

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
UNIVERSITY OF TENNESSEE - KNOXVILLE

DSE 511 - Computational Methods and Data Science
Fall 2025 Syllabus (Draft)

Time and Location:

Tuesday & Thursday 4:05 pm - 5:20 pm, Min Kao Engineering 406

Instructor:

- Dr. Scott Emrich
 - Office: 608 Min Kao; 865-974-3891; semrich@utk.edu.
 - Tentative office hours: Mon 2:00pm-3:00pm, and by appointment.
 - If my office door is open, you are welcome to come in and ask questions.

An Invitation

This course is designed to support PhD students from diverse backgrounds as they begin to use data science in their research. My goal is to give you the computational fluency, confidence, and habits necessary for future research. We will blend foundational skills with hands-on, research-relevant applications.

Expect a balance of programming, collaboration, and discussion. No matter your starting point, you will leave this course better able to build reproducible workflows, analyze data, and communicate your results.

Course Webpage: <https://web.eecs.utk.edu/~semrich/dse511/>

Short Course Description:

This course will start the journey to equipping PhD students with the computational thinking, data fluency, and collaborative development skills essential for modern data-driven research. It blends foundational computer science concepts, applied statistics, and core data science workflows to foster a deep understanding of how to design and implement reproducible, scalable computational frameworks.

Designed for students with varied programming backgrounds, DSE 511 emphasizes real-world application, cross-disciplinary problem-solving.

Prerequisites:

- Graduate standing
- Basic statistics recommended
- Some programming experience (any language) helpful but not required

Course Outcomes

By the end of the course, you will be able to:

1. Apply computational reasoning to real-world research problems
2. Design and evaluate algorithms and data pipelines
3. Collaborate using modern software engineering tools (e.g., Git)
4. Build reproducible, transparent, and ethical data workflows
5. Communicate technical findings clearly to diverse audiences
6. Analyze, visualize, and interpret research data using appropriate tools
7. Stay current with evolving trends in data science, including Responsible AI

Major Topics & Approximate Hours

- **Weeks 1–2:** Computational thinking & programming foundations (pseudocode, Python)
- **Weeks 3–4:** Data structures & algorithmic thinking (lists, dictionaries, searching/sorting, recursion)
- **Weeks 5–6:** Data wrangling & transformation (pandas, NumPy, dplyr, EDA workflows)
- **Weeks 7–8:** Collaborative software development (Git/GitHub, modular design, testing)
- **Weeks 9–10:** Statistics for data science (inference, modeling, simulation)
- **Weeks 11–12:** Machine learning foundations (regression, clustering, evaluation)
- **Weeks 13–14:** Data visualization (principles, dashboards in Python/R)
- **Week 15:** Responsible data science (reproducibility, FAIR data, bias, fairness, explainability)

(Topic timing may shift based on class progress.)

Tools and Languages

- **Primary:** Python
- **Comparative:** R (selected modules)
- Git/GitHub, Jupyter, Docker, SQL
- Visualization: ggplot2
- Students are required to either bring or work on domain-specific datasets/tools

Grading Policy

Assignments are due before midnight on the posted date. Late submissions are accepted with a **10% per day penalty** (pro-rated for the first day).

Grade Weighting (tentative):

- 20% Exercises & labs (programming drills, algorithm design, data wrangling challenges)
- 25% Collaborative projects (Git-based teamwork on software tasks)
- 20% Systems design project (planning a reproducible/scalable workflow)
- 25% Capstone project (end-to-end analysis on research-relevant data)
- 10% Technical communication (documentation, peer feedback, short presentations)

Letter Grades:

A: 95–100% | A-: 92–94.9% | B+: 88–91.9% | B: 85–87.9% | B-: 82–84.9% |
C+: 78–81.9% | C: 75–77.9% | C-: 72–74.9% | D: 65–71.9% | D-: 62–64.9% | F: <62%

Class Participation

Engagement in assigned work, class discussions, and peer code reviews will be a critical part of your success. Most points will be earned through:

- Contributing to in-class exercises and assigned group work
- Collaborating effectively in team projects
- Providing peer feedback on code and presentations

Attendance & Time Management

Programming is best done steadily throughout the week, not the night before. Expect to spend at least twice the number of credit hours outside of class, especially if you have little or no prior experience.

If you're falling behind, contact us early – we can help you catch up.

Academic Integrity

All work must be your own or explicitly attributed. Collaboration is encouraged, but copying code or analysis without credit will result in referral to **Student Conduct**.

AI Policy: Generative AI tools may be used *only when explicitly permitted* (e.g., brainstorming, documentation drafting). For all other assignments, use of AI without disclosure is prohibited.

ADA Statement

If you need an accommodation based on a disability, contact Dr. Emrich privately. Full accommodation will be made once approved by **Student Disability Services** (<https://sds.utk.edu/>).

Tips to Succeed in DSE 511

1. Start assignments early and work consistently.
2. Use multiple resources (slides, labs, textbooks, office hours, peer discussions).
3. Treat programming as a practice-based skill.
4. Get help early – do not wait until deadlines.
5. Collaborate responsibly – commit your own work to GitHub, don't copy others'.
6. Learn from feedback and refine your workflows.
7. Stay curious – explore how data science connects to your research domain.