MDSC 403
Computation for Bioinformatics

Instructor:

Quan Long, PhD
quan.long@ucalgary.ca
403-220-5580

Office Hours/Policy on Answering Student Emails
Office Hours: Friday 2pm – 4pm. (Students must arrange a minimum of 3 days in advance to meet.) HS 1151 (1173)
Student emails will be answered within 3 business days. Please note that most questions should be posed at the beginning or end of class.

Time and Location:
Thursdays, 3 – 6 pm
This course will take place **online** via Desire2Learn (D2L) and Zoom. All sessions are in real time (i.e., synchronous model).

Prerequisite/Co-Requisite:
Permission of Instructor

Course Description:
This advanced course will focus on current computational and statistical methods in bioinformatics. Topics will shift as computational methods in bioinformatics shift. The underlying philosophy of how to use computational and statistical techniques to solve biological questions will be elaborated using hands-on examples and inquiry-based projects. Students will also learn how to acquire new techniques in the fast-moving field of bioinformatics by leveraging their 'old-version' skills in conjunction with a general understanding of the field. Currently, the course covers (1) R, as a vehicle for data analysis as well as a real programing language; (2) Statistical techniques frequently used in bioinformatics and data sciences, including hypothesis test, statistical inference, and linear models; and (3) Linux command-line and scripting, including the use of popular bioinformatic tools and high-performance computing (HPC) clusters. The course will build basic skills for both upstream computational processing and downstream statistical characterization. Experienced bioinformaticians will give guest lectures on applying computing techniques to biological projects at the end of the course.

Overarching Theme
Statistics and Computation are central to bioinformatics. For many years, statistics and R packages have been seen as essential parts of the toolkit of bioinformaticians. This course will teach basic statistical models and techniques using R as a demonstrating language. As such, it integrates a theoretical course on statistics and a practical course in R. Additionally, most bioinformatics tools are based on Unix/Linux system. Also, Linux system by itself offered powerful commands that can carry out many *ad hoc* analyses, providing convenient alternatives of writing fully documented programs. Students in this course will develop facility in Linux programming, focusing on widely used commands and approaches.

Students will leave this course with the ability to manipulate and analyze data using statistics (R) and computation (Linux), and other bioinformatics tools. There are two formats in this course. The statistical theories of this course will be fully lectured by the instructor. The R/Linux programming will be taught in a blended format with a combination of students using books and online articles at home with in-class lecture content to expand upon a subset of topics.

Global Objectives

By the end of the course, students will have:

- Developed a skillset of statistics and implement them using R.
- Developed a skillset of computation and execute them in Linux system.
- Gained a basic understanding of some of the key issues surrounding manipulation of very large files, especially with a high-performance computing (HPC) cluster.

Course Learning Outcomes

By the end of this course, students will have:

- Understood basic mathematical theories underlying frequently used statistical techniques in bioinformatics and data science
    1. Hypothesis test;
    2. Statistical inference (parameter estimation);
    3. Linear models;
    4. Power issue in experimental design.
- Experienced programming and solving problems using fundamental features of R
    1. Basic R programming;
    2. Statistical functions;
    3. Graphics functions.
- Learned Linux commands, including
    1. Basic commands necessary for bioinformatics data manipulation;
    2. Advanced commands such as ‘awk’ and ‘sort’.
- Developed the skills of using high-performance computing (HPC) clusters
    1. Submission and manipulation of HPC jobs;
    2. Considerations of optimally using shared resources of an HPC.

Learning Resources

There are no required textbooks. Norman Matloff’s The Art of R Programming is used to be part of the reading materials for R. Several reading materials on statistics, Linux, and HPC will be given in the class.

A Note regarding readings

A list of required readings for all course sections will be outlined on D2L and links and documents will be made available, where possible. Required readings have been chosen carefully to inform you and enhance the lecture material**. Students are REQUIRED to complete assigned readings BEFORE each lecture. Readings and sources will be discussed at the first lecture.** Instructors will proceed in class on the assumption that students have read completely the assigned readings. Students should be aware that many of the readings they will be assigned may be of an unfamiliar nature and style. Students should allot sufficient time to allow for several reads of the assigned material.

Learning Technology Requirements

Brightspace (by D2L) is located on the University of Calgary server and will be used extensively for communication with students. **It is the student's responsibility to ensure that they receive all posted communications and documents and that they receive emails sent by instructors or fellow students through D2L.** Only your @ucalgary.ca email address may be linked to D2L. Please ensure that you are regularly checking your @ucalgary.ca account.

In order to successfully engage in learning experiences at the University of Calgary, students taking online, remote and blended courses are required to have reliable access to the following technology:

- A computer with a supported operating system, as well as the latest security and malware updates;
- A current and updated web browser;
- Webcam (built-in or external);
- Microphone and speaker (built-in or external), or headset with microphone;
- Current antivirus and/or firewall software enabled;
- Broadband internet connection

Most current laptops will have a built-in webcam, speaker and microphone.
A laptop, desktop, tablet or mobile device is required for D2L access. If you need help accessing or using D2L, please visit the Desire2Learn resource page for students: http://elearn.ucalgary.ca/d2l-student/.

Evaluation
The University policy on grading and related matters is described in section F of the 2020-2021 Calendar.

In determining the overall grade in the course, the following weights will be used:

- 12% in class participation and performance, including performance on in-class problems
    - This addresses all components of the learning objectives
- 48% after-class assignments
    - This addresses all components of the learning objectives
- 20% Written test
    - This addresses statistical models and its R implementation.
    - Students will have a 24hour window in which they can access and complete a 120-minute test. To allow connectivity issues etc., 50% extra time is granted, which means that students are allowed to choose a 180-min window within a 24-hr period around the test.
- 20% final project
    - This addresses all components of the learning objectives
    - Note that this project component encompasses writing and presentation (see next section)

A student's final grade for the course is the sum of the separate components. It is not necessary to pass each component separately in order to pass the course.

A Note regarding Writing Assignments:

Writing skills are important to academic study in all disciplines. In keeping with the University of Calgary's emphasis on the importance of academic writing in student assignments (section E.2 of 2020-21 Calendar), writing is emphasized, and the grading thereof in determining a student's mark in this course. The Bachelor of Health Sciences values excellence in writing. Competence in writing entails skills in crafting logical, clear, coherent, non-redundant sentences, paragraphs and broader arguments, as

well as skills with the mechanics of writing (grammar, spelling, punctuation). Sources used in research papers must be properly documented. The University of Calgary offers a number of instructional services through the Students' Success Centre's Writing Support Services (http://www.ucalgary.ca/writingsupport/) for students seeking feedback on assignments or seeking to improve their general writing skills. Students are **strongly encouraged** to take advantage of these programs.

Grading Scheme:

Letter Grade	Description	Percentage
A+	Outstanding performance	96-100
A	Excellent performance	90-95
A-	Approaching excellent performance	85-89
B+	Exceeding good performance	80-84
B	Good performance	75-79
B-	Approaching good performance	70-74
C+	Exceeding satisfactory performance	65-69
C	Satisfactory performance	60-64
C-	Approaching satisfactory performance	57-59
D+	Marginal pass	54-56
D	Minimal pass	50-53
F	Did not meet course requirements	0-49

Missed Components of Term Work:
At each instructor's discretion, late assignments will not be accepted and will automatically receive a mark of zero, (or students will lose 5% per day late past the deadline for all assignments. In this case, assignments will **NOT** be accepted more than 72 hours after the posted deadline and students failing to submit any assignment within this time frame will receive a mark of zero. **Students who miss a quiz will receive a mark of zero unless the instructor has been previously notified. There will be NO exceptions to this policy.**

It is the agreement of all Faculty involved in MDSC 403 that **extensions will NOT be granted** on any assignment or quizzes. The only exceptions to this are those in keeping with the University Calendar (debilitating illness, religious conviction, or severe domestic affliction) that are received in writing and with supporting documentation. Traffic jams and late or full buses are common events in Calgary and are NOT acceptable reasons for late arrivals to class, meetings and examinations. Please note that while absences are permitted for religious reasons, students are responsible for providing advance notice and adhering to other guidelines on this matter, as outlined in the University Calendar (https://www.ucalgary.ca/pubs/calendar/current/e-4.html).

Course Evaluations and Student Feedback

Student feedback will be sought at the end of the course through the Universal Student Rating of Instruction (USRI) and a qualitative student evaluation. Students are welcome to discuss the process and content of the course at any time with the instructor.

Guidelines for Zoom Sessions

Zoom is a video conferencing program that will allow us to meet at specific times for a 'live' video conference, so that we can have the opportunity to meet each other virtually and discuss relevant course topics as a learning community.

To help ensure Zoom sessions are private, do not share the Zoom link or password with others, or on any social media platforms. Zoom links and passwords are only intended for students registered in the course. Zoom recordings and materials presented in Zoom, including any teaching materials, must not be shared, distributed or published without the instructor's permission.

The use of video conferencing programs relies on participants to act ethically, honestly and with integrity; and in accordance with the principles of fairness, good faith, and respect (as the Code of Conduct). When entering Zoom or other video conferencing sessions, you play a role in helping create an effective, safe and respectful learning environment. Please be mindful of how your behaviour in these sessions may affect others. Participants are required to use names officially associated with their UCID (legal or preferred names listed in the Student Centre) when engaging in these activities. Instructors/moderators can remove those whose names do not appear on class rosters. Non-compliance may be investigated under relevant University of Calgary conduct policies. If participants have difficulties complying with this requirement, they should email the instructor of the class explaining why, so the instructor may consider whether to grant an exception, and on what terms. For more information on how to get the most out of your zoom sessions visit: https://elearn.ucalgary.ca/guidelines-for-zoom/.

If you are unable to attend a Zoom session, please contact your instructor to arrange an alternative activity (where available). Please be prepared, as best as you are able, to join class in a quiet space that will allow you to be fully present and engaged in Zoom sessions. Students will be advised by their instructor when they are expected to turn on their webcam (such as for group work, presentations, etc).

The instructor may record online Zoom class sessions for the purposes of supporting student learning in this class – such as making the recording available for review of the session or for students who miss a session. Students will be advised before the instructor initiates a recording of a Zoom session. These recordings will be used to support student learning only.

Attendance

Given the large class performance mark, it is essential that students attend each meeting of this class, unless under conditions indicated above.

Conduct During Lectures

The classroom should be respected as a safe place to share ideas without judgement - a community in which we can all learn from one another. Students are expected to frame their comments and questions to lecturers in respectful and appropriate language, always maintaining sensitivity towards the topic. Students, employees, and academic staff are also expected to demonstrate behaviour in class that promotes and maintains a positive and productive learning environment.

As members of the University community, students, employees and academic staff are expected to demonstrate conduct that is consistent with the University of Calgary Calendar, the Code of Conduct and Non-Academic Misconduct policy and procedures, which can be found at https://www.ucalgary.ca/policies/forms/title.

Students are expected to take notes during class and should not rely solely on material supplied by the instructors.

Use of Internet and Electronic Communication Devices in Class

The Bachelor of Health Sciences program aims to create a supportive and respectful learning environment for all students. The use of laptop and mobile devices is acceptable when used in a manner appropriate to the course and classroom activities. However, research studies have found that inappropriate/off-topic use of electronic devices in the classroom negatively affects the learning of others during class time.

Students are responsible for being aware of the University's Internet and email use policy, which can be found at https://www.ucalgary.ca/policies/files/policies/electronic-communications-policy.pdf .

UNIVERSITY OF CALGARY POLICIES AND SUPPORTS

Copyright

All students are required to reach the University of Calgary policy on Acceptable Use of Material Protected by Copyright (https://www.ucalgary.ca/policies/files/policies/acceptable-use-of-material-protected-by-copyright-policy.pdf) and requirements of the Copyright Act (https://laws-lois.justice.gc.ca/eng/acts/C-42/index.html) to ensure they are aware of the consequences of unauthorized sharing of course materials (including instructor notes, electronic versions of textbooks, etc.). Students who use material protected by copyright in violation of this policy may be disciplined under the Non-Academic Misconduct Policy https://www.ucalgary.ca/pubs/calendar/current/k.html.

Instructor Intellectual Property

Course materials created by instructors (including course outlines, presentations and posted notes, labs, case studies, assignments and exams) remain the intellectual property of the instructor. These materials may **NOT** be reproduced, redistributed or copied without the explicit consent of the professor. **The posting of course materials to third party websites such as note-sharing sites without permission is prohibited**. Sharing of extracts of these course materials with other students enrolled in the course ***at the same time*** may be allowed under fair dealing.

Academic Accommodations

It is the student's responsibility to request academic accommodations according to the University policies and procedures listed below. The Student Accommodations policy is available at https://ucalgary.ca/student-services/access/prospective-students/academic-accommodations. Students needing an accommodation based on disability or medical concerns should contact Student Accessibility Services (SAS) in accordance with the Procedure for Accommodations for Students with Disabilities (https://www.ucalgary.ca/policies/files/policies/procedure-for-accommodations-for-students-with-disabilities.pdf). Students who require an accommodation in relation to their coursework based on a protected ground other than disability should communicate this need in writing to their instructor.

SAS will process the request and issue letters of accommodations to instructors. For additional information on support services and accommodations for students with disabilities, visit www.ucalgary.ca/access/ .

Academic Misconduct

The University of Calgary is committed to the highest standards of academic integrity and honesty. The University of Calgary has created rules to govern all its members regarding the creation of knowledge and the demonstration of knowledge having been learned.

Academic Misconduct refers to student behaviour that compromises proper assessment of a student's academic activities and includes (but is not limited to): cheating, fabrication, falsification, plagiarism, unauthorized assistance, failure to comply with an instructor's expectations regarding conduct required of students completing academic assessments in their courses, and failure to comply with exam regulations applied by the Registrar.

For information of the Student Academic Misconduct Policy and Procedures, please visit;
https://ucalgary.ca/policies/files/policies/student-academic-misconduct-policy.pdf
https://ucalgary.ca/policies/files/policies/student-academic-misconduct-procedure.pdf

Additional information is available on the Academic Integrity website at: https://ucalgary.ca/student-services/student-success/learning/academic-integrity.

Recording of Lectures
Audio or video recording of lectures is prohibited except where explicit permission has been received from the instructor.

Freedom of Information and Protection of Privacy Act
Student information will be collected in accordance with typical (or usual) classroom practice. Students' assignments will be accessible only by the authorized course faculty. Private information related to the individual student is treated with the utmost regard by the faculty at the University of Calgary

Appeals
If there is a concern with the course, academic matter or a grade, first communicate with the instructor. If these concerns cannot be resolved, students can proceed with an academic appeal, as per Section I of the University Calendar. Students must follow the official reappraisal/appeal process and may contact the Student Ombuds' Office (http://www.ucalgary.ca/ombuds) for assistance with this and with any other academic concerns, including academic and non-academic misconduct. Students should be aware that concerns about graded term work may only be initiated **within 10 business days** of first being notified of the grade. https://www.ucalgary.ca/pubs/calendar/current/i-2.html

MEDIA RECORDING (if applicable)
Please refer to the following statement on media recording of students: https://elearn.ucalgary.ca/wp-content/uploads/2020/05/Media-Recording-in-Learning-Environments-OSP_FINAL.pdf

Media recording for lesson capture

The instructor may use media recordings to capture the delivery of a lecture. These recordings are intended to be used for lecture capture only and will not be used for any other purpose. Although the recording device will be fixed on the Instructor, in the event that incidental student participation is recorded, the instructor will ensure that any identifiable content (video or audio) is masked, or will seek consent to include the identifiable student content to making the content available on University approved platforms. Lecture PowerPoints will be shared via D2L. But the full video will not be shared.

Media recording for self-assessment of teaching practices

The instructor may use media recordings as a tool for self-assessment of their teaching practices. Although the recording device will be fixed on the instructor, it is possible that student participation in the course may be inadvertently captured. These recordings will be used for instructor self-assessment only and will not be used for any other purpose.

Sexual Violence Policy
The University recognizes that all members of the University Community should be able to learn, work, teach and live in an environment where they are free from harassment, discrimination, and violence. The University of Calgary's sexual violence policy guides us in how we respond to incidents of sexual violence, including supports available to those who have experienced or witnessed sexual violence, or those who are alleged to have committed sexual violence. It provides clear response procedures and timelines, defines complex concepts, and addresses incidents that occur off-campus in certain circumstances. Please see the policy available at https://www.ucalgary.ca/policies/files/policies/sexual-violence-policy.pdf

Resources for Support of Student Learning, Success, Safety and Wellness

Student Success Centre	http://www.ucalgary.ca/ssc/
Student Wellness Centre	http://www.ucalgary.ca/wellnesscentre/
Distress Centre	http://www.distresscentre.com/
Library Resources	http://library.ucalgary.ca

Wellness and Mental Health Resources
The University of Calgary recognizes the pivotal role that student mental health plays in physical health, social connectedness and academic success, and aspires to create a caring and supportive campus community where individuals can freely talk about mental health and receive supports when needed. We encourage you to explore the excellent mental health resources available throughout the university community, such as counselling, self-help resources, peer support or skills-building available through the SU Wellness Centre (https://www.ucalgary.ca/wellnesscentre/services/mental-health-services) and the Campus Mental Health Strategy (http://www.ucalgary.ca/mentalhealth/).

Student Ombuds' Office
The Student Ombuds' Office supports and provides a safe, neutral space for students. For more information, please visit www.ucalgary.ca/ombuds/ or email ombuds@ucalgary.ca

BHSc Student Faculty Liaison Committee (SFLC)
The BHSc SFLC, with elected representatives from all majors, serves to raise issues of interest to BHSc students to the program administration, including items pertaining to curriculum, scheduling and events. A list of current representatives can be found on the BHSc website.

Student Union (SU) Information
The SU Vice-President Academic can be reached at (403) 220-3911 or suvpaca@ucalgary.ca; the SU representatives for the Cumming School of Medicine can be reached at medrep1@su.ucalgary.ca or medrep2@su.ucalgary.ca.

Student Success Centre
The Student Success Centre provides services and programs to ensure students can make the most of their time at the University of Calgary. Our advisors, learning support staff, and writing support staff assist students in enhancing their skills and achieving their academic goals. They provide tailored learning support and advising programs, as well as one-on-one services, free of charge to all undergraduate and graduate students. For more information visit: https://www.ucalgary.ca/student-services/student-success

Emergency Evacuation/Assembly Points

As part of the University of Calgary Emergency Evacuation plan, students, faculty, and staff should locate the closest Assembly Point in case of Fire Alarm. Safety signage is posted throughout the campus showing the locations and the possible route to these locations. All students, faculty, and staff are expected to promptly make their way to the nearest Assembly Point if the Fire Alarm is activated. No one is to return into campus facilities until an all clear is given to the warden in charge of the Assembly Area. For more information, see https://www.ucalgary.ca/emergencyplan/building-evacuation/assembly-points

Safewalk
Campus security will escort individuals, day or night, anywhere on campus (including McMahon Stadium, Health Sciences Centre, Student Family Housing, the Alberta Children's Hospital and the University LRT station). Call 403-220-5333 or visit http://www.ucalgary.ca/security/safewalk. Use any campus phone, emergency phone or the yellow phone located at most parking lot pay booths. Please ensure your personal safety by taking advantage of this service.

Class Schedule
The following is a list of topics for class, associated readings, and assignment / exam due dates. Please note that unforeseen circumstances may cause changes to the schedule with respect to the timing of topics and readings. Students will be notified of all changes in a timely manner by way of email and D2L announcements. The project and assignment due dates will not be altered.

NOTE: The project due date is 5 pm of the 4th day after the final class. Project format will be discussed in class. Project final materials should be submitted to the instructor by uploading to D2L dropbox. It is the student's responsibility to assure that submissions are successful.

The course activities are synchronous.

Date MM/DD/YY	**Module / Topics**	**Instructor/Guest Lecturer**	**Assignments & Due Dates**
09/10/20	1: R: Data structures and basic functions.	Quan Long	
09/17/20	2: R: Operations and data frame	Quan Long	
09/24/20	3: R: Programming	Quan Long	
10/01/20	4: Random variables	Quan Long	
10/08/20	5: Hypothesis test	Quan Long	
10/15/20	6: Statistical inference	Quan Long	Final project handout
10/22/20	7: Linear models	Quan Long	
10/29/20	8: Power estimation	Quan Long	
11/05/20	9: R Statistics & Graphics	Quan Long	
11/19/20	10: Linux: Basic commands	Quan Long	Written exam handout
11/26/20	11: Linux: advanced commands and HPC	Quan Long	
12/03/20	12: Guest lecture: Advanced topics in bioinformatics research	Paul Gordon	

			Project due 12/09/2020 by 5 pm