STAT 7500

Statistical Programming (3 credits)
Fall 2025
Section 001, CRN 29062
Mendel Hall 258
Wednesdays 6:15 – 8:45PM

Instructor
Katie Fitzgerald, PhD Statistics
Assistant Professor of Statistics
kaitlyn.fitzgerald@villanova.edu
Office: SAC 370

Student hours – come talk to me!
Tuesdays 11AM – 12PM
Wednesdays 4 – 5:30PM
Thursdays 1 - 2PM

Course Objectives
Develop statistical programming skills using SAS and R. This includes reading in data in various formats, creating and modifying variables, assigning variable attributes, combining and subsetting datasets, statistical summaries, data visualization, sampling, simulations, and working with strings.

Prerequisite: STAT 7404 (Statistical Methods) or its equivalent is a pre-/co-requisite

Required Course Materials

Textbook: None. We will use a combination of lectures and online resources for the course. There are a few recommended textbooks on the course website.

Software: We will use both R and SAS in this course. Both can be accessed for free online - see the course website for information on how to download them.

Hardware: Students are expected to bring a laptop to all class sessions.

How to access support for this class

- Our course website will be the central hub for accessing course materials, assignments, and announcements: https://kgfitzgerald.github.io/stat-7500/. Blackboard will be used primarily for assignment submissions and grade posting.
- Piazza will be our preferred platform for questions about homework, code, and general course questions. The system is integrated with Blackboard and is highly catered to getting you help quickly and efficiently from classmates and the instructor. Rather than emailing questions to the instructor, you should post your questions on Piazza. Email should be reserved for personal questions such as those regarding absences or grades.
- Come to my student hours! Zoom available upon request.
    - Tuesdays 11AM – 12PM
    - Wednesdays 4 – 5:30PM
    - Thursdays 1 – 2PM
- Contact me about any concerns. Best way to reach me is via email (kaitlyn.fitzgerald@villanova.edu). I do my best to respond within 24 hours Monday – Friday.

Assessment Factors Contributing to Final Grade

Homework

Homework will be assigned approximately weekly and will usually involve coding in SAS or R. Each assignment's instructions and due date will be posted on Blackboard, with most assignments being due on Blackboard by class time on Wednesdays (6:15pm).

Attendance & Engagement

You are expected to contribute to a meaningful learning environment for yourself and your peers. Attendance, participation, and active engagement with the material are expected. A small percentage of your course grade is allocated to the following:

- **Weekly check-ins** – each week you will be asked to fill out a brief survey that provides the opportunity for you to give feedback about what was unclear or difficult about that week's material as well as to self-assess your own engagement with the course.
- **Attendance** – if you miss a class (life happens!), contact a classmate to see what you missed and get the material for that day. Consulting Blackboard is also a good idea. Don't hesitate to reach out if a larger life circumstance is interfering with your ability to engage in the course or if you need additional support to get back on track.

Data Ethics Readings & Community Annotations

We live in an Information Age where data is all around us. Statistical literacy and data competencies can provide us with tools to be better *stewards of information*. Throughout the semester, you will encounter many ways that data and statistics are used (and misused) in our lives and society and how data can be harnessed to unearth knowledge, to illuminate injustices, and to reason critically about uncertainty.

Each week you will engage with an assigned article, book excerpt, or video about applications of statistics and/or data ethics. You will engage with the reading via the community annotation tool Perusall. Annotations are due by class-time on Wednesdays.

Late/makeup work

Here is how deadlines work in the real world: They exist and they're important. However, there's a certain amount of flexibility with them. If you need a little longer on something, you communicate with whoever has set you the deadline and ask if you can have a few more days. This is usually not a big deal, but if it happens a lot, people will start asking you if everything is all right.

That is also how deadlines work in this class. You may communicate with me via an extension request form (available on Blackboard) to ask for an extension on anything you need, and that's mostly fine. If you ask for lots of extensions, we'll work together to find ways to help you keep up with the work in the course.

If there are life circumstances that are having a longer-term impact on your academic performance or well-being, come talk to me, and we can work towards a solution and connect you to the support you need.

Grading

Homework	30%
Attendance & Engagement	5%
Community Annotations	5%
3 Mini-projects	60%

Final letter grades will be assigned according to the scale below

A 93-100%	B+ 87-89%	C+ 77-79%	F 0-69%
A- 90-92%	B 83-86%	C 73-76%	
	B- 80-82%	C- 70-72%	

Important Dates

August 27 (Wed)	First day our class meets
August 31 (Sun)	Add/Drop deadline
Oct 15 (Wed)	NO CLASS due to Fall Break
November 26 (Wed)	NO CLASS due to Thanksgiving break

Course Community & Policies

Inclusive Community

It is my intent that this course models and fosters justice, equity, diversity, and inclusion. We will engage with these values both in content and in practice as we do data science and statistics in community with one another and critically engage with ethical issues in the discipline. You are expected to engage your peers and new perspectives with curiosity, empathy, and intellectual humility. It is my intent that all students be well-served by this course, that your learning needs are met inside and outside the classroom, and that the diversity that you bring to this class be valued and utilized as a resource and strength.

I (like many people) am continually learning how to honor diverse perspectives and identities. If something was said in class (by me or a peer) that made you feel uncomfortable, please let me know. You will also have the opportunity to express concerns anonymously via check-in surveys. Villanova also encourages community members to submit any campus climate concerns at the following website: https://www1.villanova.edu/university/diversity-inclusion/report-climate-concern.html

Villanova affirms that diversity, equity and inclusion are integral components of the teaching and learning experience and an essential element of the ongoing intellectual, social and spiritual development of every member of the Villanova community. We believe that an inclusive community fosters an understanding and appreciation for diversity among our students, faculty, staff and administrators. We are committed to cultivating an academic environment that is marked by genuine curiosity about different perspectives, ardent receptivity to knowledge generated through intercultural connections, and a genuine sensitivity to the variety of human experiences.

Academic Integrity Policy

TL;DR: Don't cheat!
Please abide by the following as you work on assignments in this course:

- You may discuss individual homework assignments with other students; however, you may not directly share (or copy) code or write-ups with other students.
- Citing code/solutions: Unless explicitly stated otherwise, you may make use of online resources for coding on assignments. **However, the work must be primarily your own and may not be completed, in whole or in substantial part, by other humans or chatbots, AI, etc.** If you directly use code from an outside source (or use it as inspiration), you must explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism.

All students are expected to uphold Villanova's Academic Integrity Policy and Code. Any incident of academic dishonesty will be reported to the Graduate Studies Dean of the College of Liberal Arts and Sciences for disciplinary action. You may view the University's Academic Integrity Policy and Code for a

detailed description. If a student is found responsible for an academic integrity violation, which results in a grade penalty, they may not WX the course unless they are approved to WX for significant medical reasons. Students applying for a WX based on significant medical reasons, must submit documentation and their request for an exception will be considered.

Office of Disabilities (ODS) and Learning Support Services (LSS)

If there is any portion of this class that is not accessible to you due to course format or challenges with technology, please let me know so we can make appropriate accommodations.

It is the policy of Villanova to make reasonable academic accommodations for qualified individuals with disabilities. All students who need accommodations should go to Clockwork for Students via myNOVA to complete the Online Intake or to send accommodation letters to professors. Go to the LSS website http://learningsupportservices.villanova.edu or the ADS website https://www1.villanova.edu/university/student-life/ads.html for registration guidelines and instructions. If you have any questions please contact LSS at 610- 519-5176 or learning.support.services@villanova.edu, or ADS at 610-519-3209 or ods@villanova.edu.

Absences for Religious Holidays

Villanova University makes every reasonable effort to allow members of the community to observe their religious holidays, consistent with the University's obligations, responsibilities, and policies. Students who expect to miss a class or assignment due to the observance of a religious holiday should discuss the matter with their professors as soon as possible, normally at least two weeks in advance. Absence from classes or examinations for religious reasons does not relieve students from responsibility for any part of the course work required during the absence.
https://www1.villanova.edu/villanova/provost/resources/student/policies/religious holidays.html

Notice: Students should be aware that Mediasite Lecture Capture is being used to automatically record each class session. Some recordings may be available to students upon request.

Acknowledgements: My teaching practices and course materials incorporate many ideas and resources shared with me by many colleagues at Northwestern University, Azusa Pacific University, Villanova University, and in the Project NExT and statistics education communities. In particular, this document uses inspiration and language from Drs. Sharon McCathern, Mine Çetinkaya-Rundel, Keegan Kang, Yimin Zhang, Jesse Frey, Michael Posner, and Villanova University.

Course Calendar (Tentative)

Week	Date (Wed)	Topics	DUE (Wed, 11:59pm)
1	Aug 27	Course Overview / Introductions Intro to SAS	Week 01 check-in
2	Sept 3	Entering/Reading Data Attributes of Data (Labels/Formats)	SAS HW 01 Annotations 01 Week 02 check-in
3	Sept 10	Attributes of Data (Labels/Formats) Combining Datasets	SAS HW 02 Annotations 02 Week 03 check-in
4	Sept 17	Working with Data	SAS HW 03 Annotations 03 *Hand out SAS Mini Project* Week 04 check-in
5	Sept 24	Common PROCs and output SAS Graphics	SAS HW 04 Annotations 04 Week 05 check-in
6	Oct 1	Intro to R and Quarto	SAS Mini Project Week 06 check-in
7	Oct 8	Working with Data Data Viz I	R HW 01 Annotations 05 Week 07 check-in
	Oct 15	FALL BREAK	
8	Oct 22	Data Viz II	R HW 02 Annotations 06 Week 08 check-in
9	Oct 29	Data Viz III Data Wrangling - Isolating	R HW 03 Annotations 07 Week 09 check-in *Hand out RDV Mini Project*
10	Nov 5	Data Wrangling – Piping Data Wrangling - Deriving	R HW 04 Annotations 08 Week 10 check-in
11	Nov 12	Data Wrangling – Combining	R Data Viz Mini Project Week 11 check-in
12	Nov 19	Data Wrangling – Tidy Data Data Wrangling – Scraping	R HW 05 Annotations 09 Week 12 check-in *Hand out RDW Mini Project*
13	Nov 26	THANKSGIVING – NO CLASS	
14	Dec 3	Working with Text Bootstrapping and Simulation	R HW 06 Annotations 10 Week 14 check-in
15	Dec 10	Additional Topics Parting Thoughts	End of Course Reflection
	Dec 17	NO CLASS	RDW Mini Project

Course schedule, topics, exams and assignments may be changed at the instructor's discretion

AI / LLM policy (i.e. usage of ChatGPT, Gemini, etc.)[1]

TLDR: you're responsible for understanding how to solve problems, cite any use of AI

In general, we treat AI-based assistance, such as ChatGPT, the same way we treat collaboration with other people; you are welcome to talk about your ideas with other people, both inside and outside the classroom, as well as with AI-based assistants.

However, **all work you submit must be primarily your own, and may not be completed, in whole or in substantial part, by other humans or chatbots, AI, etc.** You also **must properly acknowledge (cite) any ideas / code / solutions that did not originate from you.** In all cases, you are responsible for understanding all work that is turned in and may be periodically asked to orally explain your answers.

I expect you *will* use AI / LLMs periodically to assist you in this course. Responsible use of AI is not "against the rules" and you should not feel the need to hide it. **If/when you use AI while working on an assignment, you are expected to provide the following with your submission:**

- A statement acknowledging your use of AI and which tool you used
- A precise description of the prompt(s) you used on which problem(s)
- A brief reflection of your takeaway / analysis of the output provided by the tool, including your level of confidence in it. A couple of sentences will suffice.

Considerations for responsible AI / LLM usage

AI / LLMs are likely to be used in your future workplace and can be an effective tool for the modern statistician / data scientist. However, there are both effective and detrimental ways that LLMs can be used in a learning context. Here are a few things to consider when choosing whether/how to use AI in your coursework:

- AI / LLMs can hallucinate and provide incorrect answers or unintentional code. You must develop your own foundational knowledge of a subject in order to effectively judge and verify whether an LLM's output is trustworthy and code works as intended.
- AI / LLMs use an enormous amount of energy. In order to be climate-conscious, we need discerning use of AI and should be careful not to over-rely on it when other methods (e.g. Google search, non-AI computational tools, human effort) will suffice.
- Employers are interested in people who can, among other things, achieve Task X (with or without LLMs) correctly and efficiently and who can effectively document and communicate *how* they achieved Task X so that it can be verified and reproduced by someone else. This motivates the policy described above.
- To be an effective data scientist, you must "get your hands dirty" with the data and work with it thoroughly in order to know how to properly clean, analyze, and interpret it. Automating too much with AI can shortcut this process and limit your understanding of the data you're analyzing. This can lead to flawed, misleading, or incomplete results and interpretations.
- One of the broader / more existential threats posed by AI is its potential to diminish human connection. Be mindful of how often you are turning to AI for help when you otherwise would be turning to a human. There is value in day-to-day interactions with classmates, tutors, and professors that go beyond efficiently completing an assignment.

Tips for when AI assistance can be especially useful and appropriate

- Debugging code or interpreting error messages

- Clarifying course concepts. For example:
    - "provide me with an intuitive explanation of XYZ"
    - "help me understand why ABC happens in this context"
    - "I'm confused about the difference between ABC and XYZ"
- Deepening understanding of posted solutions
- Generating additional practice questions and step-by-step explanations when studying

(Non-exhaustive) scenarios when AI assistance would be inappropriate

- Copying / typing a homework problem into ChatGPT and having it generate a full solution from start to finish. Homework is intended to build your general problem solving and coding intuition, and you are responsible for coming up with the steps to solve the problem.
- Copying output from ChatGPT directly into your submission. Just as you should not let a classmate write content directly into your submission, so also should you avoid using AI assistance in such a way that directly adds content to your submission.
- Using AI on an open-ended problem or writing assignment that asks for your reflection, opinion, or meta-cognitive thought-processes. It is considered academic irresponsibility to use AI to generate an answer that does not reflect what *you* truly think and believe. I am interested in what you think, not what an LLM thinks.
- Submitting an assignment that has ideas, code, or solutions that originated from AI but is not properly cited. This is plagiarism.

If at any point you are unsure whether a particular use-case of AI is appropriate, please ask!

Notice: Students should be aware that Mediasite Lecture Capture is being used to automatically record each class session. Some recordings may be available to students upon request.

Acknowledgements: My teaching practices and course materials incorporate many ideas and resources shared with me by many colleagues at Northwestern University, Azusa Pacific University, Villanova University, and in the Project NExT and statistics education communities. In particular, this document uses inspiration and language from Drs. Sharon McCathern, Mine Çetinkaya-Rundel, Keegan Kang, Paul Bernhardt, Yimin Zhang, Michael Posner, and Villanova University.