

# CS 185C Section 2

## Advanced Practical Computing Topics

### Solving Big Data Problems

Fall 2017: Tu/Th 07:30-08:45  
Instructor: James Casaletto

---



- Build sandbox cluster on laptop
- Use distributed file systems
- Write MapReduce programs

Build data pipelines and workflows

Perform data transformations

Store and retrieve data at scale

APACHE  
HBASE

kafka

elastic

hadoop

Spark



Use data science methods

Build clustering models

Perform classification and regression

#### Prerequisites:

- Java programming (CS 146) and ability to build Maven projects
- 64-bit laptop with at least 8GB RAM and 30G disk space