ACKNOWLEDGEMENT

UBC's Point Grey Campus is located on the traditional, ancestral, and unceded territory of the xwməθkwəyəm (Musqueam) people. The land it is situated on has always been a place of learning for the Musqueam, who for millennia have passed on their culture, history, and traditions from one generation to the next on this site.

COURSE INFORMATION

Course Title	Course Code Number	Credit Value
Data Analysis in Medical Genetics and Genomics	MEDG 525	3

PREREQUISITES

UBC BIOL 335 (76% minimum) or equivalent genetics classes (evaluated by course coordinator).

UBC BIOL 300 (76% minimum) or equivalent introductory statistics classes (evaluated by course coordinator).

Non-MEDG students interested in taking MEDG 525 must obtain permission from the coordinator. Email the coordinator a list of all upper level genetics and statistics courses that you have taken, your full transcripts & grades, and your CV; include your UBC student number.

COREQUISITES

None.

CONTACTS

Course Coordinator(s)	Contact Details	Office Location	Office Hours
Dr. Wyeth Wasserman	wyeth@cmmt.ubc.ca 604-875-3812	BC Children's Hospital Research Institute 938 West 28th Ave Office 3-3103	Instructors will be available for approximately 30 min after class, or upon request by email.
Dr. Jessica Dennis	jessica.dennis@ubc.ca 604-875-2000 x7871	BC Children's Hospital Research Institute 938 West 28th Ave Office 371A	
Dr. Adi Steif	adi.steif@ubc.ca 604-675-8000 ext. 7980	BC Cancer Research Institute 675 West 10th Ave Office 8.111	

OTHER INSTRUCTIONAL STAFF

Teaching Assistant: TBD

OVERVIEW

MEDG 525 (3 credits) is an elective course designed to introduce Medical Genetics graduate students to data analysis methods. The course covers introductory computer programming, foundational biostatistics, and data science methods as applied to problems in medical genetics and genomics. All enrolled students are expected to have a strong background in the principles and fundamentals of genetics and genomics. Prior experience in computer programming is helpful but not required.

COURSE STRUCTURE

The course consists of six 2-week blocks covering topics in genetic and genomic data analysis, with two blocks per faculty member (see schedule). Students will be assigned four hands-on coding assignments throughout the term. The course will conclude with a final project.

SCHEDULE OF TOPICS

Block	Instructor	Topic	Dates	Location
1.1	Wyeth Wasserman	Foundations in biostatistics and computer programming: R	Week 1 Sept <#>	TBD
2	Jessica Dennis	Study design, random and systematic error	Weeks 2-3 Sept <#>	TBD
3	Jessica Dennis	Association testing and regression	Weeks 4-5 Oct <#>	TBD
1.2	Wyeth Wasserman	Foundations in biostatistics and computer programming: Python	Week 6 Oct <#>	TBD
4	Adi Steif	Unsupervised machine learning: dimensionality reduction and clustering	Weeks 7-8 Nov <#>	TBD
5	Adi Steif	Supervised machine learning: classification and performance evaluation	Weeks 9-10 Nov <#>	TBD
6	Wyeth Wasserman	Advanced and emerging topics in genomic data science	Weeks 11-12 Dec <#>	TBD

LEARNING OUTCOMES

By the end of this course, students should be able to:

1. Interpret and critically assess the methods section of a data intensive research paper (e.g. data used, model selected, generalizability, reproducibility)
2. Design, justify, and implement a data analysis plan based on a genetics/genomics research question using statistical analysis software (R, Python)
3. Describe and apply ethics and principles in data science and applied genetics (use, storage, retention, reporting, ownership)
4. Map emerging methods to statistical frameworks and concepts

LEARNING ACTIVITIES

The course includes a mixture of lectures and hands-on programming assignments. For each block, students will be given a dataset and set of analytical aims. With guidance from the instructors and TA, students will develop an analysis workflow, select appropriate models/statistical tests and visualise the results. Emphasis will be placed on showing one's work in Jupyter notebooks and explaining the reasoning behind each methodological choice. At the end of the course, students will select their own research question, design and carry out the analysis and prepare a final project report.

LEARNING MATERIALS

Reading material as well as assignments for students will be assigned by the TA for each class on UBC's interactive learning interface, Canvas (https://canvas.ubc.ca).

ASSESSMENTS OF LEARNING

The course will be graded based on:

- Class participation (20%)
- Student assignments (4x15% = 60%)
- Final project (20%)

UNIVERSITY POLICIES

UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community.

Harassment and discrimination are not tolerated; nor is suppression of academic freedom. UBC values academic honesty and students are expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions (see UBC Academic Honesty and Standards).

UBC provides appropriate accommodation for students with disabilities and for religious observances. Details of student support policies and how to access support are available on the UBC Senate website.

Students are encouraged to stay at home if they have a communicable illness (such as flu-like symptoms) to prevent the further spread of illness to other students, staff, or faculty. If you are too ill to attend class, you should contact your instructor immediately and submit a Student Declaration of Absence Due to Illness or Injury form.

Please maintain a respectful environment: UBC Respectful Environment Statement.

LEARNING ANALYTICS

Learning analytics includes the collection and analysis of data about learners to improve teaching and learning. In this course, it is planned to:

- View overall course progress

- Acquire personalised feedback via a mandatory questionnaire at the conclusion of the course

LEARNING RESOURCES

Source material required to cover content in this course will be shared on Canvas and/or available through PubMed or the UBC library.

Students will be required to use a laptop/notebook computer that is capable of running specific open-source software (Jupyter notebooks, Python and R). It is expected that most computers purchased within the past 3-4 years will be suitable. If students are concerned about the specifications of their computer or not able to afford or obtain a suitable computer, please contact a course instructor prior to the start of classes to resolve the concerns.

COPYRIGHT

LOCATION OF CLASSES

MEDG 525 is expected to take place in-person in <Room X, at the Life Sciences Centre (LSC), 2350 Health Sciences Mall, BC V6T 1Z3>.

Students with disabilities are encouraged to reach out to the Centre for Accessibility to address any concerns. Please email the course coordinators if you need further information.