STT 843: MULTIVARIATE ANALYSIS

Spring 2026 SYLLABUS

CONTACT INFORMATION

INSTRUCTOR: Dr. Guanqun Cao   **OFFICE:** C426 Wells Hall
E-mail: caoguanq@msu.edu   **Phone:** 517-353-7154
TA: Mr. Andrews Boahen
E-mail: boahenan@msu.edu

CLASS TIME & PLACE:

MW 8 : 30 am - 9 : 50 am, Wells Hall A136.

OFFICE HOURS:

M: 10 : 00 am - 11 : 00 am or by appointment.

TEXTBOOK:

1. *Applied Multivariate Statistical Analysis*, 6th edition, by Johnson and Wichern, Prentice Hall (ISBN: 9780131877153).

SUPPLEMENTARY TEXTBOOK AND USEFUL SOURCES:

1. *Applied Multivariate Statistics with R*, by Daniel Zelterman, Springer (ISBN 978-3-319- 14092-6, ISBN 978-3-319-14093-3 (eBook))

2. *Aspects of Multivariate Statistical Theory*, 2nd edition, by Robb J. Muirhead, Wiley- Interscience (ISBN: 9780471769859)

PREREQUISITES:

Students MUST have a working knowledge of probability (STT 441 or STT 861), and statistical inference (STT 442 or STT 862) before enrolling in this course. Experience with Linear Algebra (MTH 415) is also highly recommended. Students are expected to be familiar with basic level of matrix algebra and R programming.

WEBSITE:

There are three places you need to access to get content for this course:
1. Content for the class will be posted on the course webpage: `https://caoguanqun.github.io/STT843/`
Nearly every class will have a skeleton of slides, as well as a Rmarkdown for coding portions.

2. The D2L website has your grades kept as up to date as possible. Announcements will also be facilitated on D2L.

3. We use Crowdmark (`https://crowdmark.com/`) for grading. You will upload your homeworks here, and you will also have access to graded quizzes and tests from this website. Once we have the first homework assignment, you will get an invitation from the website to be able to upload your stuff.

COURSE DESCRIPTION:

This course will introduce statistical methodologies based on multivariate analysis: matrix methods in multivariate statistical set-up, multivariate normal distribution, tests of hypotheses for means and covariance matrices, multivariate analysis of variance, discriminant analysis, principal components and factor analysis. All class computing and or programming will be done in R. We will discuss using R for the methods introduced in the course, including several examples with different data sets.

COURSE GOALS:

Course develops students' abilities to apply regression model building methods in a wide variety of application environments.

COURSE ORGANIZATION:

During Class: No cell phone use in call. Phone should be in silence mode. Courtesy is extremely important in maintaining effective learning environment. Please be respectful of the individual rights to learn of your colleagues.

COURSE POLICIES:

Attendance: You are strongly encouraged to attend classes on a regular basis. In the event that you miss class it is solely your responsibility to obtain class notes, assignments and announcements. Furthermore, I intend to keep an informal record of attendance. I reserve the right use attendance and class participation to determine borderline grades.

Reading Ahead: I expect you all to read the material ahead of time, before class if assigned. This will help you to follow lectures and ask questions.

Homework: Homework will be assigned tentatively every two week (posted on the website or in lecture). Homeworks are due on the days marked in the schedule at midnight. Homeworks submitted before midnight the following day will still be graded with a 5% penalty. Homeworks submitted before midnight two days after will still be graded with a 15% penalty. After midnight 2 days after, no credit will be given, but see the dropped grades policy below. For example, if homework is due on Friday, those turned in before midnight on Saturday have a 5% penalty; those turned in before midnight Sunday will have a 15% penalty; and no homework will be accepted after that. While late homework will not be accepted, please see the missed work policy section below for informations on automatically dropped grades. The material presented draws from many disparate backgrounds, and it is highly unlikely any one student will have experience with all of it. To that end, we encourage students to work together, discuss the problems, and teach each other while struggling together. Our collaboration policy is as follows.

We do assume you will talk to each other to work on things.

We do assume you will google definitions while you are working on things.

We do assume you will include an acknowledgement section in your homework mentioning the people and resources you used in the course of answering problems. *Example: I worked with Maryam Mirzakhani and Grigori Perelman while completing this assignment. I also used wikipedia to understand Galois cohomology and stack overflow to get me started on an algorithm for enumerating Hamiltonian cycles in a graph.*

We do not assume you will copy each other's work or copy from the internet.

Project: There will be a final project where you will take a publicly available data set, use tools learned in class to analyze the data, and submit a written report of your results. This project will be due at the end of the semester, but we will have several checkpoints throughout the semester.

Exam: There will be one in-class exam during the course. Acceptable resources for the exam are a calculator for numerical calculations and one-page cheat sheet. No other resources are acceptable. The tentative exam date is: **March 11, 2026** (In-Class and closed book : subject to change)

Quizzes: Once per three weeks, there will be a short in class quiz on basics from the previous week's course. It will be at the beginning/end of class and will last for about 15 minutes. The sorts of things I will put on quizzes include checking understanding definitions and interpretations of the content to ensure that we are all at the same place with the material. Note that no makeups will be given, see the missed work policy below.

A student who misses an examination because of a medical reason must provide documented evidence of medical incapacitation to the professor. Other reasons for missing an examination must be discussed with the professor as soon as possible before the day of the examination and each case will be considered on an individual basis. A student who misses an examination without a valid excuse will receive a grade of zero on that examination. Please plan ahead and make sure that you can attend the exams.

Missed Work Policy At the end of the semester, I will drop the two lowest homework grades; and the two lowest quiz grades. For this reason, there is no late homework accepted and no makeups are available for quizzes. Students with an issue (i.e. medical, bereavement) needing to miss either the midterm or the final exam need to get in touch with Dr. Cao as soon as possible, preferably before the exam is administered. A time to make up the exam will be organized with the student.

FINAL GRADES

Final Grades will be based on:

Mid-term Exam: 40% (**March 11, 2026**, subject to change)

Quizzes: 20% (drop one lowest grade)

Homework assignments: 20% (drop one lowest grade)

Project: 20%

Grading

Grading is guided by the table below:

Grade	% Points
4.0	≥ 90
3.5	≥ 85
3.0	≥ 80
2.5	≥ 75
2.0	≥ 70
1.5	≥ 65
1.0	≥ 60
0.0	< 60

Questions about Grading

Questions about grading must be asked within one week of the graded work's return.

Course Activities

There may be times when we cannot meet in person, e.g., when the instructor is sick. We will post an announcement to D2L about the change in location as soon as we know that it needs to change.

Required Materials

You will need to ensure that you have the following:

A laptop to bring to class. We will be coding examples of topics nearly every day, so be sure to bring it with you. In the case that you do not have access to a laptop to bring to class every day, please speak to me as soon as possible so we can figure out alternative accommodations.

You need to have access to the course website before every class, as we will be posting skeleton slides and RMarkdown prior to class. These should be posted before noon the day of class. If you want the ability to take notes on the slides, you need to be sure to either have access to a tablet, or print out the slides in advance of class. We will not be providing print-outs of the slide decks.

Acceptable and Unacceptable Use of Generative AI

This course is all about the use of machine learning, and so understanding the potential for use of generative AI tools such as ChatGPT, Dall-e, etc., is directly applicable to our learning objectives. That being said, the extremely fast emergence of this tool means that we are all still learning about how it can be used, and working to understand the use of this tool in an ethical manner, particularly in our classroom context. In class, we will draft a list of examples of how we as a class think that that these tools can and cannot be used for the purposes of this course. Your use of AI tools must be properly documented and cited in order to stay within university policies on academic integrity and the Spartan Code of Honor Academic Pledge. Clearly indicate where generative AI has been used in your assignments and provide proper citations for the AI tools and models you utilized, along with the prompts you used to generate content. This can be done through comments in your code, annotations in your documentation, or a dedicated section in your homework or project reports. Properly citing both the tools and the specific prompts for each piece of generated content helps ensure transparency, reproducibility, and gives credit to the creators of the AI tools. Any assignment that is found to have used generative AI tools in unauthorized ways is subject to being given a 0. When in doubt about permitted usage, please ask for clarification.

ACADEMIC MISCONDUCT:

The Department of Statistics and Probability adheres to the policies of academic honesty as specified in the General Student Regulations 1.0, Protection of Scholarships and Grades, and in the All-University of Integrity of Scholarship and Grades which are included in Spartan Life: Student Handbook and Resource Guide.

ADDRESSING ISSUES OF DIFFERING ABILITIES:

To arrange for accommodation a student should contact the Resource Center for People with Disabilities at (517)353-9642 or http://www.rcpd.msu.edu/

EMAIL CORRESPONDENCE:

In order to protect your privacy, all course e-mail correspondence must be done through a valid MSU account.

Any of the above is subject to change. Such changes, if any, will be announced in class. Please note that it is your responsibility to attend classes and keep track of the proceedings.