STAT:5400 Computing in Statistics
Fall 2022

Lectures **MW 5:00 PM – 6:15 PM** 71 SH

Instructor	Boxiang Wang, 261 SH, `boxiang-wang@uiowa.edu`
Teaching Assistant	Yikai Zhang, `yikai-zhang@uiowa.edu`
Department	Statistics & Actuarial Science, 241 SH. `http://www.stat.uiowa.edu`
DEO	Professor Kung-Sik Chan, DEO, 319-335-0712, `kung-sik-chan@uiowa.edu`

Course Website.
Announcements, homework problems, lecture notes, etc: `http://icon.uiowa.edu`.

Course Schedule.
`https://docs.google.com/spreadsheets/d/1JZCKP984yb8NTPdMhXiYIR50VXrS7Vkiz0OCLKbJeLU/edit#gid=0`

Course Description and Objectives.
This is a required course for master students in statistics. Through hands-on experience with real problems, we will learn computing skills essential in applied statistics and in research in methodological and theoretical statistics.

This course is bilingual: coding is primarily in R and secondarily in Python.

The course mainly consists of two components.

- Component 1: Introduction to Statistical Computing Methods. Topics include generating random variables, Monte Carlo studies, resampling methods (Bootstrap, jackknife, and permutation tests), computing linear models, computations in high-dimensional statistics, optimization algorithms (e.g., bisection search, Newton's method, gradient descent), and machine learning algorithms.
- Component 2: Tech Guide. Topics include the Linux operating system, R Markdown, LATEX (mathematical document preparation language), SAS, GitHub (version control), Sublime (text editor), ggplot2 R package, C and Fortran, writing R packages, parallel computing, among other topics.

Prerequisites.
CS:1210 and STAT:3200 and (STAT:3120 or STAT:3101 or STAT:4101).

Corequisites.
STAT:5100 and STAT:5200 if not already completed.

Textbook.
No required textbook.

Optional Reference:

- Maria L. Rizzo (2008). *Statistical Computing with R.* Chapman & Hall/CRC.
- Trevor Hastie, Robert Tibshirani, and Jerome H. Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction.* The 2nd Edition, New York: Springer.
  `https://hastie.su.domains/ElemStatLearn/printings/ESLII_print12.pdf`
- Norman Matloff (2011). *The Art of R Programming: A Tour of Statistical Software Design.* No Starch Press.
- Hadley Wickham and Grolemund Garrett (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data.* O'Reilly Media, Inc.
  `https://r4ds.had.co.nz/`
- Rafael A. Irizarry (2019). *Introduction to Data Science.* Chapman & Hall/CRC.
  `https://rafalab.github.io/dsbook/`

Grading.
Your semester grade will consist of the following components:

Homework	30%	
Midterm Exam	30%	Oct 26
Instructional Project	15%	
Computing Project	25%	
Total	100%	

Grading will be on a curve, with +/− grade used. A grade of A+ represents exceptional work and is given only in extraordinary situations.

Homework.
In general, weekly homework will be assigned each Wednesday and will be due in class the following Wednesday. Exceptions to this schedule will be announced in class. Homework should be submitted electronically through the ICON submission tools for this course (`https://icon.uiowa.edu`). Show your work when solving written homework problems. Complete code and output must be submitted for computer problems. Always document your answer.

You are encouraged to study with others. However, if you work with others on homework, please: a) write up your own assignment and make sure you completely understand all solutions that you submit, and b) write the names of the others in your study group on your assignment.

Late homework is accepted only as required by the university policy, i.e. due to "illness, mandatory religious obligations, or other unavoidable circumstances or University activities." This is a strict rule; however, you are allowed to drop the lowest homework grade.

Office Hours.
In person: 3:30 PM – 5:00 PM Wednesday
Online: 1:00 PM – 2:30 PM Thursday
Reserve your slot using the following google sheet and send me a zoom link before the scheduled time.
`https://docs.google.com/spreadsheets/d/1I4dUIErEgweMItzSmVooVJXYZ857zImzQb84qae3dMQ/edit#gid=0`
You may use your first name or simply write "Reserved" to reserve a time slot. You can reserve consecutive slots if you want to request more than 15 minutes.

Proficiency Exam.
The Incoming Statistics Graduate Students Computer Programming Proficiency Test is designed to determine whether you have already met the programming proficiency requirement. Students who are not deemed proficient will be required to do an extra programming practice homework in this class. The exam result will not affect the grade of this course. The exam results of statistics graduate students will be reported to DGS of the department.

The exam will be held online. You are allowed to use any resource but you cannot communicate with others about the exam. You may use your choice of the following programming languages, R (preferred), Python, C, C++, and Fortran.

The proficiency exam will be posted at noon on Aug 23 and will be due by 11:59 PM Aug 25. No late submission is accepted.

Midterm Exam.
There is a 75-minute in-class hands-on computing midterm exam on 5:00 PM Oct 26, 2022 in 71 SH. The computing exams are open book and open notes. You are not allowed to discuss with others about the exam.

Students must take exams as scheduled except in cases of officially university-approved absence such as class conflict with official exam time, illness, religious observance, and NCAA athletic competition.

Makeup exams are not available for other reasons, including student org field trips, club competitions, job interviews, and personal events.

A student who is absent for more than five days or miss an exam due to an emergency or an illness may request the UI Service Center, 17 Calvin Hall, registrar@uiowa.edu, to

notify me of the reason for the absence.

More teaching policies and resources are seen in:
`https://clas.uiowa.edu/faculty/teaching-policies-resources-examination-policies`.

Instructional Project.
Students will work in groups of three to four to study and present a topic that is not covered in this class. You are expected to give one half lecture to introduce the topic to other students. You are free to use any resource, including some tutorials available online, but you must cite them clearly. You are encouraged to design your lecture in the format of a lab, allowing students to have hands-on experience of the tools that you introduce. You are also required to make a short homework question based on your material.

Possible topics include: R packages shiny, caret; Hadoop, Julia, web scraping.

The project is collaborative. Each student on a team is expected to complete a similar amount of work and to contribute equally to the final project. Students who misrepresent themselves as equal partners in this collaborative project but who are actually letting others do the bulk of the work will be reported to the College for academic dishonesty.

Computing Project.
Students will work in groups of three or four to carry out projects involving application of the statistical computing methods covered in the course to problems of their own choosing. Examples of possible types of projects are:

- Design and carry out a simulation study to compare the properties of two or more statistical procedures.
- Analyze a complicated or large-scale real data.
- Choose a research paper. Use simulation and find data to verify results in the paper.

The computing project will be carried out in two stages. Please meet with me while you are working on each stage.

- Project proposal (due 5:00 PM Nov 7. Grade is 30% off if submission is late within one less day, otherwise no grade is given.)

  Each team should write a two-page summary. This is a detailed description of what you plan to do, including questions to be addressed, software to be used, and methods to be applied.

- Project presentation.

  Each team will give a thirty-minute presentation during the week of Dec 5. A list of which tasks each team member has performed should be submitted.

It is encouraged to work with different people for the instructional and computing projects.

Extra Bonus 1.
You are encourage to attend the Travelers Modeling Competition, which will be held in late Oct and concluded in mid Nov, and is in form of a Kaggle competition. The modeling competition is a great opportunity to bring you hands-on modeling experience and give them exposure to a real-world case. You will be given an opportunity to present your work to Travelers employees if your performance is among the top. Without top performance you will be given an opportunity to present your work in class anyway. You will have extra 3% - 5% bonus on the FINAL GRADE based on your participation and in-class presentation, not on your model performance.

Extra Bonus 2.
Extra 1% bonus on the FINAL GRADE will be considered if you attend our department colloquia regularly.

Attendance.
Attendance is required. See instruction for Absence from Class due to illness, emergence, or extenuating circumstances in `https://registrar.uiowa.edu/absence-class` .

See COVID-related questions in `https://coronavirus.uiowa.edu/`.

A student who is absent for more than five days or miss an exam due to an emergency or an illness may request the UI Service Center, 17 Calvin Hall, registrar@uiowa.edu, to notify me of the reason for the absence.

Copyright.
The lectures, slides, homework, quiz and exam questions are for use only by students attending STAT:5400 (0001). It is illegal to share with anyone else the course materials without instructor's permission.

"Incomplete" Grades A grade of "Incomplete" will only be given under extraordinarily extenuating circumstances that prevent the student from completing the course requirements. Having a failing grade in the course is not an extraordinarily extenuating circumstance.

Academic Honesty and Misconduct All students in CLAS courses are expected to abide by the CLAS Code of Academic Honesty.

Student Complaints Students with a complaint about a grade or a related matter should first discuss the situation with the instructor, and finally with the Director or Chair of the school, department, or program offering the course.
Undergraduate students should contact CLAS Undergraduate Programs for support

when the matter is not resolved at the previous level. Graduate students should contact the CLAS Associate Dean for Graduate Education and Outreach and Engagement when additional support is needed.

Drop Deadline for this Course You may drop an individual course before the deadline; after this deadline you will need collegiate approval. You can look up the drop deadline for this course here. When you drop a course, a "W" will appear on your transcript. The mark of "W" is a neutral mark that does not affect your GPA. Directions for adding or dropping a course and other registration changes can be found on the Registrar's website. Undergraduate students can find policies on dropping and withdrawing here. Graduate students should adhere to the academic deadlines and policies set by the Graduate College.

University Policies
Accommodations for Students with Disabilities
Basic Needs and Support for Students
Classroom Expectations
Exam Make-up Owing to Absence
Free Speech and Expression
Mental Health
Military Service Obligations
Non-discrimination
Religious Holy Days
Sexual Harassment/Misconduct and Supportive Measures
Sharing of Class Recordings