CSC 333: Computational Data Science

Spring 2025

Instructor:	Jared Adam McBride	**Time:**	Sec. 1 TR 9:30 – 10:45
Office:	111 Magnolia Hall		Sec. 2 TR 12:30 – 1:45
Email:	jared.mcbride@svu.edu		Sec. 3 MW 11:00 – 12:15
Phone:	(480) 695-5627	**Place:**	Von Canon Library Computer Lab

"We could compare learning to . . . the way a pianist prepares for a concert. A pianist cannot cram his knowledge of music and his playing skill into one week or even one month of solid practice to prepare for a concert. But he prepares himself through consistent diligent practice, day by day over a long period of time."

—Joseph B. Wirthlin
(*Finding Peace in Our Lives*, p. 75)

Student Hours: Student hours are just office hours with a better name. They are held at these times:

Mon, Wed, Fri: 12:30 PM – 2:30 PM Tues, Thur: 7:30 AM – 9:00 AM

or by appointment. Student hours will be held in my office or the Magnolia Conference room (right next to my office), which I have reserved for these times. Usually during these times I will be in the conference room (109 Magnolia Hall). If I am not, I will likely be in my office (again, next door to the conference room). Please, feel free to come and study. No appointment is necessary. If you wish to discuss something privately, it likely would be better for you to make an appointment for some time other than student hours. However, you are welcome to try your luck and come to student hours and see if I can meet with you privately then.

My tentative weekly schedule is as follows:

Time	Monday	Tuesday	Wednesday	Thursday	Friday
7:00 AM		Student Hours		Student Hours	
8:00 AM		Student Hours		Student Hours	
9:00 AM	MAT 115	MAT 221	MAT 115	MAT 221	MAT 115
10:00 AM	CSC 333		CSC 333		CSC 333
11:00 AM	MAT 221		MAT 221		
12:00 PM		MAT 221		MAT 221	
1:00 PM	Student Hours		Student Hours		Student Hours
2:00 PM	Student Hours		Student Hours		Student Hours
3:00 PM					
4:00 PM					

Introduction: This course is uniquely designed to match the needs and resources available at Southern Virginia University. Currently, this is the only computer science course is data science. So, this course contains not only fundamentals of data science, which could well be a course in itself, it contains two theoretical technologies that are fairly common and which I have used quite a bit in my data-driven modeling work. These are Principal Component Analysis (PCA) and Least Squares Regression. Then the course finishes with a brief introduction to Natural Language Processing (NLP). So, I think of this course as being the cream skimmed from the top of at least three mainstream courses (courses like: Fundamentals of data science, Linear Algebra for data scientists, Statistical Machine Learning, and Natural Language Processing).

This is a new course, and you are all now test subjects as well as codevelopers. I welcome any input you may have about the topics and the pace of the course.

Text Books: There will be no textbook for this course. All the material will from official or unofficial Python documentation or class notes that will be given to you.

Required Materials:

- Access to a computer. Preferably one with a python installation.

Course Description: This course introduces the fundamental concepts and practices of data science. Students will explore the full data science lifecycle, from data collection and cleaning to analysis, visualization, and communication, with exposure to current tools and technologies used in data science. Students will also develop core skills in programming, statistics, and data wrangling.

Prerequisites: MAT 221: Statistics.

Class Policy:

- Regular attendance is essential and expected.
- It is suggested to spend at least **2 hours** outside of class for each hour in class in order to understand the material. To achieve excellence, more time may be required. So, please budget at least 6 hours each week (in addition to class time) to spend learning this material. Note that if you are taking 15 credits that comes to about 45 hours of work a week.
- Cell phone usage during class is very distracting (to yourself as well as others) and is strongly discouraged. Unless you are expecting an emergency call, keep your cell phone on silent. Please do yourself a favor and allow yourself to focus by not using your phone.
- Class time is active learning time. During the times we meet, I expect you to be ready to engage your brain. This is not a lecture series in which you can sit and listen passively. My goal will be to get each of you to exercise your minds each day in class.

General Instructional Format: This course will be taught in the style of a task-oriented *flipped classroom*. For a good explanation of a flipped classroom as used here, see this video. Of course, for a more general explanation see Wikipedia here. Basically, a flipped classroom is one in which the students study assigned material before class, then the class time is devoted to practicing and solidifying their understanding of the material. The assignments will be reading guides, homework assignments, tasks, projects, and presentations.

Grading Scheme:

Read the Docs	20%
Homework	10%
Tasks	25%
Projects	9%
Presentations	21%
Final Project/Presentation	15%
Total	100%

(The percentages guarantee at least the following letter grades.)

A	93%	C+	77%
A-	90%	C	73%
B+	87%	C-	70%
B	83%	D+	67%
B-	80%	D	63%

Read the Docs (20%): Before we discuss a topic in class, you will be expected to read a portion of the Python documentation on the topic. To do this you will be given a reading guide which will direct you in what to read and

emphasis certain information. These are to be completed and submitted prior to the class in which the topic will be discussed. We will not be following a set schedule because the pace of the class will be determined by in part the students. so make sure you are coming to class and checking all communications from me to know when those assignments are due.

Homework (10%): These will be posted on Canvas with the associated due date. They will be graded on completion and correctness. These will take the form of Python Notebooks that you will fill in or work sheets be be completed and submitted as a PDF.

Tasks (25%): Tasks consist of a specified task to be completed and a report to be written. The report will be submitted on Canvas. These will also be in the for of a Python Notebook.

Projects (9%): Projects are like big tasks, but they are more open-ended and will require more time and effort. The project must be approved by me. Associated with each project will be a presentation. There will be 2 projects in the semester.

Presentations (21%): You will be required to give a presentation on your project. The presentation will be graded on content, delivery, clarity, and accuracy.

Final Project/Presentation (15%): This is very much like a project, but it is the final project. It will be even more involved and will require even more time and effort than projects. The project must be approved by me. Associated with the final project will be a presentation. Final presentations will be given during the final exam period, which is Wednesday, May 7, 2025, 9:00 – 11:50 am.

Late Work: If you need to turn in an assignment late, you must contact me before the due date. Generally, late work will not be accepted without prior approval.

Very tentative schedule of topics:

Week	Topic	Assignment
1	Syllabus, What is Data Science?	What is Data Science?
2	IPython, Jupyter Notebooks, Pandas	RTD: Pandas
3	Git, Regex, os	RTD: Git, Regex, os
4	Web Scaraping and Data visualization	RTD: Requests, BeutifulSoup, MatPlotLib, Seaborn
5	Project Work	Project 1
6	Project Work	Presentation
7	Probability and Linear Algebra for Data Science	Course Notes
8	Multiple Regression	Course Notes
9	Regularized Multiple Regression	Course Notes
10	PCA	Course Notes
11	Spring Break	Enjoy!
12	Project Work	Project 1
13	Project Work	Presentation
14	NLP	RTD: Scipy
15	Final Project Work	Final Project
16	Final Project Work	Final Project
17	Final Project Work	Final Presentations

Core Learning Goals (CLG): Students completing Southern Virginia's core curriculum are expected to attain the

following college-level competencies:

- *Careful Writing:* Ability to present arguments in clear written English with proper spelling, grammar, vocabulary, and transitions.
- *Critical Thinking:* Ability to separate a conclusion from its premises and evaluate how well the premises support the conclusion.
- *Analytic Reasoning:* Ability to evaluate a body of evidence and draw appropriate conclusions.

Core Values: Students demonstrate:

- *Scholarship* by coming prepared to class, focusing on activities in class, and meticulously completing homework assignments.
- *Discipleship* by being respectful of others and showing courtesy and concern for others and the learning environment.
- *Accountability* by faithfully attending class and coming on time, coming prepared, being aware of assignments given and completing them on time.
- *Enthusiasm* by participating in class discussions and completing in-class activities, showing interest in cultural material and discussions about mathematics in general.
- *Refinement* by going above and beyond in living the dress and grooming standards, their manner of speech and actions, cleaning up the classroom and following proper classroom etiquette. This includes not using cell phones (texting, etc.) or other electronic devices that may disrupt learning in the class.

Student Learning Outcomes (SLO) for the Mathematics Major: A student successfully completing this major will;

1. Competently work with the concepts of calculus, differential equations, geometry, analysis, algebra, and statistics. (Foundation-Formulate)
2. Learn to read, understand, analyze, and produce proofs at increasing depth. (Mathematical Proofs-Prove)
3. Use appropriate technology to enhance mathematical thinking and understanding, solve mathematical problems, and judge the reasonableness of results. (Technology-Solve)
4. Exhibit knowledge of formulating and solving problems, interpreting solutions, and modeling techniques central to applications of mathematics. (Mathematical Modeling -Model)
5. Demonstrate the ability to effectively communicate mathematics and other quantitative ideas in written and oral forms. (Communication-Communicate)

Course Goals: After completion of the course you should be able to do the following:

- Become proficient with some tools and technologies used in data science, including Python, Jupyter Notebooks, and Pandas.
- Become capable of learning new tools and technologies in the field of data science, specifically though documentation.
- Become familiar with data the visualization tools matplotlib and seaborn.
- Become conversant in the theory and implementation of PCA and least squares regression.
- Become conversant with some tools of natural language processing.

Title IX: Southern Virginia University is committed to providing an environment that promotes personal integrity, civility, and mutual respect and is free of all forms of sex discrimination and gender-based violence, including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence, and stalking.

Title IX of the Education Amendments of 1972 states that "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance." Title IX applies to all employees and students, and covers discrimination in programs, admissions, and activities.

You can report any concerns to the Title IX Coordinator (titleix@svu.edu), your professor, the Wellness Center, or any university administrator. Please be aware that all instructors are mandated reporters, so by law, we must disclose details relayed to us about any possible Title IX infraction to the University's Title IX coordinator.

SVU's Title IX can be found at https://svu.edu/titleix/.

Accommodations: Southern Virginia University is committed to providing a learning environment that accommodates and supports students with disabilities through accommodations. The Office of Accessibility Services oversees Academic Accommodations. Students with any type of physical, medical, psychological, or emotional disability may be eligible - through the Americans with Disabilities Act (ADA) - to receive accommodations to provide equal access to academic settings. Any student seeking accommodations should contact the Office of Accessibility Services by emailing accessibility@svu.edu. Students are advised to take care of the necessary documentation and paperwork early in the semester. Accommodations are not made until the student follows the procedures outlined on the website, (https://catalog.svu.edu/reasonable-accommodations/) receives an official accommodation letter AND discusses it with their professors.

Academic Honesty: Cheating, plagiarism, and other forms of dishonesty are against the code of conduct (https://svu.edu/campus-life/code-of-conduct/) and will be reported to the provost and office of the registrar. Students have the right to appeal decisions on academic honesty (https://catalog.svu.edu/academic-honesty/) to the appropriate division chair following the Student Grievance Procedure (https://catalog.svu.edu/student-grievance-procedure/).