

DATA 367 – Statistical Methods in Sports Analytics

ENGR 304 MW 3-4:15pm – Spring 2026

About the Course: This course will introduce statistical methods and training in statistical consulting aimed to analyze sports by using observational data on players and teams. With an emphasis on statistical inference and modeling, the students will learn how to analyze a sports related problem, utilize statistical tools to find a solution and interpret those results to sports professionals. The course will also offer the opportunity to focus on a semester long sports analytics project.

Course Prerequisites: MATH 129 or MATH 263

It is also recommended that students have experience in a programming language. We will primarily use Python in this class, but no experience with Python is necessary to succeed in the class.

Instructor and Contact Information

Primary Instructor: Jason Aubrey (jaubrey@arizona.edu)

Office hours: Mondays 1-2pm in Math 220, Tuesdays and Wednesdays 1-2pm in Math 219

Graduate Teaching Assistant: Jeffrey Mei (jmei@arizona.edu)

Office Hours: TBA

Our D2L page will serve as the course home page - check it often. Course materials will be published on D2L and important announcements will be made on D2L.

Course Format and Teaching Methods

This class is scheduled to be taught in the in-person modality.

Course Communications

Announcements and important course information may be sent out via official University email or through D2L. It is the student's responsibility to check for messages and announcements regularly. Email should and will be used for notification purposes - however, it is a poor tool for discussion. Mathematical questions should be asked, and will be discussed, in our class meetings - either during class (or after class if time permitting), during office hours or by appointment.

Required Texts or Readings

There is no required text for the course. Supplementary reading material will be provided to students via D2L.

Required or Special Materials

You are required to bring a laptop to every class meeting.

This course will use the programming language Python, Jupyter notebooks, Quarto, git and github, and Positron. You can obtain this software as follows:

- Python - <http://python.org>
- Jupyter - <http://jupyter.org>
- Quarto - <http://quarto.org>
- Positron - <https://positron.posit.co/>
- Git - <http://git-scm.com>

- Github - <http://github.com>

It is your responsibility to have this software installed on the laptop you bring to class. The websites above have documentation and downloads for the most common operating systems, such as MacOSX, Windows, and many Linux distributions. If you run an unusual or niche operating system, you may have to install this software from source.

The software above is what we will use in lecture. Assignments will be turned in as Quarto Markdown (QMD) files often by making a commit to a git repository. If you are an advanced student, in some cases you may use R, and you may use an IDE other than Positron, but we will not be teaching R in this class, and we may not be able to help you with other IDEs.

Equipment and software requirements: For this class you **MUST** bring a laptop to class with a reliable internet signal that can:

- Access D2L
- Run the software above
- Access Gradescope
- Scan and upload written work to Gradescope
- View pdf documents

Class Meetings

Meeting Times: This class will meet on Mondays and Wednesdays from 3-4:15 pm in ENGR 304. Our meetings will give us the opportunity to develop our understanding of the ideas and methods of Sports Analytics. Most days we will utilize group work. Expect to turn in a sample of group work to be graded almost every class meeting.

Class attendance:

- If you feel sick, or if you need to isolate or quarantine based on University protocols, stay home. Except for seeking medical care, avoid contact with others and do not travel.
- Notify your instructor if you will be missing a course meeting or an assignment deadline.
- Non-attendance for any reason does not guarantee an automatic extension of due date or rescheduling of examinations. Please communicate and coordinate any request directly with your instructor.
- If you must miss the equivalent of more than one week of class, you should contact the Dean of Students Office DOS-deanofstudents@email.arizona.edu to share documentation about the challenges you are facing.

Class Recordings: For lecture recordings, which are used at the discretion of the instructor, students must access content in D2L only. Students may not modify content or re-use content for any purpose other than personal educational reasons. All recordings are subject to government and university regulations. Therefore, students accessing unauthorized recordings or using them in a manner inconsistent with UArizona values and educational policies are subject to suspension or civil action.

Absence and Class Participation Policy

Participating in the course and attending lectures and other course events are vital to the learning process. As such, attendance is required at all lectures and discussion section meetings. If you anticipate being absent, are unexpectedly absent, or are unable to participate in class activities, please contact me as soon as possible. Students who miss the first two class meetings, and do not contact me within 24 hours of the second class meeting, may be administratively dropped.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, <https://drc.arizona.edu/>) to establish reasonable accommodations.

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Your responsibilities as a class member

- Be fully engaged in the mathematics, with your peers while in the classroom. This means put aside non-math conversations, texting, social media, and anything else that may make this time less mathematically productive for you and your peers.
- Be ready and willing to participate in many different forms of interactive activities, including small-group discussion, explaining ideas and R-code to others, working out code individually and in a group, and adding/modifying other's solution code.
- Listen to your peers' arguments and the instructor's lead discussion(s) respectfully, politely and engagedly - be willing and ready to contribute whenever appropriate.
- Come to class mentally prepared, so that you (and your peers) may benefit from being in an interactive class.
- Be on time and ready to start right when class is scheduled to start, and remain until the class is dismissed.

Course Objectives and Expected Learning Outcomes

This course will have students utilize statistical tools to solve sports analytics problems, including but not limited to, factors influencing game outcomes and individual performance metrics. This course will begin with an examination of the history of sports analytics and continue on to discuss how to numerically and visually analyze sports related data. The course will also present methods of evaluating team and player performance data using a variety of techniques, including data visualization, regression and hypothesis testing. An emphasis will be placed on learning how to describe outcomes from analysis in a non-technical manner.

Along with the learning outcomes from this course, all students will participate in a project to assist in the data collection and analysis from one of the participating University of Arizona athletic programs. Students will use analysis techniques learned in class to provide expected outcomes based on coach input, while also improving on current techniques in data collection and analysis to provide new insights for the coaching staff to utilize.

By the end of this course, students will be able to:

- Utilize statistical tools to solve sports analytics problems, including but not limited to, factors influencing game outcomes and individual performance metrics.
- Numerically and visually analyze sports related data.
- Utilize methods of evaluating team and player performance data using a variety of techniques, including data visualization, regression and hypothesis testing.
- Describe outcomes from analysis of sports-related data in a non-technical manner.

Assignments and Examinations: Schedule/Due Dates

In Class Group Activities - 50 points - almost every day

Research Paper - 75 points - week 9

Applied Statistical Analysis in Sports Paper - 75 points - week 13

5 Python Assignments - 100 points - weeks 3, 4, 5, 6, 7

Team Sport Presentation 60 points - week 14

Final Results Presentation 80 points - week 17 or 18

Final Results Paper - 50 points - week 18

Reflection on Project - 10 points - week 17

In Class Group Activities

Graded group activities will be assigned during class meetings. Many will involve Python programming - others will involve project updates. According to the calendar, we expect to have 27 graded class activities, each worth 10 points. In general, no make-up activities will be offered. When computing your final class activities grade (out of 50 points) we will take the total number of class activities points earned and divide by 230 and multiply by 50 (to a maximum of 50).

Python Assignments

Graded Python assignments will be assigned in the first half of the course. These assignments will build on class instruction and activities and will be group assignments. It is expected that groups of 3 – 5 students will work together, and deductions may be applied for failing to contribute appropriately to the group work, or not submitting the assignment properly or on time. There will be 5 Python assignments in the semester, each worth 20 points.

Research Paper

The purpose of this paper is to analyze a positional sports paper or article that includes advanced analytics (such as WAR, PER, QBR, Real plus/minus, etc.). It should be a 3 – 5 page paper that includes:

- A summary of the author's thesis and argument.
- A thorough examination of the most important analytical part of the article.
 - A thorough examination (and explanation) of this analysis/metric. This should be very detailed - walk your reader through how to calculate it. If you can get hold of the data used and reconstruct the analysis or the metric, this would be best.
 - What are the positive features of this analysis/metric? How is it an improvement over previous methods/metrics?
 - What are the negative features of this analysis/metric? What improvements need to be made?
 - How does this analysis/metric supports the author's argument? Are there any issues with using this analysis/metric to support their argument?
- A brief description of any other analytics in the paper. Include a brief comment on how the analysis/metric supports the author's argument and if there are any issues with using this analysis/metric.

Applied Statistical Analysis in Sports Paper

The purpose of this paper is to use the techniques learned in Math 367 (or elsewhere) to explore an idea of your own in sports analytics. It should be a 2 – 4 page paper (longer if needed).

- Sport: Pick whichever sport you would like. Give enough information about the sport so that the reader can understand the terms you use your thesis question.
- Thesis: Pick whatever question you would like to explore. Be sure to discuss the potential impact an answer would have on the sport.

- **Data:** It is up to you to find data to work with. Obviously, this may limit the questions you may be able to explore. If you want to explore the impact of going for it on fourth down in college football, and are planning to write a simulation that samples from fourth down plays, then trying to collect all fourth down plays in NCAA Division I football history may be too ambitious. Start with a simulation on a much more limited set, do the appropriate analysis, reach the appropriate conclusion (which would be more limited), and then discuss how one would improve the simulation by increasing the data set in the areas of improvement part of the paper.
- **Analysis:** Be sure to include some advanced analytic:
 - it could be a metric (such as WAR, PER, etc...)
 - it could be a method (such as regression, a simulation, clustering, or hypothesis testing, etc...)
 - Whatever metric or method you choose (and it could be more than one), it should be appropriate for the question.
 - Go through your metric or method thoroughly. The reader should be able to reproduce your results.
 - For example, if you use WAR, it is clear how one calculates WAR (and the reader could calculate WAR for other players by following your work).
 - Use comments in your code to describe what the Python commands you use are doing.
- **Conclusion:** The conclusion you reach should be appropriate to the methods and results you get. (If you do a hypothesis test and do not meet the 5% threshold, then concluding “we do not have enough evidence to conclude that...” is the appropriate conclusion.) Do not overreach. Do not make statements here that are not supported by your work.
- **Areas of Improvement:** Discuss ideas for areas of improvement. This could include expanding your data sets, or starting to collect data that nobody has collected yet, or improvements to a simulation to add more realism, or ideas for new metrics that haven’t been thought of yet, or...
- **Format:** Your paper should be submitted as an Quarto markdown file (QMD). You should also submit any data files necessary to run your Python code. You will lose points on your paper if the Python code does not successfully execute for the grader.

Final Project

A final project, consisting of four parts, will be completed in the last half of the class. The four parts are: the Team Sport Presentation, the Final Results Presentation, the Final Results Paper, and the Reflection on Project. The final projects are group projects, and you will be given significant class time to work on a project - however, you will also need to allocate time outside of class to complete the project. Failure to be present or participate during days allocated to project work, may impact your grade on the Final Results Presentation.

The Team Sport Presentation will be given on the 14th week, which your team will:

- Introduce your sport, and the overall project that you are undertaking.
- Identify the semester goal the team has for this project.
- Build benchmarks for the entire project which specify each task that must be completed
- Discuss the deadline for each task
- Identify work as assigned to each team member

The Final Results Paper will be due May 6th (the last day of class)

- This paper should include all of the relevant work on your project. It should include all the analysis that supports the conclusions in the project presentations. If you think of the presentation as the “highlights” of the project, think of this paper as where the audience goes to find all the details.

Categories	
In Class Group Activities	50 pts (10%)
Research Paper	75pts(15%)
Applied Statistical Analysis in Sports Paper	75 pts (15%)
5 Python Assignments	100 pts (20%)
Reflection	10 pts (2%)
Team Sport Presentation	60 pts (12%)
Final Results Paper	50 pts (10%)
Final Results Presentation	80pts(16%)

Table 1: Total possible points

- Your paper should be submitted as an Quarto markdown file (QMD). You should also submit any data files necessary to run your Python code. You will lose points on your paper if the Python code does not successfully execute for the grader.

Project presentations will be completed the last two weeks of class (including the time scheduled for our final exam, May 12th from 3:30-5:30 pm) and will:

- Discuss the motivation of the project and how it can assist the team for which it was designed
- Detail any significant results
- If you have code, demonstrate how it works
- If you performed analysis, show all results and what they mean (This section should take up the most time and you should translate results for coaches to understand)
- Wrap up your presentation by detailing how current results can be utilized
- Discuss recommendations for future work in this area

The Reflection on Project will be due May 6th. It is a one to two single-spaced typed page document where you are asked to reflect and describe your work and discoveries on the sports analytics project this semester. Be sure to include a description of your contributions and also your thoughts on how or whether this activity helped your professional development or influenced your professionalism in team work and collaboration, communication, and problem solving.

If you have more than two absences during days devoted to project work, your final project grade will be reduced by 10%, and more than four absences on those days your final project grade will be reduced by 20%, and more than six absences on those days, your final project grade will be reduced by 30%.

Please note the following:

- University rules relating to final examinations may be found at: <https://www.registrar.arizona.edu/courses/final-examination-regulations-and-information>
- The University final exam schedule may be found at: <http://www.registrar.arizona.edu/students/courses/final-exams>

Grading Scale and Policies

You will earn a grade of:

- A if you earn at least 450 points (90%)
- B if you earn at least 400 points (80%)
- C if you earn at least 350 points (70%)
- D if you earn at least 300 points (60%)

Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at <http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete> and <http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal> respectively.

You may drop the class without a W through January 28 using UAccess. The class will appear on your UAccess record, but will not appear on your transcript. You may withdraw with a W through April 1 using UAccess. The University allows withdrawals through April 15, but only with the Dean's approval. Late withdrawals are dealt with on a case by case basis, and requests for late withdraw without a valid reason may or may not be granted.

Administrative Drops: Administrative drop is an instructor's option, not an obligation. Instructors are not required to drop students who fail to attend class. Since students may add courses beyond the official start date, instructors should be attentive to student enrollment dates when assessing adequate participation for the purposes of administrative drop. Students may be administratively dropped if they miss the first two class meetings.

Dispute of Grade Policy: In general, any questions regarding the grading of any assignment, quiz, or exam need to be cleared up within one week after the graded item has been returned.

Confidentiality of Student Records <http://www.registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa?topic=ferpa>

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.

Scheduled Topics/Activities

Week	Dates	Topics Covered	Assessments
1	1/14	Course Introduction	
2	1/21	Linear Regression	
3	1/26, 1/28	Linear Regression	Python Assignment #1
4	2/2, 2/4	Logistic Regression	Python Assignment #2
5	2/9, 2/11	Getting Data	Python Assignment #3
6	2/16, 2/18	Manipulating and Processing Data	Python Assignment #4
7	2/23, 2/25	Simulations	Python Assignment #5
8	3/2, 3/4	Bootstrapping Introduction to Data Analytics	Research Paper
	3/9, 3/11	Spring Break!	
9	3/16, 3/18	Hypothesis Testing	
10	3/23, 3/25	Group Meetings	
11	3/30, 4/01	Group Meetings	
12	4/6, 4/8	Group Presentations - Introduce your Sport	Team Sport Presentation
13	4/13, 4/15	Group Presentations	Applied Statistical Analysis in Sports Paper
14	4/20, 4/22	Group Meetings	
15	4/27, 5/29	Group Meetings	
16	5/04, 5/06	Final Presentations	Final Results Presentation & Reflection
17	5/12	Final Presentations from 3:30-5:30pm	Final Results Presentation & Final Results Paper

Additional Resources for Students

UA Academic policies and procedures are available at <http://catalog.arizona.edu/policies> Student Assistance and Advocacy information is available at <http://deanofstudents.arizona.edu/student-assistance/students/student-assistance>

Physical and mental-health challenges: If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call (520) 621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.

University-wide Policies link

The Links to the following UA policies are provided here, <https://academicaffairs.arizona.edu/syllabus-policies>:

- Absence and Class Participation
- Policy Regarding Absences for Any Sincerely Held Religious Belief, Observance or Practice
- Threatening Behavior Policy
- Accessibility and Accommodations
- Code of Academic Integrity
- Nondiscrimination and Anti-Harassment Policy