 Statistics and Sports MS 251

Fall 2020

**Instructor**: Michael Lopez

**Office**: Basement

**Email:** mlopez1@skidmore.edu

**Twitter:** @StatsbyLopez

**Office Hours:** By appointment

**Class Meetings:** WF 8:40-10:00

**Class Website:** <https://github.com/statsbylopez/StatsSports20/>

**Text (optional):** *Statistical Inference via Data Science (Modern Dive)*: by Ismay and Kim

**Link:** www.moderndive.com

**Course Goals**: Students will:

-Gain an understanding of how statistical tools have been applied in the sports world

-Learn advanced methods for analysis, visualization, and implementation

-Gain familiarity with nonparametric and parametric statistical tools

-Contrast in-sample versus out-of-sample statistics with respect to sports

**Course Skills:**

-Computing skills using *R* & *R-studio*

*-*Reproducible analysis using *RMarkdown*

-Multivariate linear regression, logistic regression, model diagnostics

**Textbook**: Readings will be posted on Github. There is no textbook. See <https://github.com/statsbylopez/StatsSports20/blob/master/Readings.md>

**Computing:** We’ll use R & RMarkdown throughout the semester. Use your laptop, download the R statistical environment (downloadable from <http://www.r-project.org/>) and the RStudio interface (downloadable from [www.rstudio.org](http://www.rstudio.org))

**Writing:** Your ability to communicate results, which may be technical in nature, to your audience, which is likely to be non-technical, is critical to your success as a data analyst. The assignments in this class will place an emphasis on the clarity of your writing.

**Grading:**

**Homework & Participation [15%]** Homework is the most effective way to reinforce concepts learned in class. There will be homework assignments. Most often, questions will relate to material in the reading that will be covered in class. Doing statistics is the best way to learn statistics. 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.

RMarkdown will be used as to ensure reproducible work and of a standard format. Using RMarkdown, *only* HTML or PDF copies will be accepted. These must be emailed to [mlopez1@skidmore.edu](mailto:mlopez1@skidmore.edu) prior to the start of class

On the top of every homework, identify 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

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

Students will be responsible for grading their own homeworks using a solutions key provided on Github

**Project Presentation and Technical report [35%]** The major milestone in this class will be conducting (with a partner or 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, if 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 partner) will give a 10-minute virtual 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 and submit it.

**Exams [50% total]** There will be two take-home exams. These will involve write-ups of a statistical analysis, done using RMarkdown, and submitted to [mlopez1@skidmore.edu](mailto:mlopez1@skidmore.edu). The following scales will be used for each student:

Best exam: 30%

2nd best exam: 20%

**Disability accommodations**: If you are a student with a disability and believe you will need academic accommodation, you must formally request accommodation from Meg Hegener, Coordinator of Student Access Services (mhegener@skidmore.edu). You will also need to provide documentation which verifies the existence of a disability and supports your request. For further information, please call 580-8150 to contact Student Academic Services in Starbuck Center.

**Title IX**: Skidmore College considers sexual and gender-based misconduct to be one of the most serious violations of the values and standards of the College. Unwelcome sexual contact of any form is a violation of students’ personal integrity and their right to a safe environment and therefore violates Skidmore’s values. Sexual and gender-based misconduct is also prohibited by federal and state regulations. Skidmore College faculty are committed to supporting our students and upholding gender equity laws as outlined by Title IX. If a student chooses to confide in a member of Skidmore’s faculty or staff regarding an issue of sexual or gender-based misconduct, that faculty or staff member is obligated to tell Skidmore’s Title IX Coordinator or Title IX Deputy Coordinator. The Title IX Coordinator or Deputy Coordinator will assist the student in connecting with all possible resources for support and options for reporting both on and off campus. Identities and details will be shared only with those who need to know to support the student and to address the situation through the college’s processes. If the student wishes to confide in a confidential resource, the Counseling Center Staff, Health Services, and Victim Advocates (anonymous) are all options available.

More information can be found at the Sexual and Gender-Based Misconduct website or by contacting the Title IX Coordinator, Joel Aure (jaure@skidmore.edu), 580-5708, or Deputy Coordinator for Student Affairs, Gabriela Melillo (gmelillo@skidmore.edu), 580-5022.

**Attendance**: Your attendance in class is crucial, as is your punctuality. We are all going to learn this material together, so we need to have everyone present and working. During the Fall 2020 semester, attendance will not be considered when determining your grade in this course. Students remain responsible for making up any missed work.

**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.

**Calendar (subject to change)**

|  |  |  |
| --- | --- | --- |
| Date | Topics | Assignments |
| 26-Aug | What is `statistics in sports’? |  |
| 28-Aug | Lab: Intro to R/RStudio/RMarkdown | HW 0 |
| 2-Sep | Player valuation in baseball |  |
| 4-Sep | Lab: Correlation, R-squared, scatter plots | HW 1 |
| 9-Sep | Multiple linear regression modeling |  |
| 11-Sep | Lab: Baseball analysis via Lehman database | HW 2 |
| 16-Sep | Linear regression & prediction |  |
| 18-Sep | Lab: Prediction for MLB batter/pitcher stats | HW 3 |
| 23-Sep | NFL - Logistic regression, kickers, expected points |  |
| 25-Sep | Lab: NFL kickers | HW 4 |
| 30-Sep | Logistic regression review |  |
| 2-Oct | *Exam 1 – take home* |  |
| 7-Oct | Football - game theory: 4th downs, field goals |  |
| 9-Oct | Football – Expected points and win probability | HW 5 (proj proposal) |
| 7-Oct | Basketball – possessions, shot difficulty |  |
| 9-Oct | Lab: NBA shot maps & hot-hand (?) | HW 6 |
| 14-Oct | Hockey stats: Stein’s estimator, shrinkage/shooting % |  |
| 16-Oct | Lab: Shrinkage and player prediction | HW 7 |
| 28-Oct | Soccer & expected goals |  |
| 30-Oct | Lab: NWSL games | HW 8 |
| 4-Nov | Team ratings & BTM |  |
| 6-Nov | Lab: Team ratings | HW 9 |
| 11-Nov | TBD |  |
| 13-Nov | *Exam 2 – take home* |  |
| 18-Nov | Project work/speakers |  |
| 20-Nov | Project work/speakers |  |
| N 28-30 | *Study Days* |  |
| D 1-4 | *Exam (project presentation)* |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |