

LS 88-2: Data Science Connector Sports Analytics

Alex Papanicolaou

Spring 2018

E-mail: apapanicolaou@berkeley.edu

Web: piazza.com/berkeley/spring2018/ls882/home

Class Room: Cory 105

Class Hours: T 12-1:59pm

Office Hours: TBD

Course Description

The principles of data science meet sports analytics. What makes a good hitter in baseball? How do you measure that? What are the flaws of plus/minus in basketball? Do Steph Curry or Klay Thompson ever get a hot hand? When should a coach go for it on 4th down? This course covers a wide range of topics on the analytical thinking behind the data revolution in sports and explores data science through the lens of sports analytics.

Materials

There are no required materials but here are some interesting resources (among many)

- *The Book: Playing The Percentages In Baseball* by Tom Tango, Mitchel Lichtman, and Andrew Dolphin
- *Analyzing Baseball Data with R* by Jim Albert and Max Marchi
- *Mathletics* by Wayne L. Winston
- *Basketball on Paper* Dean Oliver

Prerequisites

This connector course moves quite quickly trying to immerse students in hands-on labs. It is highly recommended this connector be taken after Data 8 because many of the topics will assume a grounding in Python and the knowledge of the topics in Data 8.

Course Objectives

This course will demystify the analytical thinking behind the data revolution in sports and learn data science through a wide range of topics. We will discuss the theory, development, and application of data science analytics in sports. Students will learn about measuring performance, inference, regression modeling, and the idiosyncracies and subtleties of data as well as common pitfalls. At the end of the course, students will have deeper appreciation of how the methods of data science applied to sports have broader applicability and will be well on their way to engaging with data in their own projects, education, or career.

Course Structure

Class Structure

The course will consist of 1 lecture per week that will also feature a lab component with hands-on work. The labs will tie-in with the lectures.

Homework

There will be (near weekly) homeworks that will build on the lectures and the labs. The homeworks will consist of Jupyter Notebooks and will be submitted through a private Dropbox File Request.

Participation

Attend the lab-lectures and participate.

Website

Piazza will be used for disseminating labs and homeworks as well as hosting any discussions. It would also be nice to see a vibrant discussion of related topics that are not exactly within the class but relevant and interesting. The system is highly catered to getting you help fast and efficiently from classmates or the teaching staff. Rather than emailing questions, I encourage you to post your questions on Piazza. If you have any problems or feedback for the developers, email team@piazza.com.

Find the link to the class page above.

Class Project

There will be a group project for the class.

- Groups of size 3-4 are preferred but can allow smaller. No larger.
- Proposals due Week 2-3. The proposal need to succinctly describe the research question, the data, and the methodology.
- Milestones:

1. Data acquisition (~Week 5)
 2. Preliminary results (~Week 8)
 3. Finals results, report, and presentation (~Week 14/End)
- Final product: a 3-5 page report, accompanying code or notebooks, and a short presentation for the class. Presentations will be in the last week.

Grading Policy

Attendance will be graded as checkmark for whether you are present or not.

Homeworks will be graded as check, check plus, or check minus.

- 20%: Attendance
- 40%: Homework
- 60%: Course Project

Course Policies

During Class

Please attend lecture and participate. Phones are prohibited as they are rarely useful for anything in the course. Please refrain from eating in class. Beverages are okay. Food and beverage policy is superseded by that of the lecture room.

Attendance Policy

Attendance is expected in all lab-lectures. Valid excuses for absence will be accepted before class. In extenuating circumstances, valid excuses with proof will be accepted after class. For every class missed the participation grade will be dropped 1 percentage point.

Policies on Incomplete Grades and Late Assignments

Standard Berkeley policy will be followed regarding incomplete grades.

Late assignments will be accepted for no penalty if a valid excuse is communicated to the instructor before the deadline. After the deadline, assignments will be accepted for a 50% deduction to the score up to 2 days after the deadline. After this any assignments handed in will be given 0.

Academic Integrity and Honesty

Students are required to comply with the standard UC Berkeley policy on academic integrity. This shouldn't be an issue but this is here on the rare chance something happens.