

STATS 504: Practice and Communication in Applied Statistics

Octavio Mesner, Ph.D.

Fall 2021

E-mail: omesner@umich.edu

Class Time: MW 8:30am – 10am EST

Location: 1324 EH

GSI: Pramit Das (pramitd@umich.edu)

Office Hour: M 10am (in person after class), Tu 10am (zoom)

Meeting Dates: Aug 30, 2021 – Dec 10, 2021

Canvas: <https://umich.instructure.com/courses/458613>

Zoom link: <https://umich.zoom.us/j/93116796505> Passcode: 504

Piazza: piazza.com/umich/fall2021/stats504

Course Description

This course provides students with hands-on experience using a variety of techniques from modern applied statistics through case studies involving data drawn from various fields. Lectures provide background on case studies, along with reviews of relevant methodology. Students then conduct independent data analyses for each case study and produce written reports. Evaluation is based on attaining insight from the data, effective communication of findings, and appropriate use of statistical methodology, as well as active participation in class discussions.

Overview

The goal of this course is to provide Masters level students with hands-on experience using a variety of techniques from modern applied statistics. Most course material is presented through case studies involving data drawn from various fields. Lectures will provide background about each case study along with discussion of relevant methodologies. Students will then conduct independent data analyses and produce brief written reports. Evaluation will be based on attaining insight from the data, effective communication of findings, and appropriate use of statistical methodology, as shown in the written reports.

Participation in class discussions is an essential part of the class. Regular attendance and active participation is expected from all students.

Major themes:

- Formulating meaningful and tractable questions based on research goals and consistent with available data
- Devising, documenting, and implementing analysis strategies

- Communicating findings
- Leveraging knowledge of statistical foundations and theory when engaging in applied research
- Understanding the capabilities and limitations of statistical methods
- Understanding the value and limitations of different types of data
- Interpretation of analytic results
- Developing data manipulation and computing skills, especially for large and complex data sets
- Building knowledge about several applied research areas where data-driven investigation plays a major role

Prerequisites

Students are expected to have mastered the essential foundations of statistics at the undergraduate and masters levels, and should possess a solid understanding of ideas such as sampling, variation, bias, and uncertainty. Students should have substantial prior exposure to core statistical methods including regression and multivariate analysis.

Coursework

Students will conduct independent analyses of datasets throughout the semester and write about their findings. Datasets, reading materials, and writing prompts will be provided by the instructor. There is no textbook or other materials to purchase for the course.

Grading

- 95% of the course grade will be writing assignments, 5% will be class participation.
- Participation will be assessed on productive contributions and questions during class.
- Writing assignments will be due roughly every 2 weeks, and will be submitted on canvas.
- There are no exams for this class.
- Writing assignments will generally be short (2 to 3 pages). See below for guidelines on what we are expecting in the writing assignments.
- The instructors or GSI will read and evaluate every assignment, providing individual feedback. Assignments will be graded on a 100 point scale.
- All assignments will count equally toward 95% of the final grade. Since assignments are submitted electronically and are announced at least 7 days in advance, late assignments will generally not be accepted. Exceptions to this policy may be made at the instructors discretion, but only in cases of serious and unanticipated personal emergencies.

- All writing is to be done individually, and should primarily reflect each students own ideas. Students are welcome and encouraged to discuss statistical methods, coding strategies, and topics relating to the data and motivating questions for each assignment.
- Plagiarism, including copying any material prepared by another person, copying from other students (with or without permission), or allowing other students to write all or part of your assignment, will be handled strictly following U-M policies regarding academic integrity.

Computing

You are free to use whatever computing tools you choose, but it must be clearly documented for someone not familiar with the language you are using. Code will be assessed on clarity and reproducibility. It is generally recommended to explain the thought process underlying the code when documenting. The instructors will mainly use R or Python when providing code for illustration.

In addition to statistical analyses, substantial data manipulations may be required. If you are using R, you will likely want to use the `data.table` or `dplyr` libraries. If you are using Python, you will likely want to use the Pandas library. However these are only suggestions and you are free to use whatever software and libraries you choose.

The code used to produce the analyses discussed in the course lectures will be available on the course Github site here: <https://github.com/omesner/UMSTATS504>.

Expectations for student writing

This is both an applied statistics course and a technical communications course. The coursework will consist of writing and data analysis. All statisticians must frequently communicate and document their findings in written form. Writing about your analytic plans and research findings, and writing reviews and critiques of other peoples writing are all excellent ways to organize your thoughts, strengthen your arguments, and identify weaknesses in your claims.

For most of the writing assignments in this course, you should imagine that you are writing a memo or email to be read by collaborators or colleagues. Your audience consists of people familiar with the data and scientific (or industrial) context behind the data, but not necessary about statistics.

Below are some guidelines for writing in this course. These are a few of the most important things to keep in mind. We will expand on this a lot during the semester.

- Write in an appropriate academic or business tone. Do not write informally or casually, but also avoid excessively formal language.
- Organize your content so that your writing has a focused message, with each paragraph contributing distinctly to communicating this message.
- Use simple, direct language. Avoid convoluted expressions, hidden or needlessly subtle meanings, unusual vocabulary, and digressions that do not contribute to your main message.
- Express your arguments in as plain and simple terms as is practical. Do not write in a way that requires the reader to re-read your writing multiple times to understand your point. Favor short sentences in the active voice with one clause, and paragraphs that focus on one topic and do not extend for more than half a page.

- Write for an audience that may include non-native English speakers. Avoid colloquialisms and obscure cultural references.
- You are expected to be able to write in grammatically-correct English, appropriate for graduate level coursework. It is understood that there are many non-native English speakers in the course, and occasional minor grammatical issues will be overlooked.

Student Well-Being

Students may experience stressors that can impact both their academic experience and their personal well-being. These may include academic pressure and challenges associated with relationships, mental health, alcohol or other drugs, identities, finances, etc.

If you are experiencing concerns, seeking help is a courageous thing to do for yourself and those who care about you. If the source of your stressors is academic, please contact me so that we can find solutions together. For personal concerns, U-M offers many resources, some of which are listed at Resources for Student Well-being on the Well-being for U-M Students website. You can also search for additional resources on that website.