FINAL PROJECT KICKOFF**

Rob Hall

DAT13 SF // March 23, 2015

---

## FINAL PROJECT KICKOFF

---

### *Final Projects in Brief:*

- Select a dataset
- Formulate a hypothesis / problem statement
  - What question do you want to answer? / What do you want to predict?
  - Why is it useful to do so?
- Wrangle the data
- Build *at least one* model
- Test that model
- Write up your project and what you learned
- Present your project to the class (last week of class)

---

## FINAL PROJECT KICKOFF

---

### *What is the purpose of doing a final project?*

- Opportunity to apply everything you will have learned in this course
- Opportunity for your curiosity to come out and play
- Required to receive a “Letter of Completion” for this course
- Becomes part of your portfolio when seeking employment in data science
- There will be beer

### *Sources of inspiration:*

- Your work
- Cool projects you see online
- Stanford's CS 229 projects:  
<http://cs229.stanford.edu/projects2013.html>  
<http://cs229.stanford.edu/projects2012.html>
- Each other

---

## FINAL PROJECT KICKOFF

---

### *The “Elevator Pitch” (due Wednesday 4/1 in class)*

- A one-paragraph write up of the project you propose
- A concise (<90 seconds) verbal pitch of your project to the rest of the class
- Include:
  - A concise statement of the goal of your project
  - Where you will get the data / what dataset you plan to use
  - What type of machine learning problem this is (from our 2x2)
  - Why you think this is a cool project
- *NOTE: This does not lock you into that project. You can change your idea and your ML technique as we learn more. The point is to get moving on it :-)*