

# STAC33: Introduction to Applied Statistics

## Course outline

# This course

- STAC33, Introduction to Applied Statistics
- Online:
  - 2-3 hours of recorded lectures / week, asynchronous (watch when you like)
  - 1 hour of live tutorial / week, Monday morning, on Zoom
  - “office hours” also on Zoom, by appointment
- Instructor:
  - Ken Butler, [ken.butler@utoronto.ca](mailto:ken.butler@utoronto.ca)
  - I like sharing what I know about applied statistics and the use of R
  - This course will help you develop the skills to use software to analyze data, and to communicate the results.
  - Effective communication of results is key, since in applied work you are always trying to convince someone of what you found
  - I want you to demonstrate your understanding and insight about the data you're working with

# Prerequisites

- You *need* a second Statistics course that covers statistical inference from a mathematical viewpoint, one of these:
  - STAB57
  - STA 248
  - STA 261
- I check prerequisites.
- You need STAB57 for the Statistics Minor/Major/Specialist programs.

# Getting help

- Option 1: come to tutorial and ask the TA.
- Option 2: Quercus discussion board:
  - If it's a question about course procedures, look in this course outline first.
  - Then: look at the Quercus discussion forum. Your question might have an answer there.
  - If you see a question there that you can answer, go ahead and answer it.
  - If you don't see an answer to your question, post your question on the discussion forum.
  - For an issue that is more personal, email me (below).
- Option 3: I am also willing to meet at other times on Zoom if you want to chat. I will set up a mechanism for you to make an appointment.

# E-mail

- Use a U of T email (so that I know who you are)
- Use a suitable subject, eg “STAC33”
- Begin with “Dear Dr Butler”
- State *clearly* how I can help you. (The quicker I can figure this out, the better your chances of a quick reply.)
- Include your name (as U of T knows it) *and* student number
- I aim to reply within one working day, but with so many of you in the course, it may take longer.

# Course objectives

- You will learn how to use software to analyze data, and to communicate what you found.
- This will be assessed through assignments (approximately weekly).
  - Some questions will ask for software output, and some for explanation.
  - The grading will reward evidence of clear understanding and insight.
- Along the way you will learn how to use R Notebooks and how to turn these into reports.

# Course materials

The first thing on each line is a link:

- Quercus: lecture videos, hand in assignments/projects, receive grades.
- Course website: course news, reminders, lecture notes and code, assignment solutions, other stuff I want to share with you
- Problems and Solutions in Applied Statistics: what it says; PASIAS for short.
- Computer requirements
  - to get the most out of tutorials, you will want a working camera and microphone
  - can use `rstudio.cloud` (needs only web browser) or install R and R Studio on your computer. Instructions in lecture notes.
- Reference: R for Data Science

# Assessment

- Assignments: weekly.
  - 10 total, due in weeks 3 through 12. Best 8 count (12.5% each).
  - available on Tuesdays, due on Thursdays at 11:59pm
  - based on the previous week's recorded lectures
  - you have a certain amount of time (eg 3 hours) between the time you open an assignment and the time you need to have handed it in.
  - these assignments are like little take-home exams, so *do not ask for or expect to get help on the assignments themselves*. The time to get help is at tutorial or by working through the practice problems that I will attach to each recorded lecture.



# Grading

- Assignments and projects will be graded out of 10 on this scale:
  - 6/10 or less: Falls short of expectations for this course, with serious errors in coding or absent or deficient explanations.
  - 7/10: As expected in this course. May be small coding errors or misconceptions.
  - 8/10: Good work, no obvious coding errors, mostly accurate explanations.
  - 9/10: Excellent work, showing clear understanding throughout.
  - 10/10 (rare): Excellent work that shows additional insight.
- Graders will also give feedback on how you can improve for the next assignment.
- Grades are the grader's best assessment of the quality of the work, and are unlikely to change on appeal.
- My (detailed) solutions will be available after each assignment is handed in.

# Policies

- I do not generally accept late work, because I will want to post my solutions for everyone to learn from.
- I am willing to be a little flexible (at my discretion), but work that is handed in after my solutions have been posted will definitely not be graded.
- You have two “free” assignments (since only the best 8 count); you may miss these without explanation. If circumstances make it impossible to hand in work beyond that, email me and we’ll see what we can work out, *before* the assignments in question are due.
- There are no make-up assessments.
- Things may change during the semester. I reserve the right to make changes to this course outline as we proceed, including taking class votes if needed.

# Course material

- Installation / connection for R
- Reading data files (of different sorts)
- Making graphs
- Numerical summaries
- Statistical inference
- Reports
- Tidying and organizing data
- Case studies
- Functions
- Dates and times
- Miscellaneous tasks
- Vector and matrix algebra in R
- The bootstrap
- Bayesian statistics with Stan

# Accessibility statement

*Students with diverse learning styles and needs are welcome in this course. In particular, if you have a disability/health consideration that may require accommodations, please feel free to approach me and/or the AccessAbility Services Office as soon as possible. I will work with you and AccessAbility Services to ensure you can achieve your learning goals in this course. Enquiries are confidential. The UTSC AccessAbility Services staff (located in S302) are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations: (416) 287-7560 or by e-mail at [ability@utsc.utoronto.ca](mailto:ability@utsc.utoronto.ca).*

# Expectations 1/3

- I expect that you will do your own work in this course, and not seek help from outside this course. Expect consequences if you do not. (The Chair's Designate for Academic Integrity is currently investigating an incident from my Fall course.)
- All the ideas you need to succeed in this course are contained in the lecture notes, with a few exceptions that I will tell you about. Seeking help from outside the course may confuse you more than it helps.

## Expectations 2/3

- You may freely use anything that you learn from the lecture notes, PASIAS, solutions to previous assignments, or from conversation with the TAs or me.
- Anything from outside the course, for example, material you find on the Internet, **must be cited**. You can do this by giving the URL that it came from, for example “according to <https://r4ds.had.co.nz/data-import.html>”, and then say whatever it is you learned from there. If you learned it in another course, say which course, but be aware that things may be done differently in other courses (see first point on next page).
- If you want credit for any ideas that come from outside the course materials, you *must* say where they came from.
- Citation is a way of making sure that people get credit for ideas that are theirs; not citing when you should is a way of erasing someone's point of view.

# Expectations 3/3

- I expect you to learn how to do things *as they are done in this course*.
- If you tell me you're having trouble, I will not judge you or think less of you. I hope you'll extend me the same grace. (From here.)
- This (video).
- I am doing my best to give you an opportunity to learn something that will help you in your future. I would like you to take that opportunity.