

CS 185C Advanced Topics

Solving Big Data Problems

Spring 2018: 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



Use data science methods

Build clustering models

Perform classification and regression

Prerequisites:

- *Java programming (CS 146) and basic UNIX/Linux command-line skills*
- *64-bit laptop with at least 8GB RAM and 30G disk space*