**STAT718/BIOL703**
**Genomic Data Science**
**Fall 2026**
**MW 2:20-3:35 LeConte 206**

**Instructor**: Yen-Yi Ho
Associate Professor, Department of Statistics
**Office**: LeConte 216A
**E-mail**: hoyen@stat.sc.edu
**Class Meeting Times**: MW 2:20PM -3:35PM LeConte 206
**Office Hours**: Thurday 15:00-17:00PM, Friday 10-11:30AM LeConte 216A or by appointment
**Course Website**: https://people.stat.sc.edu/hoyen/STAT718/STAT718.html

## Course Description:

If you're interested in data analysis and interpretation, then this is the course for you. To become an expert data scientist, you need practice and experience. By the completing this course, you will get an opportunity to apply the knowledge and skills in analyzing high-dimensional large-scale genomic data using programming such as Python, R and PyTorch. The course will also include popular artificial intelligence and deep learning applications. We will focuses on quantitative concepts for interdisciplinary applications in genetics as well as "hands-on" experience in analyzing genomic data.

Topics include false discovery rate, gene set enrichment analysis, machine learning algorithms, RNA-seq data analysis, analyzing single-cell data. Advanced topics such as spatial single-cell data analysis, analyses of multi-omics, ChIP-seq data will be included if time permits. In this course, students will have programming exercises in using analysis tools to conduct genome-wide analysis, annotation and interpretation of genetic data.

## Prerequisites

Students will be expected to have taken introductory Statistics course (STAT205 or STAT 201 or STAT515 or equivalent with grade C or better).

## Learning Outcomes

After successful completion of this course, you will be able to:

1. demonstrate advanced knowledge of statistical tools applied to human genetics using open source software, including R and Bioconductor
2. demonstrate advanced skills for genomic visualization, reproducible analysis, data architecture, and exploration of large-scale genomic data

3. apply advanced knowledge, analytical skills and interpret findings for a typical genome-scale assays.
4. analyze appropriate literature to develop a literature review based on a selected topic.

**Course Materials**

***Required***

1. Introduction to Data Science: Data Wrangling and Visualization with R, by Rafael A. Irizarry (link on course website).
2. Python Data Science Handbook by Jake VanderPlas available at https://jakevdp.github.io/PythonDataScienceHandbook/
3. Python for Biologists Tutorial. Available at https://www.pythonforbiologists.org/
4. Machine Learning for Biology Tutorial. Available at https://pythonforbiologists.com/
5. Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville.
   https://github.com/janishar/mit-deep-learning-book-pdf
   https://www.deeplearningbook.org/

***Recommended***

1. Python Crash Course, 3rd Edition: A Hands-On, Project-Based Introduction to Programming by Eric Mattens.
2. Deep learning with PyTorch: build, train, and tune neural networks using Python tools by Eli Stevens, Luca Antiga and Thomas Viehmann.

**Technology Requirements**

Access to a computer (64-bit operating system) with R (available as a free download from the CRAN home page), R Studio and Python/Anaconda. Notice that a Chromebook will not have sufficient computing power required by this course. **It is encouraged to bring your own laptop** to the class and practice along with the programming code during the class. The lecture presentations, programming codes, links to articles, assignments are located on the course website. To participate in learning activities and complete assignments, you will need:

- This course involves hand-on data analysis during the class section. It is encouraged that students bring a laptop during class sections.

- Access to a working computer that has a current operating system with updates installed;
- RStudio program and report writing using R Markdown installed;
- Jupyter and Python/Anaconda, and MiKTeX installed, and
- Reliable Internet access and a USC email account;
- A current Internet browser that is compatible with Blackboard (Google Chrome is the recommended browser for Blackboard);
- Reliable data storage for your work, such as a USB drive or Office365 OneDrive cloud storage.

If your computer does not have RStudio or Jupyter program. It is free and can be downloaded from https://posit.co/download/rstudio-desktop/ and https://www.anaconda.com/download. If you have further questions or need help with the software, please contact the Service Desk (https://www.sc.edu/about/offices_and_divisions/university_technology_services/support/servicedesk.php).

**Minimal Technical Skills Needed**

Minimal technical skills are needed in this course. All work in this course must be completed and submitted using compile pdf file from R Markdown. Therefore, you must have consistent and reliable access to a computer and the Internet. The minimal technical skills you have include the ability to:

- Organize and save electronic files;
- Use USC email and attached files;
- Check email daily;
- Download and upload documents;
- Locate information with a browser; and
- Use Blackboard.

**Technical Support**

If you have problems with your computer, technology, IT-related questions, support, including Blackboard, please contact the Division of Information Technology (DoIT) Service Desk at (803) 777-1800 or submit an online request through the Self-Service Portal (https://scprod.service-now.com/sp) or visit the Carolina Tech Zone (https://www.sc.edu/about/offices_and_divisions/university_technology_services/support/ctz.php)). The Service Desk is open Monday – Friday from 8:00 AM – 6:00 PM (Eastern Time Zone). If you are located in the Columbia, SC area, the Thomas Cooper Library at USC has computers for

you to use in case you encounter computer issues/problems. If you are not located in the Columbia, SC area, most regional campuses and public libraries have computers for public use.

## Course Assignments and Grading

Weekly assignments are due on the day indicated on the course website. Each student will have time to present the final project during class time. The final project will be due on TBA. No late final project report will be accepted.

**Evaluation and Grading Scale**

| **Grading Rubric** | |
|---|---|
| Homework Assignments | 30% |
| Participation/<br>Journal Discussion | 25% |
| Final Project | 30% |
| Fianl Project Presentation | 15% |
| Total | 100% |

The grades will be based on performance on homework assignments (30%), course participation/Journal Discussion (25%), a final course project (30%) and presentation (15%).

**Grading Scale**

89.5% - 100% = A
84.5% - 89.4% = B+
79.5% - 84.4% = B
74.5% - 79.4% = C+
69.5% - 74.4% = C
64.5% - 69.4% = D+
59.5% - 64.4% = D
0% - 59.4% = F

## Course Policies and Procedures

- Please submit your homework via blackboard. Detailed guideline about homework submission can be found on the course website https://people.stat.sc.edu/hoyen/STAT718/Homework/SubmissionGuide.html
- I encourage you to work together in computing and discussing the problems. However, each student is expected to independently write up the submitted assignment using her or his own computing and giving explanations in her or his own words.
- **Late homework will lose 10% of total points per day,** unless arrangements have been made with the instructor for an extension. Homework will not be accepted after the time at which graded

homework are returned.

- Incomplete Contracts: A grade of incomplete “I” shall be assigned at the discretion of the instructor when, due to extraordinary circumstances (e.g., documented illness or hospitalization, death in family, etc.), the student was prevented from completing the work of the course on time. The assignment of an “I” requires that a contract be initiated and completed by the student before the last official day of class, and signed by both the student and instructor. If an incomplete is deemed appropriate by the instructor, the student in consultation with the instructor, will specify the time and manner in which the student will complete course requirements. Extension for completion of the work will not exceed one year (or earlier if designated by the student’s college).

**Attendance Policy**

When you miss class, you miss important information. If you are absent, you are responsible for learning material covered in class. If you are absent when an assignment is due, you must have submitted the assignment prior to the due date to receive credit. If you miss more than 20% of the classes unexcused, your grade will be dropped one letter grade.

**Academic Integrity**

You are expected to practice the highest possible standards of academic integrity. Any deviation from this expectation will result in a minimum academic penalty of your failing the assignment, and will result in additional disciplinary measures. This includes improper citation of sources, using another student's work, and any other form of academic misrepresentation. The first tenet of the Carolinian Creed is, "I will practice personal and academic integrity."

Below are some websites for you to visit to learn more about University policies:

Carolinian Creed (http://www.sa.sc.edu/creed)
Academic Responsibility (http://www.sc.edu/policies/staf625.pdf)
Office of Student Conduct and Academic Integrity (https://www.sa.sc.edu/academicintegrity/)
Network Guidelines for Responsible Computing (http://www.sc.edu/about/offices_and_divisions/university_technology_services/policies_procedures/networkguideline.php)

**Plagiarism**

Using the words or ideas of another as if they were one’s own is a serious form of academic dishonesty. If another person’s complete sentence, syntax, key words, or the specific or unique ideas and information are used, one must give that person credit through proper citation.

**Class Conduct**

Professionalism will be expected at all times. Because the university classroom is a place designed for the free exchange of ideas, we must show respect for one another in all circumstances. We will show respect for one another by exhibiting patience and courtesy in our exchanges. Appropriate language and restraint from verbal attacks upon those whose perspectives differ from your own is a minimum requirement. Courtesy and kindness is the norm for those who participate in my class.

**Generative AI Use Policy**

**Generative AI tools (e.g., ChatGPT) are valuable assistants in your professional toolkit. In this course, you will build skills using Python and R for machine learning in genomic data analysis. The AI policy for this course is:**

1. **AI tools for code** is **allowed**
   - You *may* use AI tools for coding tasks.
   - You **must explicitly cite** their use (with guidance on how to cite).
   - For example:

```
# AI Use Statement:
# I used ChatGPT (OpenAI, Jan 2026) to help debug a recursion
#error  in the function `compute_entropy`. The AI suggested
#checking the base case. I implemented the fix and verified it
#myself.
```

   - Example 2:

> **AI Assistance Disclosure**
> I used ChatGPT (OpenAI, January 2026) to help generate an initial version of the data-cleaning function and to suggest ways to optimize a loop. I modified the code to fit the assignment requirements and tested all outputs independently.

   - Do not **paste assignment prompts** directly into AI tools.
   - You must **create your own prompt**.
2. **AI tools for narrative** is **not allowed**
   - Unless instructed, do not use generative AI to write narratives you submit verbatim.
   - You may not have AI edit a narrative and submit it as your own.
3. **AI tools for learning** is **allowed**
   - You are **encouraged to ask AI questions** that help you learn and understand the course material.

- However, students are cautioned: **don't believe everything AI says**

**Instructional Methods**

The course will be taught using multiple instructional methods. These methods will include lecture, computer programming exercise sections.

**Diversity and Inclusion**

The university is committed to a campus environment that is inclusive, safe, and respectful for all persons, and one that fully embraces the Carolinian Creed. To that end, all course activities will be conducted in an atmosphere of friendly participation and interaction among colleagues, recognizing and appreciating the unique experiences, background, and point of view each student brings. You are expected at all times to apply the highest academic standards to this course and to treat others with dignity and respect.

**Expectations of the Instructor**

I am expected to facilitate learning, answer questions appropriately, be fair and objective in grading, provide timely and useful feedback on assignments, maintain adequate office hours, and treat you as I would like to be treated.

**Course Communication**

I will be communicating with you regarding grades and assignments. If you need to get in touch with me, the best method is via email.
If you are having trouble with this course or its material, you should contact me via email to set up a appointment to discuss the issues.
Emails will be sent your email address. In addition, announcements may be posted on the course website. If you primarily use another email account, you should make sure that the Blackboard account is linked to that address. It is your responsibility to ensure that your email account works properly in order to receive mail.

Please be sure that the email you check regularly is set in Blackboard:

- Click on the My USC tab along the top of the page in Blackboard
- In the Tools module, click on "Personal Information"
- Click on "Edit Personal Information"
- Scroll down to the listing for Email
- In the box will be listed what Blackboard has as your email address. If you wish to change it, delete the email address in the box and type in the email address you want to use.
- Click on the Submit button at the top or bottom of the page.

**Disability Services**

Student Disability Resource Center (http://www.sa.sc.edu/sds/): The Student Disability Resource Center (SDRC) empowers students to manage challenges and limitations imposed by disabilities.

Students with disabilities are encouraged to contact me to discuss the logistics of any accommodations needed to fulfill course requirements (within the first week of the semester). In order to receive reasonable accommodations from me, you must be registered with the Student Disability Resource Center (1523 Greene Street, LeConte Room 112A, Columbia, SC 29208, 803-777-6142). Any student with a documented disability should contact the SDRC to make arrangements for appropriate accommodations.

**Student Success Center**

In partnership with USC faculty, the Student Success Center (SSC) offers a number of programs to assist you in better understanding your course material and to aid you on your path to success. SSC programs are facilitated by professional staff, graduate students, and trained undergraduate peer leaders who have previously excelled in their courses. Resources available to you in this course may include:

- Peer Tutoring: You can make a one-on-one appointment with a Peer Tutor (www.sc.edu/success). Drop-in Tutoring and Online Tutoring may also be available for this course. Visit their website for a full schedule of times, locations, and courses.
- Supplemental Instruction (SI): SI Leaders are assigned to specific sections of courses and hold three weekly study sessions. Sessions focus on the most difficult content being covered in class. The SI Session schedule is posted through the SSC website each week and will also be communicated in class by the SI Leader.
- Peer Writing: Improve your college-level writing skills by bringing writing assignments from any of your classes to a Peer Writing Tutor. Similar to Tutoring, you can visit the website to make an appointment, and to view the full schedule of available drop-in hours and locations.
- Success Consultations: In Success Consultations, SSC staff assist you in developing study skills, setting goals, and connecting to a variety of campus resources. Throughout the semester, I may communicate with the SSC via Success Connect, an online referral system, regarding your progress in the course. If contacted by the

SSC, please schedule a Success Consultation. Success Connect referrals are not punitive and any information shared by me is confidential and subject to FERPA regulations.

SSC services are offered to all USC undergraduates at no additional cost. You are invited to call the Student Success Hotline at (803) 777-1000, visit the SSC website (www.sc.edu/success), or stop by the SSC in the Thomas Cooper Library on the Mezzanine Level to check schedules and make appointments.

Writing Center

Writing Center (http://artsandsciences.sc.edu/write/university-writing-center)

This course has many of writing assignments. The University Writing Center is an important resource you should use! It's open to help any USC student needing assistance with a writing project at any stage of development. The main Writing Center is in Byrnes 703.

Library Resources

Library Resources (http://library.sc.edu)

The university library has great resources for finding out how to cite materials in your projects. Remember that if you use anything that is not your own writing or media (quotes from books, articles, interviews, websites, movies – everything) you must cite the source in MLA format.

Blackboard and Technology

Blackboard and Technology (http://www.sc.edu/about/offices_and_divisions/university_technology_services/)

As a student in this course, you have access to support from the Division of Information Technology (DoIT) for Blackboard and computer issues. The service desk can be reached at 803-777-1800.

Counseling Services (https://sc.edu/about/offices_and_divisions/student_health_services/medical-services/counseling-and-psychiatry/index.php): The University offers counseling and crisis services as well as outreach services, self-help, and frquently asked questions.