 Sports Analytics

Fall 2018

**Instructor**: Michael Lopez

**Office**: Harder 221

**Email:** mlopez1@skidmore.edu

**Office Hours:** By appointment

**Class Meetings:** MW 6:30-7:50

**Course Description:**

There is increasing demand for the understanding and implementation of analytical thinking and tools in the world of sports. This course will cover the development of sports analytics, from discovering methods for novel player and team analysis, to ways to improve in-game decision making, to developing an understanding of referee behavior, and to how the future of big data can transform the world of sports. No prior statistics knowledge is required, though a willingness to code and an interest in sports will help students maximize what they get out of this course.

**Learning Goals**

This course will introduce students to the world of sports analytics. Discussions and material will be drawn from several sports, both professional and amateur, and include extensive reading into the background of how methods for different sports have been developed. Readings will cover topics such as statistics, economics, psychology, business, history, and communication. Students will acquire basic statistical and programming skills. This course is about finding ways to identify problems, formulate productive questions, and to go about answering this questions. Students will demonstrate an ability to:

-Distinguish among, and formulate, types of questions asked within different disciplines (including, as examples, management, player personnel, in-game decision making)

-Read critically, and gather evidence

-distinguish among the evidence and methods appropriate to different disciplines

-Make connections among ideas

-Recognize choices, examine assumptions, and ask questions of your own work and of yourself

-Learn statistical methods for analysis and visualization

*-*Reproducible statistical analysis using *RMarkdown*

-Formulate conclusions based upon evidence. Consider and address complexities and ambiguities.

-Communicate ideas both orally and in writing

-General skill sets: research, presentation, and teamwork skills

-Relate the results of the course to educational goals

**Texts (optional: chapters provided)**

Lewis, Michael. *Moneyball: The art of winning an unfair game*. WW Norton & Company, 2004.

Baumer, Ben, and Zimbalist, Andrew. *The Sabermetric Revolution.* Penn Press, 2014

Diez, David M., Christopher D. Barr, and Mine Cetinkaya-Rundel. *OpenIntro statistics*. Vol. 12. CreateSpace, 2012.

Additional readings will also be posted online ahead of class. Readings to be done before class meetings.

**Grading:**

*Homework (15%)*Homework is the most effective way to reinforce concepts learned in class. There will be weekly homework assignments. Most often, questions will relate to material in the reading that will be covered in class. You are welcome to collaborate with other students, but you must turn in your own work and write up all assignments in your own words. Wherever possible, RMarkdown will be used as to ensure reproducible work and of a standard format.

On the top of every homework, write the names of *everyone* you collaborated with in doing problems. A basic principle of scholarship is that once gives credit to all who contributed to the findings. Copying and pasting sentences, paragraphs, or blocks of *R* code from another student is not acceptable and will receive no credit. All students, staff and faculty are bound by the Skidmore College Honor Code.

Late homework is not accepted. I will drop the lowest homework score.

Homeworks are graded out of 5 points:

1-3 out of 5 points: Most questions attempted, minimal effort

4 of 5 points: All questions attempted, complete effort, graded questions incorrect

4.5 of 5 points: All questions attempted, complete effort, graded questions partially correct

5 of 5 points: All questions attempted, graded questions perfect

*Project Presentation and Technical report [25%]* The major milestone in this class will be conducting (with a group) a statistical investigation on a question of interest to you. For each, you may collect primary data by hand or you can use data available on the Internet. You may also use data that we have used in class, as long as it is for a different project.

You will prepare a project proposal describing your study and obtain approval from me before you begin the investigation. During the last week of class, you (and your group) will give a 10-minute oral presentation of your study. We will spend time in class looking at what data is available on the web and about writing a project proposal. You will write up your report using RMarkdown.

*Exams [25% each]* There are 3 exams, with the final exam scheduled during finals week.

*Class participation (10%)* This includes equal parts attendance, preparation of the readings, bringing an overall positive attitude to class, and a willingness to learn and think creatively about certain topics. Participation is self-judged at the end of the semester

**Fourth hour activities**

I have scheduled a 4th hour for the class meetings: Tuesdays at 11:30. The fourth hour will involve a range of course and college topics, including visits to campus resource areas (library, writing center, career center, etc), and discussions about topics such as substance abuse, academic integrity, etc. Some of these activities will be led by your peer mentor. Additionally, one fourth hour meeting will discuss academic etiquette, how to take effective in-class notes, both of which are critical for transitioning to college. Please refer to the course outline for a tentative schedule of 4th hour actitivies.

**Disability accommodations**: If you need academic accommodation, you must formally request accommodation from Meg Heneger, Coordinator for Students with Disabilities. For further information, call 580-8150, or stop by the office of Student Academic Services (Starbuck Center). Please plan ahead as we will cannot accommodate you without approvals from Academic Services.

**Title IX:** Skidmore seeks to provide an environment that is free of bias, discrimination, and harassment. It is our goal that you feel able to share information about your experiences as a student. If you have been subjected to sexual or gender-based misconduct (such as sexual harassment/misconduct/assault), we encourage you to report it. In deciding to whom you should report such an incident, you should know that if you report this to a faculty member, she or he must notify our college’s Title IX coordinator about the basic facts of the incident. For more information regarding your options, see <http://www.skidmore.edu/sgbm>

**Miscellaneous:**

*Attendance*: Your attendance in class is crucial, as is your punctuality. Accommodations for an unavoidable absence can be made in advance via email; one necessary absence during the semester is not unusual; having more than two is uncommon, and can risk effecting your final grade.

*Collaboration*: Much of this course will operate on a collaborative basis, and you are expected and encouraged to work together with a partner or in small groups to study, complete homework assignments, and prepare for exams. However, every word that you write must be your own. Copying and pasting sentences, paragraphs, or blocks of *R* code from another student is not acceptable and will receive no credit. No interaction with anyone but the instructor is allowed on any exams or quizzes. All students, staff and faculty are bound by the Skidmore College Honor Code. \*\*If you and a friend collaborate on homework questions, please write the names of your collaborators on your turned-in assignment.”

*Computing*: The use of the R statistical environment (downloadable from <http://www.r-project.org/>) with the RStudio interface (downloadable from www.rstudio.org) is thoroughly integrated into the course. A willingness to learn how to code is critical to your enjoyment of the class! Access on or off campus can be found at: <http://r.skidmore.edu>.

*Electronic devices:* Cell phones and other devices are allowed in class meetings (but not exams) only for purposes of accessing material from the course or taking notes. Otherwise, they should not be out. If you will be using a laptop, I ask that you please sit in the back of the class to prevent distracting your classmates

**Calendar (subject to change)**

|  |  |  |
| --- | --- | --- |
| Date | Topics | Readings/HW |
| 9/5 | Class intro, Introduction to R/RMarkdown: Lab 0 |  |
| 9/10 | What is `statistics in sports’? |  |
|  | Data manipulation & spreadsheet basics: Lab 1 | HW 0 |
| 9/17 | Player valuation in baseball |  |
|  | Correlation, R-squared, scatter plots: Lab 2 | HW 1 |
| 9/24 | Modeling of player statistics |  |
|  | Basketball – possessions, shot difficulty: Lab 3\* | HW 2 |
| 10/1 | Shot charts, hot hand theory |  |
|  | Review: Lab 4 | HW 3 |
| 10/8 | *Exam 1* |  |
|  | Why is football analytics so hard? Lab 4 | HW 4 |
| 10/15 | Binary outcome analysis |  |
|  | Expected points and NFL kickers: Lab 5 |  |
| 10/22 | Football - game theory: 4th downs, field goals |  |
|  | Simulate a football game: Lab 6 | HW 5 |
| 10/29 | Hockey – shooting stats, zone entries\*\* |  |
|  | Shooting metrics, repeatability, and salary: Lab 7 | HW 6 |
| 11/5 | Mean reversion & shooting percentages |  |
|  | Shootouts & shooting percentages: Lab 8 | HW 7 |
| 11/12 | *Exam 2* |  |
|  | NWSL exploration: Lab 9 |  |
| 11/19 | Soccer models: expected goals |  |
|  | Expected goal models in soccer: Lab 10 |  |
| 11/26 | Team rankings, Elo, Bradley Terry | HW 8 |
|  | BTM in R, power rankings: Lab 11 |  |
| 12/3 | Referee analytics | HW 9 |
|  | Project work |  |
| 12/10 | Project work |
|  | Project presentations | *Technical Report* |
|  | *Exam 3* |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Confirmed guest speakers:

\*Seth Partnow, Milwaukee Bucks, 9/26

\*\*Namita Nandakumar, Philadelphia Eagles, 10/29

Typical 4th hour activities: note taking, library resources, alcohol and substance abuse, academic requirements, career counseling, diversity/inclusion, anxiety, civic engagement.