BIA-652-A Multivariate Data Analytics

Monday | 6:30 PM - 9:00 PM | North Building 105

Instructor

Dr. Feng Mai
Associate Professor of Information Systems & Analytics, School of Business
Email: fmai@stevens.edu Please include **BIA652** in the email subject
Teaching Assistants: Muyao Hu mhu12@stevens.edu; Zihan Chen zchen61@stevens.edu

Office Hours

Professor: Monday 4:30 - 6:15 PM
TA: TBD
Please book office hours appointment using the link: https://bit.ly/3EN916V. You can book an in-office appointment or Zoom appointment

Course Description

This course introduces basic theory and methods underlying multivariate analytics. Students will study techniques used for regression, classification, dimension reduction, and clustering. They will build expertise in applying these techniques to real data through class exercises. This proficiency will enable students to become sophisticated data analysts, and to help make more informed design, marketing, and business decisions. Python will be the programming language used for the course.

Prerequisite

Calculus (e.g., derivative and integration) and Linear Algebra (e.g., vector and matrix operation)

Learning Goals

After taking this course, the student will be able to:

- ❑ understand the probability behind basic statistical models
- ❑ use Python to analyze multivariate data
- ❑ think critically about data and research findings

Textbooks

- Practical Multivariate Analysis. (6th Edition) Afifi, A., May, S., & Clark, V. A. ISBN: 9781032088471

The following books are optional references:

- Mathematical Statistics and Data Analysis. Author: John A. Rice, ISBN: 9780534399429
- Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python 2nd Edition. Author: Peter Bruce, Andrew Bruce, Peter Gedeck, ISBN: 149207294X
- An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics) 1st ed. 2013, Corr. 7th printing 2017 Edition. Authors: Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, ISBN: 1461471370. Free Download from Author's Website

Grading

Graduate students need a grade of **C or above** to pass the course.

- **DataCamp Assignments: 50%:** Due next week before class starts; they usually take 3-4 hours to complete.
- **Weekly Quiz: 10%:** There is a multiple-choice quiz due next week before class starts; 4-5 questions each week. Most are conceptual questions from readings and the lecture.
- **Take home final project: 30%:** You will be given a dataset to analyze and several research questions to answer at the end of the semester. You will be asked to write a report.
- **Attendance: 10%**
- **Late Policy:** If you are not able to complete the DataCamp assignment on time, you can still complete later for half credit.
- The lowest scored quiz and the lowest scored assignment will be dropped when computing your final grade.

Academic Integrity

Graduate Student Code of Academic Integrity

All Stevens graduate students promise to be fully truthful and avoid dishonesty, fraud, misrepresentation, and deceit of any type in relation to their academic work. A student's submission of work for academic credit indicates that the work is the student's own. All outside assistance must be acknowledged. Any student who violates this code or who knowingly assists another student in violating this code shall be subject to discipline.

All graduate students are bound to the Graduate Student Code of Academic Integrity by enrollment in graduate coursework at Stevens. It is the responsibility of each graduate student to understand and adhere to the Graduate Student Code of Academic Integrity. More information including types of violations, the process for handling perceived violations, and types of sanctions can be found at www.stevens.edu/provost/graduate-academics.

Consequences of academic impropriety are severe, including receiving an **"F" for the course** (not just for the assignment), a warning from the Dean of the Graduate School which becomes a part of the permanent student record, or expulsion.

Uploading to, downloading from, or any unauthorized usage of course material sharing sites such as Course Hero, Chegg, GitHub, etc. are considered as violations of academic integrity. If you used any material from these sites in your assignment submission, you would get an automatic F for the course (not just for the assignment) and be reported the dean of graduate studies.

Learning Accommodations

Stevens Institute of Technology is dedicated to providing appropriate accommodations to students with documented disabilities. Student Counseling and Disability Services works with undergraduate and graduate students with learning disabilities, attention deficit-hyperactivity disorders, physical disabilities, sensory impairments, and psychiatric disorders in order to help students achieve their academic and personal potential. They facilitate equal access to the educational programs and opportunities offered at Stevens and coordinate reasonable accommodations for eligible students. These services are designed to encourage independence and self-advocacy with support from SCDS staff. The SCDS staff will facilitate the provision of accommodations on a case-by-case basis. These academic accommodations are provided at no cost to the student.

Student Disability Files are kept separate from academic files and are stored in a secure location within the office of Student Counseling, Psychological & Disability Services. The Family Educational Rights Privacy Act (FERPA, 20 U.S.C. 1232g; 34CFR, Part 99) regulates disclosure of disability documentation and records maintained by Stevens Disability Services. According to this act, prior written consent by the student is required before our Disability Services office may release disability documentation or records to anyone. An exception is made in unusual circumstances, such as the case of health and safety emergencies. For more information about Disability Services and the process to receive accommodations, visit https://www.stevens.edu/sit/counseling/disability-services

Inclusivity statement

Stevens Institute of Technology believes that diversity and inclusiveness are essential to excellence in education and innovation. Our community represents a rich variety of backgrounds, experiences, demographics and perspectives and Stevens is committed to fostering a learning environment where every individual is respected and engaged. To facilitate a dynamic and inclusive educational experience, we ask all members of the community to:

- be open to the perspectives of others
- appreciate the uniqueness their colleagues
- take advantage of the opportunity to learn from each other
- exchange experiences, values, and beliefs
- communicate in a respectful manner
- be aware of individuals who are marginalized and involve them
- keep confidential discussions private

Course Schedule

Week	Date	Topic	DataCamp Assignments
1	9/12	Probability and Random Variables	Introduction to Python
2	9/19	Lab 1: Pandas and Dataframe Manipulation in Python	Intermediate Python
3	9/26	Random Variables II	Data Manipulation with pandas
4	10/3	Estimation	Statistical Thinking in Python (Part 1)
5	10/11 (Tue)	Bivariate Random Variables	Joining Data with pandas (Ch 1-2)
6	10/17	Linear Regression (Video Lecture)	Introduction to Regression with statsmodels in Python (Ch 1-3)
7	10/24	Variable Selection and Model Comparison	
8	10/31	Lab 2: Regression using Python	Intermediate Regression with statsmodels in Python (Ch 1-3)
9	11/7	Classification: Logistic Regression and Discriminant Analysis	Introduction & Intermediate Regression with statsmodels in Python (Ch 4 – both course)
10	11/14	Dimension Reduction	
11	11/21	Lab 3: Dimension Reduction and Classification using Python	Dimensionality Reduction in Python (Ch 1 and 4)
12	11/28	Clustering	Cluster Analysis in Python (Ch 1-3)
13	12/5	Factor Analysis and Other Regression	Model Validation in Python
14	12/12	Bayesian Inference	