Introduction to Computational Statistics with R

Week 1 Monday

Miles Chen, PhD

Department of Statistics

UCLA

Section 1

Happy New Year! Happy 2026!

Welcome!

- My name is Miles Chen
- You can call me: Miles, Professor Chen, or Dr. Chen, whatever name you are comfortable with. I don't like being called by my last name only (e.g. "Hi Chen!")

Go Over Syllabus

Office hours:

I am bad at email.

Office hours are my preferred method of contact.

Questions and issues are generally resolved much more quickly via office hours.

When you come to office hours, please **introduce yourself**. Do this **every** time you visit me.

I like when students come to office hours with questions about material. I love to explain things and to help students understand.

I like when students come to office hours to tell me more about themselves and to seek counsel about classes to take or next steps. I am happy to make accommodations for students who face difficult circumstances and may need extensions for assignment deadlines. Please do not hesitate to visit office hours.

I am happy to correct grading mistakes. I do not want to get in arguments with students over points. I do not like arguing whether a particular mistake should be a 5 or 10 point deduction.

Section 2

What is this class about?

Foundations for Data Science

Drew Conway's Data Science Venn Diagram consists of

- Coding Skills
- Math & Stats Knowledge
- Domain Expertise

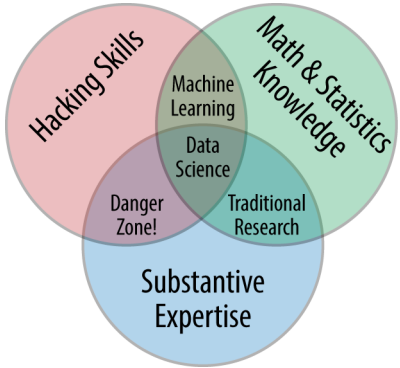

Figure 1: Drew Conway's Data Science Venn diagram

This course

- Stats 102A is an introduction to the field of computational statistics and data science.
- This course aims to prepare you for data science by further developing your coding skills in R and applying those skills to some statistics applications
- The course can be thematically split into two parts.
  - Advancing your programming and coding skills
  - Introducing the methods of computational statistics

    Some parts of the course have heavy overlap with Math 151A. Those weeks are like an "light" version of Numeric Methods that are useful for Stats majors (who aren't required to take Math 151A)

What we'll cover in lecture: Programming and Coding

1. Week 1: R Data structures, subsetting
2. Week 2: Flow control, writing functions, scope, environments
3. Week 3: Web scraping, reshaping data with tidyr
4. Week 4: dplyr, regular expressions
5. Week 5: object oriented programming, ggplot

What we'll cover in lecture: Methods of Computational Statistics

6. Week 6: Midterm Exam, floating point arithmetic
7. Week 7: Root finding
8. Week 8: Numerical optimization, Randomization tests
9. Week 9: Bootstrapping, Statistical Simulation
10. Week 10: kernel density estimation

What you will (need to) develop outside of lecture

I can't "teach" you the following skills. You develop them with practice.

- reading / writing code
- debugging code
- problem solving (decomposing a problem and thinking of a solution using the available coding tools)
- reading documentation

The homework assignments are intended to be structured exercises that will help you develop these skills. Approach the homework intentionally with the goal of developing these skills, rather than 'just trying to get it done.'

Section 3

AI: Artificial Intelligence and Academic Integrity

Artificial Intelligence and Academic Integrity

Some truths:

- AI tools like ChatGPT, Gemini, Claude, etc. are powerful enough to do all of the tasks and homework I assign.
- There's nothing much in this class that truly goes beyond what AI is currently capable of doing.
- Also, if you search hard enough, you can probably find the answers to the homework problems somewhere online.
- However, the goal of this class is not to get you to accomplish some task or even to complete the homework. The goal is that you learn.

No Pain, No Gain

Think of the gym. The goal of lifting weights at the gym is not to lift weights. Lifting weights is a means to the real goal of gaining strength.

"No Pain, No Gain": if your weight training does not result in some muscle soreness, you probably did not exert enough effort to expect muscle gain. Experiencing muscle soreness is a sign that your muscles will go through repairs and get stronger.

Your brain is similar: if your brain does not struggle when writing code, then it has no reason to create additional neuron connections that will improve your abilities as a coder. On the other hand, if your brain struggles with writing code, then your brain will try to create new connections between neurons so the next time will not be as hard. And thus you become a better coder.

No Pain, No Gain

Having the AI solve a difficult assignment (or copying a solution) is like having a stronger person lift the weights that are too heavy for you.

This would a good solution if the goal of lifting weights was to lift the weights. But this does not help for the goal of gaining strength.

Using AI is only good if the sole goal is to finish the task. It does not help towards the goal of challenging you so your brain actually learns.

Course Goals

I believe students resort to AI and plagiarism because they have confused the goals of the course.

Students who use AI believe the goal for them is to get a good grade (or avoid a bad grade) in the class. For these students, the goal of learning is secondary to the goal of getting the desired grade.

But this is wrong! The goal of the course is your learning.

I will admit, a major conflicting issue here is that I am not able to create individualized grading schemas that evaluate exactly how much each student learned over the course. All students are graded on the same criteria and evaluated on what they turn in for the assignments.

That said, I hope you can judge your performance in a class based on what you learned and not your letter grade.

My expectations

When you face a challenging homework assignment, I expect

- you work hard
- you will not seek out solutions from AI or another student or some other resource
- if you are not able to complete everything required by the assignment by the deadline
  - you submit what you have and accept a grade that is less than 100%
  - you view this not as a failure of your coding abilities, but as a indication of areas for growth and improvement

I (and the statistics department) take issues of plagiarism seriously and will escalate cases to the Dean of Students. Full details regarding academic integrity are in the syllabus.

AI Collaboration Policy

Read the course policies on AI and collaboration and be sure you understand it.

You are encouraged to discuss code that is not part of an assignment!

This is a coding class! As long as the code is not part of a homework assignment, you can post and discuss code with each other and also on Campuswire.

You can always post and discuss code that appears in lecture. You are encouraged to modify the examples the appear in lecture and discuss the effect of each change you make.

You can post and discuss code that is for the purpose of learning a particular concept or how a function works.

If AI can do everything in this class ...

Is this class worth anything if AI can do everything?

If AI can do Data Science ... will there be Data Science jobs?

... Yes! This class is worth learning!

... Yes! There will still be Data Science jobs!

AI can't do everything

There's a lot that AI still can't do. AI can handle a lot of basic data science tasks, but can't do more advanced thinking (yet).

Knowledge builds on previous knowledge.

This class helps you learn foundational concepts so you can later learn the more advanced things that AI can't do.

Section 4

Grades and Life

Your grades do not define you

You are here at UCLA. One of the reasons you got into UCLA was because you had good grades in high school and/or at community college. While you are in school, a lot of your energy is poured into your classes and I can understand why grades feel so important. That said,

Your grades do not define you

It feels good to get good grades. Grades do play a role in graduate school admissions. But they are not the most important thing in life. No one on their death bed looks back and says "I wish I got an A- instead of a B+ in that one college class."

Work - Life Balance

I like to split where we put our energies of life into three broad categories:

- Work
  - Jobs and internships
  - School and academics
  - Other professional obligations
- Relationships
  - Family
  - Friends
  - Romantic partner
  - Other social obligations
- Self
  - Care of physical health (food, sleep, exercise)
  - Care of mental health (sleep, play, entertainment)
  - Care of spiritual health (if you are spiritual/religious)

There are 24 hours in a day. It is not possible to give 100% to all categories

Work - Life Balance

Work-Life balance is achieved by consciously choosing what is important to you and devoting your time and energies accordingly.

In general, the more you put in, the more you get out.

Satisfaction can be found by accepting the natural consequences of what you have chosen to deprioritize.

Work - Life Balance

Let's say you are part of a group of friends. Let's say that one day you become involved with a romantic partner.

If you choose to invest all of your relationship hours into your romantic partner, you will likely develop a very strong relationship with your romantic partner. However, because you now invest much less into your original group of friends, those relationships will naturally become more distant. When you see distance forming, it can initially feel hostile. This is not (necessarily) the result of your friends being angry that you have a romantic partner but the natural consequence of having less time to spend with them.

As people, we have to make a choice about what is important to us.

When you accept the natural consequences of investing less time into something, you can reduce your own feelings of bitterness and jealousy.

Work - Life Balance

In the corporate and professional world, people who devote a lot of energy into the goals of the company are rewarded. The company is not necessarily punishing people who choose to have families and a life outside of work.

From the company's perspective, who would they rather promote?

- the person who did everything asked of them and then continued to stay at work and did even more
- the person who did everything asked of them and then immediately left to spend time with their family/friends/romantic partner

You have to choose what is important to you. If climbing the ranks within the company is more important, then you will spend your time accordingly. If spending time with your family/friends/romantic partner is more important, then spend your time accordingly.

Self care is important

You must not neglect taking care of your physical and mental health.

If you neglect care of self, you will likely operate at less than 100% efficiency and the time you invest in work/school/relationships will not be as productive.

Examples:

- You don't get enough sleep. A friend invites you out. You choose to accept your friend's invitation instead of sleep, but you are a bit 'out of it' and are a drag to hang around. Maybe it would have been better to decline your friend's invitation and get sleep.
- Exams are coming up. You choose to skip a meal and minimize sleep to study. You end up getting sick. Your performance on the exam suffers. Maybe it would have been better to eat properly, sleep well, and study a bit less.

Self care is important

When I tell you that your physical and mental health is important, I'm encouraging you to choose to invest your time into activities like exercise, sleep, and relaxation that will boost your physical and mental health.

Sometimes this means choosing not to complete your homework to 100%.

The natural consequence of this is a homework grade that is less than 100%.

When you can readily embrace this natural consequence of prioritizing your own physical and mental health over your homework grade, you can enjoy the quarter with less bitterness, more joy and better health.

Beware of “fruitless” entertainment

Entertainment and fun activities are important for your mental well being. It’s important to have fun.

I love hanging out with people I like, watching TV, movies, sports, playing board games, video games, going on hikes, browsing the Internet, reading a book, listening to podcasts, etc.

Participating in an entertainment activity should be a break from work and should give you mental energy so you can return to your work in a good mood.

Some activities, and even hanging out with certain people, can have the opposite effect - they drain you. Some video games, apps, and social media sites are designed to be addictive - giving your brain immediate dopamine pleasure hits while you use them so you play round after round or continue scrolling forever (and keep coming back) ... but after spending hours of doing the activity, you don’t feel good about yourself.

Be mindful and selective about your entertainment activities.

Today is day one

Today is day one. The first homework assignment has been posted, but don't feel like you have to finish it today. You should open it and take a look at it. Maybe even attempt some of it, but don't worry if you can't finish everything right now. We will cover content relevant to the homework in the coming days.