

# Introduction to: Tools & Resources

Data Science and Machine Learning I

*Data-X Plaksha Lec 2 Fall19*

Link to slides: [bit.ly/dxp-lec2](https://bit.ly/dxp-lec2)

**Alexander Fred-Ojala**  
Research Director, Data Lab  
SCET, UC Berkeley

**PLAKSHA**  
UNIVERSITY



**Berkeley**  
UNIVERSITY OF CALIFORNIA

# PLAN FOR THIS PART

- **Intro to Resources**
- **Math Test**
- **Intro to Python, Jupyter Lab, Notebooks, Git, and Setup**
- **Maybe: Project explanation**

# Data-X | Tools & Resources

# Data-X Plaksha website

[bit.ly/plaksha](https://bit.ly/plaksha)

- Current Syllabus
- HWs, deadlines, topics
- Recommended Readings
- Grading Rubric
- Contact info
- etc.

## Data-X Plaksha 2019

Course website

Lectures and HWs

[Link to Current Lecture Plan](#)

**Lecture Times (first 2 weeks):** Mon, Wed, Fri 10.30-12:15

**Location:** [Plaksha Innovation Center](#)

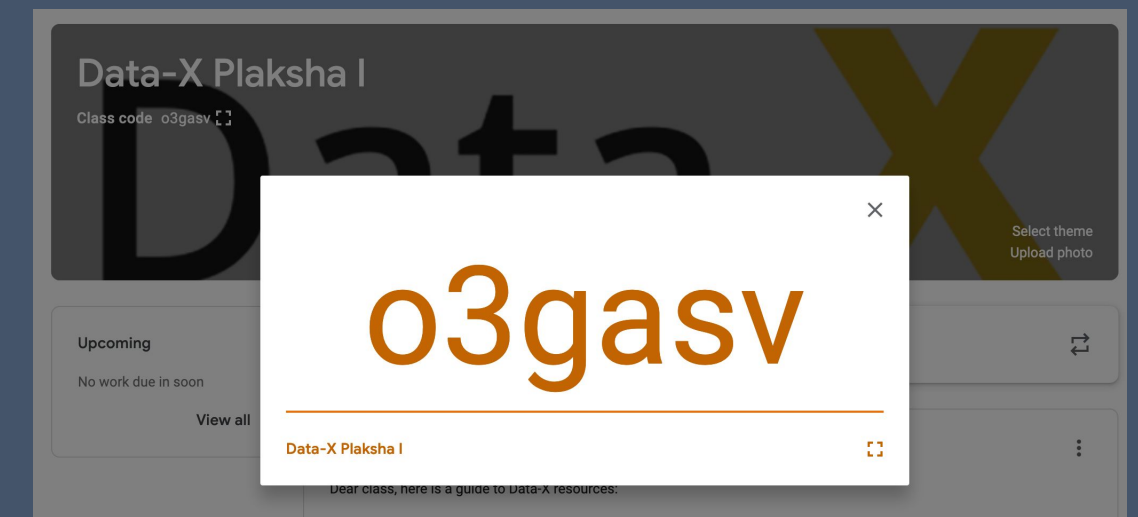
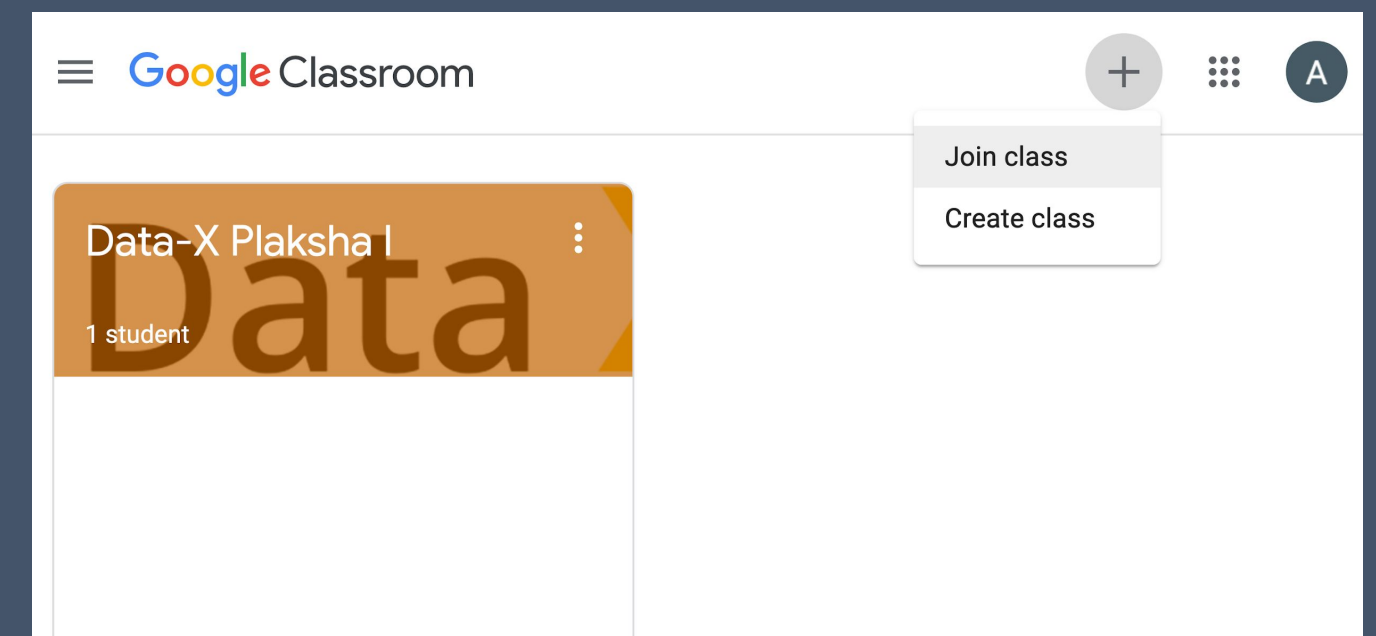
# Google Classroom

classroom.google.com

Code: o3gasv

Everyone should have been invited, if not add yourself!

- Announcements
- Assignments
- Grades
- Deadlines
- Materials



# Datacamp.com

Interactive online learning platform  
(programming, data science etc)

We have received free premium  
accounts that will be valid until Feb.

Sign up at [bit.ly/plaksha](https://bit.ly/plaksha)



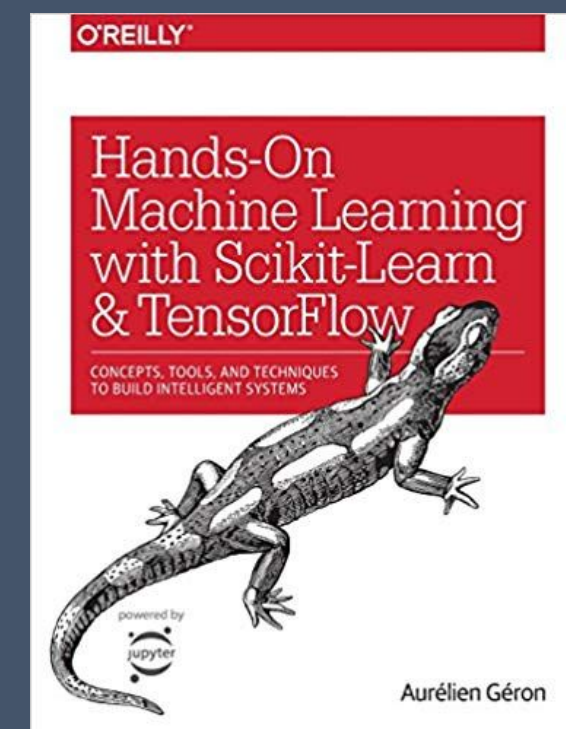
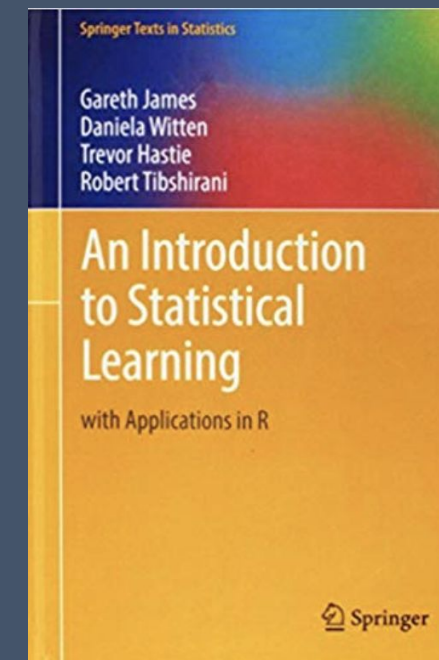
# (Informal) Course Textbooks

Introduction to Statistical Learning

Hands-On Machine Learning  
with Scikit-Learn and TensorFlow

Not mandatory, but highly  
recommended!

Find the links at [bit.ly/plaksha](https://bit.ly/plaksha)



# Cloud Credits

AWS: \$100 sign up  
+ package \$2k for 2 years

Google Cloud: \$300 sign up  
+ 6 \* \$50 in cloud credits

Email [afo@berkeley.edu](mailto:afo@berkeley.edu) if you need them.





# Python Bootcamp

14 videos (approx 14hrs) of wonderful  
Python tutorials.

Highly recommended!

Find the link at [data-x.blog/resources](https://data-x.blog/resources)



# Github

Everyone should have a Github account. Will be needed for project check-ins.

Get a premium acc at: [education.github.com/](https://education.github.com/)

All lecture slides and code samples will be pushed to this repo:

[github.com/afo/data-x-plaksha](https://github.com/afo/data-x-plaksha)

Here are Github Setup instructions:

[https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/github\\_setup.md](https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/github_setup.md)



# (Informal) Math Test

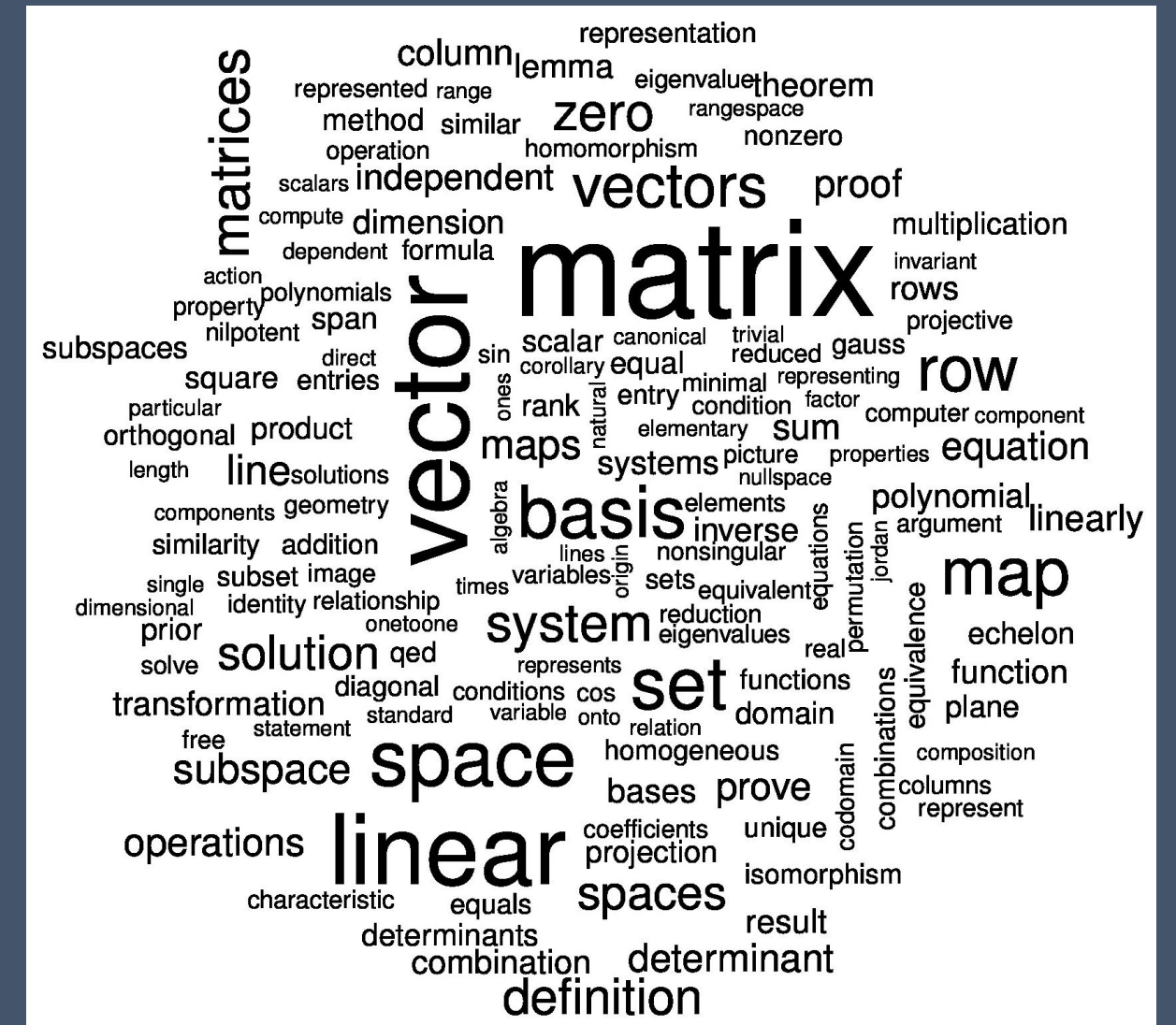
Everyone needs a piece of paper  
and a pen.

# Linear Algebra, Probability / Stats for ML

# Topics to know and materials to study published at Google Classroom (Week 0)

We will probably have a formal test (simple)  
that everyone needs to pass (you can retake it  
many times)

# Friday September 13th before lecture



# Anaconda Install Instructions

## Change to notebook

**Link to install instructions:**

**Anaconda:** [https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/anaconda\\_install\\_instructions\\_plaksha.pdf](https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/anaconda_install_instructions_plaksha.pdf)

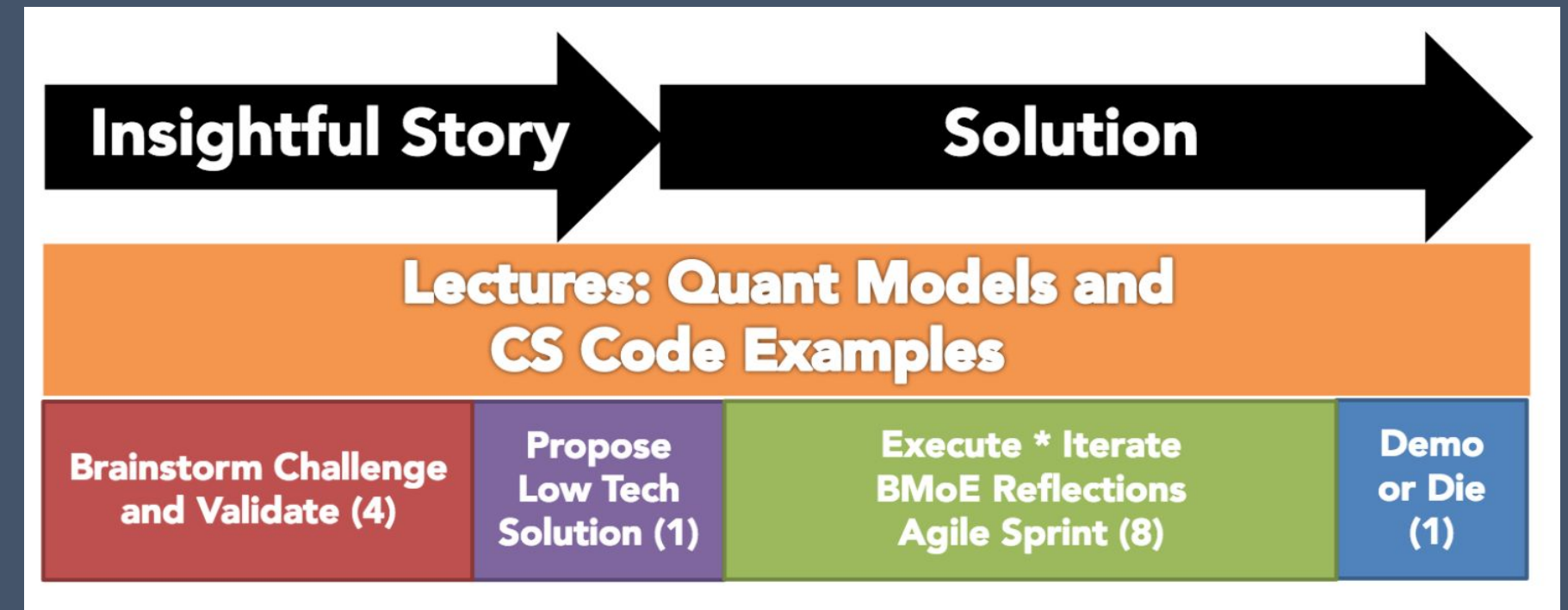
**Github:** [https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/github\\_setup.md](https://github.com/afo/data-x-plaksha/blob/master/00-install-instructions/github_setup.md)

**End of Lecture 2?**

# Project Data-X I: Explained

# Data-X I: Project

- **Totally open-ended!**
  - Collab with mentor
  - Find your own mentor
  - Come up with your own project
- **Groups of 5 (at most 6, at least 4)**
- **First weeks: Don't code.**  
**=> Lowtech demo.**



Find examples at [data-x.blog/projects](https://data-x.blog/projects)

## Project Categories:

Business Use Case	Social Impact	It's Just Cool
System Improvement	Improve Data Pipeline	Novel Research



# Lowtech Demo Template

Data-X  
Plaksha

Don't use this template, create your own,  
make it beautiful

Name of Project  
Project Overview

Name of team members

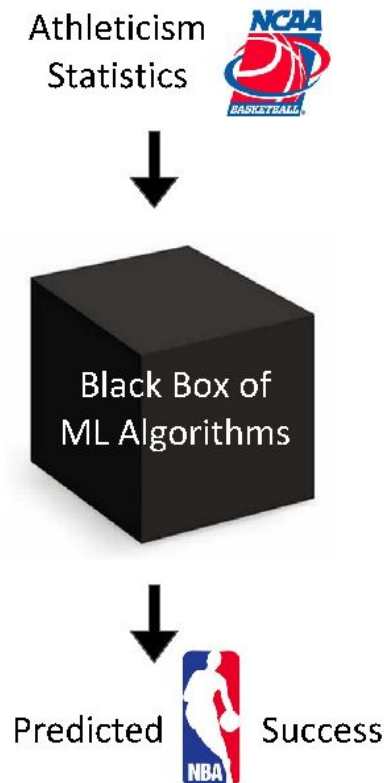
# What is Your Project / Pitch

- ✗ High Level Description
- ✗ Any validation information (did you talk with anyone who wants it)

## Example: Sport Prediction

### Basketball Player Growth Prediction - College to NBA

Members: Jessie Ji, Tu Ni, Fu-Chi Shih, Xinle Wang



Our potential Users (**coach managers, NBA scouts, etc**) are interested in:

1. key characteristics (e.g. hit rate (%), rebounds/min) of a college players in NCAA that predict the his future performance in NBA.
2. a prediction tool to evaluate players' talent in both the opposing and their own team to deploy strategies.



## Example: Predictive Policing



### Two potential purposes:

- 1) Safety app for students to alert them where there is a higher probability of crime
- 2) Analyze crime data and safety strategy data to determine where the police can optimally implement safety measures while minimizing cost

### User Validation:

- Would like notifications when walking into a red zone
- Worried about potential misuse/exploitation by other students
- Would like "optimal path" if they put in a destination

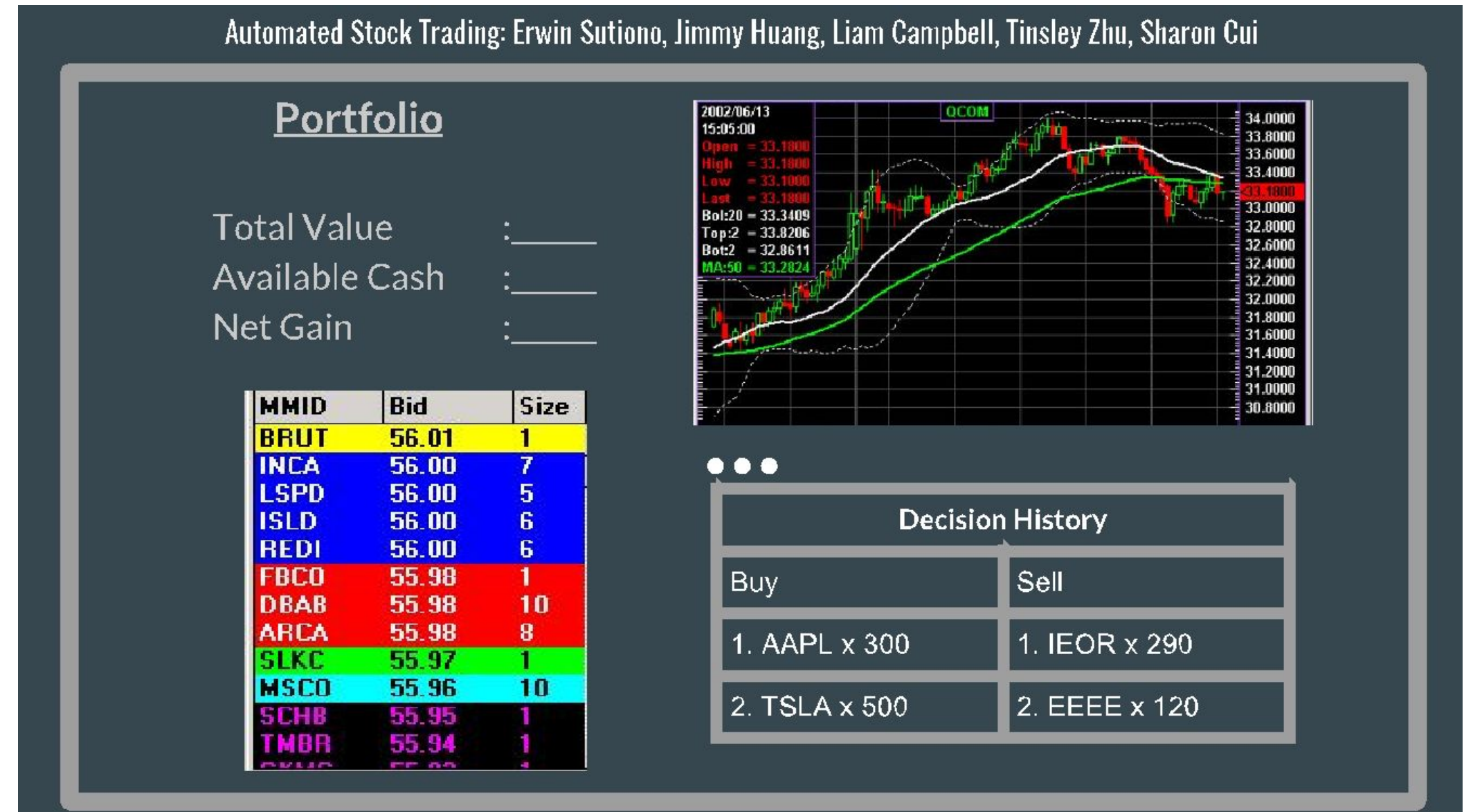
Project: Predictive Policing || Team Members: Smita Jain, Sandra Herchen, Jin Lee, Yijin Hua, Gavin Lee

# Intended User Interface

Example Intended  
Screenshot /  
Mockup

List Top 3 User  
Requirements  
(your best guess)

- Performance..
- Ease of Use..
- ...





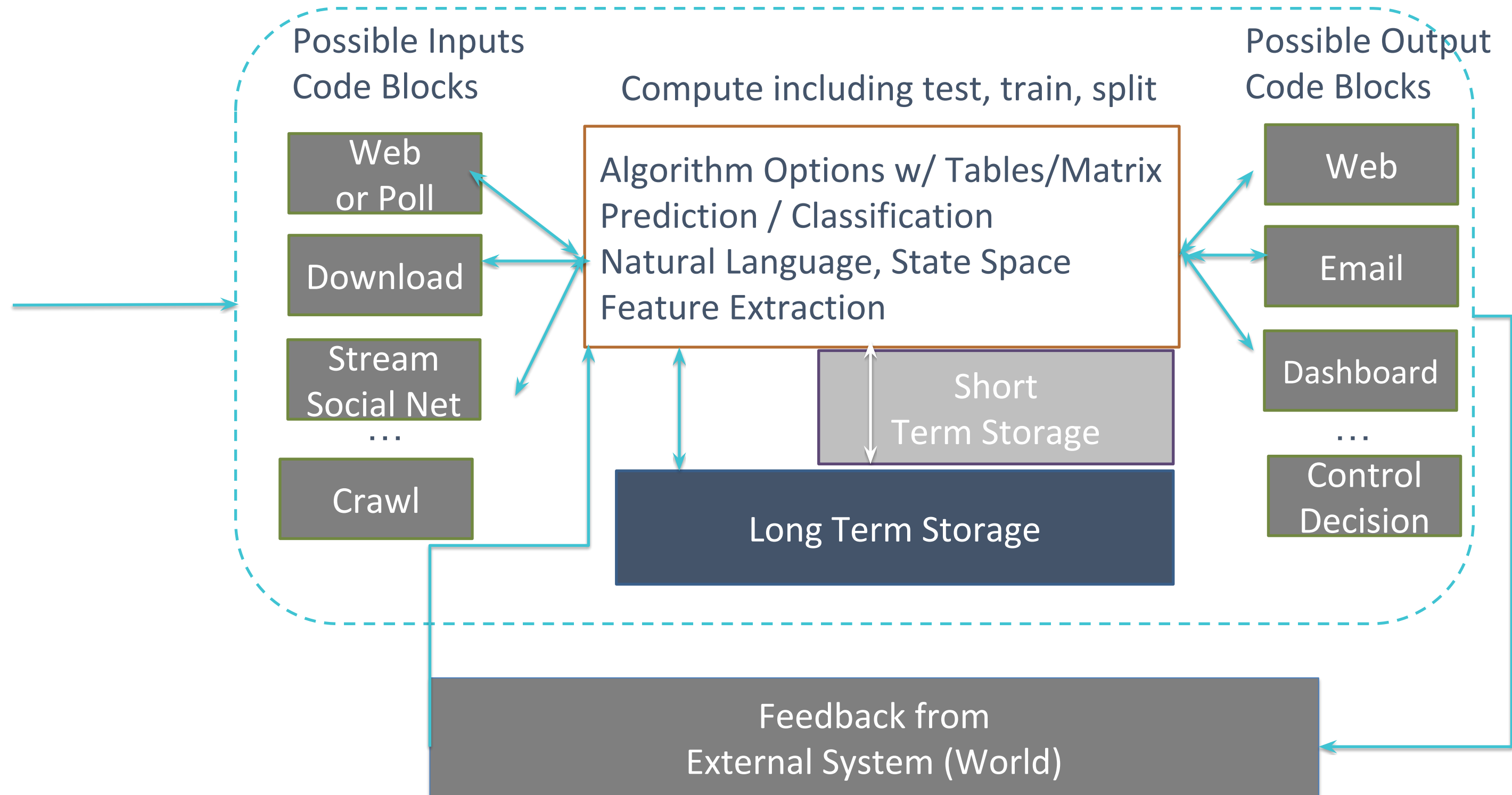
# Technical Components of Project

## Top Components in order of Importance

- UI
  - Ability to send emails
  - Use of Neural Network Algorithms
  - Where will we get the data
  - ..
- 
- Color Code Orange or Red: Lines you need to learn to do
  - Color Code Green: Lines which will be easy to develop

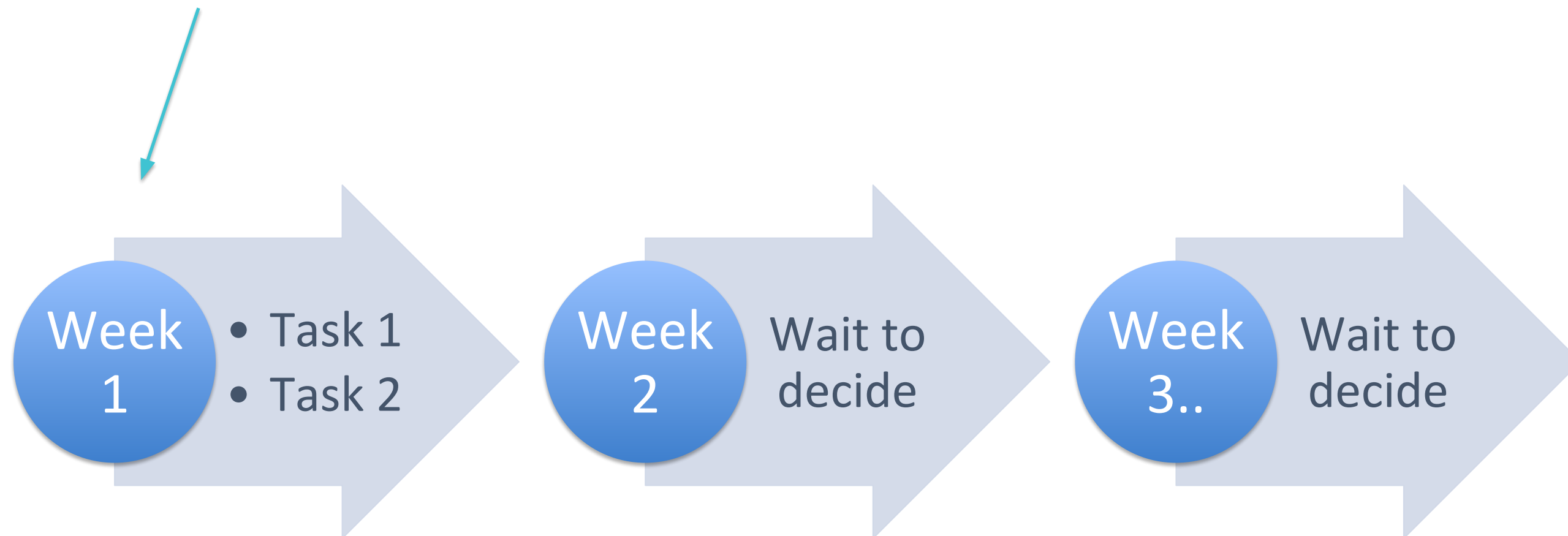
# Sample Architecture (SA)

Replace the diagram below with your SA / DM  
(it is okay to use two slides, one for SA and one for DM)



# Roadmap

- ✗ What is the initial set of tasks (3-5) tasks
- ✗ Put initials or a name next to each



Hint: start with the red items on your list of technical components

Consider Swim Lanes

# Amazon Alexa pitch



# Amazon Alexa Project

The SCET was awarded the Amazon Alexa Innovation Fellowship 2018/2019 to help improve the Alexa ecosystem.

## State of the project:

- Prototype of Berkeley specific Alex skill
- Created Alexa edu material
- Received devices and funding (Echo show and Echo dot)

## Potential Projects:

- Extend Berkeley specific skill
- Develop educational material
- Create a novel product or service utilizing Alexa

## Alexa Innovation Fellowship

Inspiring and enabling student entrepreneurs to enhance their products with Alexa

≡ FORTUNE

Amazon Alexa Fellowship Program Is Heading to 14 More Uni...

The online retail giant said on Wednesday that it plans to award fellowships to graduate students and faculty members at schools like Carnegie Mellon University, Johns Hopkins University, Dartmouth College, and the University of California, Berkeley. The 18 universities marks a major increase from last year, when the Alexa Fellowship program debuted with four schools.



amazon alexa

Email interest to [afo@berkeley.edu](mailto:afo@berkeley.edu)



Thanks!