2021-2022 CSB Graduate Modules

Students enroll in CSB 1020H or CSB1021H for Fall (F) 2021, Winter (S) 2022, or Summer 2022*, depending on the session the specific module is offered. These two course codes apply to all quarter-credit (0.25 FCE) CSB modules.
***Summer courses/modules cannot be requested on ACORN until March 14, 2022.**

Please note that each quarter-credit module has a unique teaching section, and that code must be entered when requesting a specific module on ACORN.

If you want to request two modules in the same session that have the same course code (e.g. CSB1021H/F, for Fall 2021), you may need to contact the CSB Graduate Office to arrange for enrolment in a second module.

For students in graduate programs outside of CSB, any single quarter credit module may not help complete any of your graduate program requirements.

Seminar and lab-based modules offered 2021-2022

Module: Introduction to R
CSB1020H/F, Teaching Section LEC 0142
Offered by the *Centre for the Analysis of Genome Evolution & Function (CAGEF),*
Fall 2021 session
Instructors:

Dr. David S. Guttman, CSB, CAGEF	david.guttman@utoronto.ca
Dr. Calvin Mok, CAGEF Bioinformatics	calvin.mok@mail.utoronto.ca

Time**:**
September 16 – October 28 (7 weeks)
Thurs, 2:00-5:00pm
Enrollment:
20 graduate students
Up to 40 auditors
Weight: One module (0.25 FCE)
Course Objectives
This is a beginner's introduction to R and the Jupyter Notebook environment for individuals with no prior experience or background. Individuals who complete the course will be able to:

- Work with the Jupyter Notebook environment and navigate the R programming language.
- Understand data structures and data types.
- Import data into R and manipulate data frames.
- Transform 'messy' datasets into 'tidy' datasets.
- Make exploratory plots as well as publication-quality graphics.
- Use string searching and manipulation to clean data.
- Perform basic statistical tests and run a regression model.
- Use flow control and build branching code.

Each class will consist of a short introductory section followed by 'code-along' hands-on learning that will gradually build up the lecture's topic(s). Students are expected to have access to a computer during class and are encouraged to ask questions while coding-along with the instructor. A homework assessment will be assigned after each class to reinforce the skills learned and a final project will test overall knowledge and application. The course will be provided through Quercus using Bb-collaborate.

Course Availability
This course will be presented online and will be available to all graduate students, postdocs, staff, and faculty, although only registered students will be evaluated. The course will count as a single module (0.25 credits) for CSB graduate students. All graduate students interested in taking the course for credit should enroll through ACORN.
Anyone wishing to audit the course should fill out the request form at https://bit.ly/3CqlTQB

Evaluation

Item	Note	% Mark
Homework Assignments	6 weekly assignments x 12% each	72%
Term project	Due 2 weeks after the end of the course	28%

Pre-requisites: Access to a computer. No prior programming experience needed.

Reference Material: *R for Data Science* (http://r4ds.had.co.nz/)

Syllabus

Class	Topic
1	**Introduction to R and Jupyter Notebooks**: R and Jupyter Notebook basics, best coding practices, functions and syntax, data types and structures, mathematical operations with R objects, installing R packages, getting help.
2	**How to read, write, and manipulate your data:** Importing text and Excel files, the dplyr package and functions to manipulate tabular data.
3	**Introduction to Tidy Data**: Wide versus long data formats, reshaping data with the tidyverse package.
4	**Data visualization with ggplot2**: The grammar of graphics; scatter, line, box, bar, and density plots, among other types of graphics.
5	**Data cleaning with regular expressions (RegEx):** Introduction to RegEx; inspecting, cleansing, and data wrangling using RegEx; classes, quantifiers, operators, pattern-matching, and string manipulation.
6	**Linear regressions**: Simple and multiple linear regressions, ANOVA, ANCOVA, model selection
7	**Flow control**: for loops, conditional statements (if, while, repeat, next, and break); troubleshooting loops

Subject to change

Module: The Biology of COVID-19
CSB1021H/F, Teaching Section LEC 0143
Coordinator: *Professor Timothy Westwood*
Offered: Fall 2021 session for 6 weeks running Tuesday October 19th to Tuesday December 7th, plus a final exam. No class during undergraduate reading week.
Weight: One module (0.25 FCE)
Time: Tuesdays 6-8 pm (synchronous online lecture; possibly in person last two weeks). Asynchronous lectures will be available for viewing online.

Depending on enrolment, there may be in person lectures the last two weeks of the course for student presentations. Dates and times TBD.
Location: Online. If there are in-person presentations, they will be at the Mississauga campus, building and room TBA.
Enrollment: limited to 10 students

Description:
In December 2019, there was a coronavirus outbreak in Wuhan China that lead to the global pandemic disease COVID-19 with hundreds of millions of people infected and millions of deaths. The pandemic has had a tremendous effect on worldwide economies, employment, education, social activity, and human health and wellbeing just to name a few.
This course aims to give students a good understanding of: virology including the SARS-CoV2 virus that causes COVID-19; innate and acquired immunology and how our immune system combats viral infections; and vaccines including the leading vaccines being developed against SARS-CoV2. The course will cover some the basics of epidemiology with an emphasis on the epidemiology COVID-19. Several lectures will also be spent on COVID-19 itself including the science behind the tests for the virus, disease symptoms and etiology, and various treatments.

Evaluation:

Assignment 1 (picking papers)	10%
Assignment 2 or oral student presentation	20 %
Written questions (2 x 5%)	10 %
Quizzes (2 x 10%)	20 %
Final Exam (open book) Date TBD	35 %
Participation	5 %
	100 %

Pre-requisites for module:
Third year level Molecular Biology course or second (or third) year Microbiology, Virology or Immunology course.
Reading materials:
Assigned Chapters from- Willey, J., Sherwood, L., Woolverton, C. 2020. Prescott/Harley/Klein Microbiology, 11th or 10th ed.
Assigned articles and reviews.
Website: There will be a Quercus site for this course.

Module: Self/Non-self-Recognition in Plants

CSB 1020H/F, Teaching Section LEC 0106
Coordinators: *Professors Daphne Goring and Keiko Yoshioka*
Offered: Fall 2021 session, for six weeks on October 6, 20, 27, and November 3, 17, and 24 (no class on October 13th and November 10th).
Weight: One module (0.25 FCE)
Time: Wednesdays 1 pm - 3 pm
Location: St. George campus, Earth Sciences Centre, Room TBA
Enrolment: Limited to 9 students

Schedule/Seminar topics: 2 hours/week
Week 1: Introductory lecture and reading assignments.
Weeks 2-6: Student presentations and discussion (Two presentations/week. One in week six)
Week 6: Summary Discussion

Description:
The molecular and cellular basis of self/non-self-recognition has been well-studied in the animal research field, but it is a more recent emerging topic in plant research. The molecular mechanisms of plant self/non-self-recognition is starting to be revealed in two different aspects: self-incompatibility (recognition and rejection of self-pollen) and immunity (plant resistance against micro-organisms). In this seminar course, students will investigate the current knowledge of self/non-self-recognition of plants. The course will provide a forum for an interactive discussion between the instructors and students and will be based on a selection of current high impact primary research papers.

Evaluation:
Each student will present a PowerPoint style presentation based on an assigned topic using a selection of primary research papers. A written summary on the same topic in the format of a *Science Perspective* will also be due at the end of the module. Students are expected to read all papers and participate in all discussions. However, for each presentation, two students will be assigned the task of asking questions and leading the discussion.

50% Presentation (one presentation per student on a primary research paper)
25% Leading discussions/Participating in discussions
25% Written Summary in the format of a *Science Perspective*

Pre-requisites for module: undergraduate courses in molecular and cell biology
Reading materials: TBA
Website: Quercus

Module: Current Techniques in Gene Regulation

CSB1020H/F, Teaching Section LEC 0134
Instructor: *Professor Ho Sung Rhee* (hosung.rhee@utoronto.ca)
Offered: Fall 2021 session (six weeks, from November 3rd – December 8th, 2021)
Schedule: 3:10 PM – 6:00 PM, Wednesdays
Weight: One module (0.25 FCE)
Location: Online
Enrolment: Limited to 8 graduate students

Description:
This course will examine cutting-edge techniques to study gene regulation. Techniques to be discussed include: cell (re)programming into specific cell types, long-range chromatin interaction analysis, protein-DNA interaction analysis using nuclease (e.g., CUT&RUN), current strategies for genome-editing in vivo and in vitro systems such as CRISPR-Cas9, and identifying gene regulatory networks and elements at the single cell level using next-generation sequencing technologies (e.g., scRNA-seq, snATAC-seq). Lectures and seminars will involve presentations and discussions of current topics in gene regulation. Students will present recently published research articles, lead a discussion, and write a report.

Evaluation:

1st seminar presentation:	27%
2nd seminar presentation:	27%
Participation:	22%
Written assignment:	24%

Pre-requisites: Background in cell and molecular biology
Reading materials: Required readings will be primary research articles and reviews, and will be provided during class.
Website: Quercus - https://q.utoronto.ca

Module: Fundamentals of Genomic Data Science
CSB1021H/F, Teaching Section LEC 0131
Offered by the *Centre for the Analysis of Genome Evolution & Function (CAGEF)*.
Fall 2021 session
Instructors:

Dr. David S. Guttman, CSB, CAGEF	david.guttman@utoronto.ca
Dr. Calvin Mok, CAGEF Bioinformatics	calvin.mok@mail.utoronto.ca

Date
November 4 – December 16 (7 weeks)
Thurs, 2:00-5:00pm
Enrollment:
20 graduate students
Up to 40 auditors
Weight: One module (0.25 FCE)
Course Objectives
The rise of next-generation genomics has changed the way we think about, study, and employ genetic data, enabling applications that were, until recently, merely the stuff of science fiction. These advances have dramatically increased both the size and scope of biological datasets, and consequently, increased the need for basic computational literacy for nearly all biologists.

This course is designed to serve as an introduction to genomic data science for students who do not have a background in bioinformatics. Students in the course will learn to perform several basic genomic data analyses using Galaxy, an open, web-based platform that incorporates multiple bioinformatics tools into a friendly Graphical User Interface (GUI). Students will then learn to scale up these genomic analyses using the Unix command line to tackle larger and more complex datasets. During the course, students will learn how to:

- Use Galaxy and command line tools to process and manipulate data
- Use the Integrative Genomics Viewer to visualize genomes
- Work in a Unix terminal
- Install bioinformatics software
- Connect and work on remote servers
- Understand common genomics file formats
- Perform de novo genome assembly, reference-based genome assembly, genome annotation, variant calling, and RNA-seq data analysis.

The course will take advantage of online resources for background material, while spending class time analyzing real data sets. Students are expected to have a basic understanding of genomics and molecular biology, but no prior computational knowledge is required.

Each class will consist of a short introductory section followed by ‘code-along’ hands-on learning that will gradually build up the lecture’s topic(s). Students are expected to have access to a computer during class and are encouraged to ask questions while coding-along with the instructor. A homework assessment will be assigned after each class to reinforce the skills learned. The course will be provided through Quercus using Bb-collaborate.

Course Availability
This course will be presented online and will be available to all graduate students, postdocs, staff, and faculty, although only registered students will be evaluated. The course will count as a single module (0.25 credits) for CSB graduate students. All graduate students interested in taking the course for credit should enroll through ACORN.

Anyone wishing to audit the course should fill out the request form at https://bit.ly/2VBD2Gc

Evaluation

Item	Note	% Mark
Homework Assignments	7 weekly assignments x 10% each	70%
Term project	Due 2 weeks after the end of the course	30%

Pre-requisites: Access to a computer. No prior programming experience needed.

Syllabus

Class	Topic
1	Introduction, Exploring Genomic File Formats
2	Galaxy Platform: Navigation, Quality Control, De Novo Assembly, Annotation
3	Galaxy Platform: Reference Alignment, Variant Detection, RNA-Seq
4	Galaxy Platform: RNA-Seq Command Line: Navigation, File management & manipulation, Accessing remote servers
5	Command Line: Downloading & installing software, $PATH, Testing software
6	Command Line: Quality Control, De Novo Assembly, Annotation, BLAST
7	Command Line: Reference Alignment, Samtools, Variant Detection, RNA-Seq

Subject to change

Course: Computational Genomics and Bioinformatics
Course Code: CSB 1472H/S, Teaching Section LEC 0101
Coordinator: *Professor David Guttman*
Offered: Winter 2022 session
Weight: Half credit (0.5 FCE)
Time: Wednesdays 10 am – 1 pm
Location: St. George campus, Ramsay Wright Building, Room 432
Enrolment: Limited to 10 graduate students (minimum 7 reserved for CSB grads)

*CSB1472H/S is a half-credit course that takes place during the full Winter session. It is the equivalent of two modules. Graduate students should NOT request the course using the undergraduate course code CSB472H1S, because it would not count toward graduate credit.

Description:
Recent technological advances have driven a revolution in genomics research that has had a direct impact on both fundamental research as well as direct application in nearly biological disciplines. These advances have made the generation of genomic data relatively straightforward and inexpensive; nevertheless, the data are meaningless if they cannot be properly analyzed. Computational genomics and bioinformatics are the tools we use to extract biological information from complex genomic data.

CSB1472 will teach you the fundamentals of analyzing genomic data. This course emphasizes understanding how core bioinformatic analyses work, the strengths and weaknesses of related methods, and the important parameters embedded in these analyses. CSB1472 is not an applied methods course, nor a course to for developing new bioinformatic tools, but rather a course designed to provide you with a basic understanding of the principles underlying genome analyses. We will examine the fundamentals of sequence alignment, phylogenetic analyses, genome annotation, gene prediction, and gene expression data analysis. Theoretical, applied, and statistical issues will be addressed.

The material is presented as an inverted course. Lectures are pre-recorded and available prior to class. Class time is devoted to review of the lecture material, discussion of the primary literature related to the course material, and hands-on analysis laboratories.

Recommended text: Zvelebil & Baum 2008 Understanding Bioinformatics. Garland Science, New York.

Course: Methods in Genomics and Proteomics
Course Code: CSB 1025H/S, Teaching Section LEC 0101
Coordinator: *Dr. Pauline Wang*
Offered: Winter 2022 session
Weight: Half credit (0.50 FCE)
Time: Tuesdays 12-4 pm
Location: St. George campus, Earth Sciences Centre Room 4076 & Ramsay Wright 109.
Enrolment: Limited to 2 graduate students

Students who are interested in taking this course should contact Dr. Pauline Wang at pauline.wang@utoronto.ca. The course requires instructor approval, after it is requested on ACORN.

*CSB 1025H/S is a half-credit course that takes place during the full Winter session. It is the equivalent of two modules. This course is also offered to undergraduate students as CSB 474H1S. Graduate students should NOT request this course as CSB474H1S on ACORN, because it would not count toward graduate credit.

Description:
Genomics and proteomics have revolutionized biological research. It is now theoretically possible to fully characterize the structure, organization, regulation and interaction of all genes, proteins and small bioactive molecules in an organism. CSB 1025H/S is an intensive and rigorous laboratory course that will teach students how to produce and analyze data that are central to the fields of genomics and proteomics. The course is divided into three modules, the first of which focuses on genomics, the second on transcriptomics, and the third on proteomics. Each module begins with at least two wet labs where students generate data and end with computer labs where students analyze the data. In this way students will learn how to conduct an experiment from beginning to end. Techniques taught include DNA and RNA extraction, shotgun library construction, PCR, DNA sequencing, expression profiling using microarrays, 2D-gel proteome analysis, mass spectrometry and associated bioinformatics analyses such as sequence analysis and assembly, and statistical analysis of microarray and mass spectrometry data. This is an advanced laboratory and computer-based course, and assumes a strong background in molecular genetics and some prior laboratory experience.

Required Text: No required textbook. Information will be provided through lectures presented in the first wet lab and first computer lab of each module.

Evaluation: Three quizzes (15%), three lab reports (60%), lab performance (25%). Graduate students have an additional grant proposal (20%).

Prerequisite: BIO 260H1/HMB 265H1 (Genetics), BIO 255Y1/CSB 330H1/350H1 or by permission of the instructor. Recommended Preparation: BCH 311H1/CSB 349H1/MGY 311Y1

Module: Introduction to Python
CSB1021H/S, Teaching Section LEC 0140
Offered by the *Centre for the Analysis of Genome Evolution & Function (CAGEF),* Winter 2022 session.
Instructors:

Dr. David S. Guttman, CSB, CAGEF	david.guttman@utoronto.ca
Dr. Calvin Mok, CAGEF Bioinformatics	calvin.mok@mail.utoronto.ca

Date
January 13 – Feb 24 (7 weeks)
Thurs, 2:00-5:00pm
Enrollment:
20 graduate students
Up to 40 auditors
Weight: One module (0.25 FCE)
Course Objectives
This is a beginner's introduction to Python for data science applications. The course is intended for students with no computer science background who want to develop the skills needed to analyze their own data. Students who complete this course will be able to:

- Perform data analysis in Python using the Jupyter Notebook environment.
- Understand Python data structures and data types.
- Manipulate Python objects such as lists, data frames, and dictionaries.
- Import data into Python and transform 'messy' datasets into 'tidy' datasets.
- Use flow control to develop branching code.
- Use regular expression and string manipulation to explore and clean data.
- Make exploratory plots.

Each class will consist of a short introductory section followed by 'code-along' hands-on learning that will gradually build up the lecture's topic(s). Students are expected to have access to a computer during class and are encouraged to ask questions while coding-along with the instructor. A homework assessment will be assigned after each class to reinforce the skills learned and a final project will test overall knowledge and application. The course will be provided through Quercus using Bb-collaborate.

Course Availability
This course will be presented online and will be available to all graduate students, postdocs, staff, and faculty, although only registered students will be evaluated. The course will count as a single module (0.25 credits) for CSB graduate students. All graduate students interested in taking the course for credit should enroll through ACORN.
Anyone wishing to audit the course should fill out the request form at https://bit.ly/2VGoZzl

Evaluation

Item	Note	% Mark
Homework Assignments	6 weekly assignments x 12% each	72%
Term project	Due 2 weeks after the end of the course	28%

Pre-requisites: Access to a computer. No prior programming experience needed.

Reference Material: 2016. Severance, Charles. Python for Everybody: Exploring Data Using Python 3. http://do1.dr-chuck.com/pythonlearn/EN_us/pythonlearn.pdf

Syllabus

Class	Topic
1	**Intro to Python and Jupyter Notebooks:** Generalities about Python, using Jupyter Notebooks, how to run Python code, as well as an introduction to Python variables, functions, modules, best coding practices, data types, missing data, code debugging and getting help.
2	**Python data structures, Numpy and Pandas:** List, Dictionaries, Tuples, Sets, Series, mathematical operations with Python objects, Introduction to NumPy and Pandas.
3	**How to Read, Write, and Manipulate Your Data:** The wide and long formats, reading in data, data wrangling with Pandas, and writing data.
4	**Data visualization with plotnine:** The grammar of graphics; scatter, line, box, bar, and density plots, among other types of graphics.
5	**Flow control:** Flow control, for loops, Conditionals
6	**Regular Expressions:** Classes, quantifiers, operators, pattern-matching, String manipulation.
7	**User-defined functions:** Defining a function, best practices in user-defined functions, and context managers

Subject to change

Module: Current Techniques in Neuroscience
CSB1020H/S, Teaching Section LEC 0124
Coordinator: *Dr. Baohua Liu*
Offered: Winter 2022 session, 6 classes
Weight: One module (0.25 FCE)
Location: Online
Enrolment: Limited to 10 graduate students

Schedule: Tentative – Six Fridays from 9-11 am, all online
Friday, January 21st – Friday February 25th, 2022.

Description:
This course will examine emerging cutting-edge techniques that are revolutionizing fundamental neuroscience research. Techniques to be investigated include: optogenetics, chemogenetics, current strategies for cell-type-specific transgene expression and virus-based circuit tracing, large scale electrophysiology, next generation fluorescent indicators, new imaging techniques such as two photo imaging and super-resolution microscopy. Students will take an active role in researching these techniques and presenting their theoretical foundations as well as practical applications, including advantages and disadvantages, to the class.

Evaluation:

Presentation	60%
Participation	40%

Pre-requisites for module: Background in Neuroscience
Reading materials: Required readings will be primary research articles and reviews, and will be provided during the first week of class
Website: Quercus

Module: Structural Biology in Drug Development & Biotechnology
CSB1021H/S, Teaching Section LEC 0144
Instructor: *Professor Mark Currie* (mark.currie@utoronto.ca)
Offered: Winter 2022 (nine weeks, January 11th – March 15th, 2022)
Schedule: 6:00 PM – 9:00 PM, Tuesdays
Weight: One module (0.25 FCE)
Location: Online
Enrolment: Limited to 12 graduate students

Description:
Biological, disease, and drug mechanisms are all determined by the three-dimensional arrangement of atoms within biological macromolecules. Therefore, knowledge of molecular structure is fundamental to protein engineering and the development of new therapeutics and vaccines. This course will cover the application of structural biology methods to drug development and biotechnology. Students will be introduced to the modern tools of protein structure determination including Cryo electron microscopy, X-ray crystallography, and NMR through lectures and tutorials. Lectures will focus on theory, techniques, and the advantages and limitations of each method. The applications of these methods to the pharmaceutical and biotechnology industries including protein engineering, target selection and drugability, lead identification and optimization, rational drug design, and drug mechanism of action will be explored through student presentations and discussions.

Evaluation:
Presentations: 30%
Participation: 25%
Report: 20%
Tutorials: 15%
Quizzes: 10%

Module: Neuroscience of Behavioural State Control
CSB 1021H/S, Teaching Section LEC 0123
Coordinator*: Professor John Peever*
Offered: Winter 2022 session, starting in January or February
Weight: One module (0.25 FCE)
Time: Winter session, dates and times TBA
Location: St. George campus, Ramsay Wright Building, Room TBA (or online)
Enrolment: Limited to 6 students

Description:
This course will examine the latest advances in how the nervous system controls behavioural states such as sleep, arousal, daily rhythms, breathing and movement. It will consider leading hypotheses on the function of the cell systems, organ systems and at the whole organism level that lead to appropriate and pathological control of such behaviours.

Evaluation:
Students will present 2 seminars, one from each section of the course. They will write a brief synopsis of the presentation. Students will also participate in discussion. Seminars and discussion will focus on issues raised in a selection of primary research papers. Written assignment: students will write a “News and Views” style review article on a current “hot topic” within the field. The subject chosen will accommodate the specific research interests of each student but must be different from those discussed in seminars.

Seminar = 20% (x2 = 40% total)

Synopsis = 10% (x2 = 20%)
Discussion = 10% (x2 = 20% total)
News and Views = 20%

Prerequisites: None
Reading Materials: to be determined by the specific interests of the participating students.

Module: Topics in Cell and Developmental Biology: Biology of Adherens Junctions

CSB 1020H/S, Teaching Section LEC 0145

Coordinator: *Professor Ritu Sarpal*
Offered: Winter 2022 session. The module will begin with an organizational meeting at the end of February, followed by six weeks of seminars from March to April.
Weight: One module (0.25 FCE)
Time: TBA
Location: St. George campus - Ramsay Wright Building, Room TBA. Could be online depending on circumstances.
Enrolment: Limited to 8 students

Description:
Understanding how epithelial cells stick together to form tissues and organs, how they withstand mechanical forces, and how the loss of this adhesion leads to diseases such as cancer and its metastasis, are important fundamental questions in biology with significant implications for human health. Adherens junctions (AJs) are known to be dynamic plasma-membrane structures that mediate cell-cell adhesion and promote tissue integrity. In this course, we will discuss research articles to understand how AJs facilitate multicellular development and how they impact disease.

Evaluation:
Two Seminars: 60% (2 X 30%)
Participation in discussions: 20%
Written report: 20%

Module: Data Visualization and Advanced Graphics in R

CSB1020H/S, Teaching Section LEC 0141

Offered by the *Centre for the Analysis of Genome Evolution & Function (CAGEF),* Winter 2022 session.
Instructors:

Dr. David S. Guttman, CSB, CAGEF	david.guttman@utoronto.ca
Dr. Calvin Mok, CAGEF Bioinformatics	calvin.mok@mail.utoronto.ca

Date
March 3 – April 7 (6 weeks)
Thurs, 2:00-5:00pm
Enrollment:
20 graduate students
Up to 40 auditors
Weight: One module (0.25 FCE)

Course Objectives
This is an intermediate to advanced level introduction to R and the packages associated with visualizing large or complex data sets. Participants are strongly encouraged to have prior experience in R (i.e., Introduction to R, CSB1020). Individuals who complete the course will be able to manipulate and prepare large datasets to produce publication-quality graphics. The goal of this course is to introduce the proper use and interpretation of simple, popular and complex data visualizations. Topics will include

- A deep dive into building relatable figures with the ggplot package.
- Analysis and visualization of large datasets from differential expression experiments.
- Popular visualization methods and packages for genes and genome analysis.

Each class will consist of a short introductory section followed by 'code-along' hands-on learning that will gradually build up the lecture's topic(s). Students are expected to have access to a computer during class and are encouraged to ask questions while coding-along with the instructor. A homework assessment will be assigned after each class to reinforce the skills learned. The course will be provided through Quercus using Bb-collaborate.

Course Availability
This course will be presented online and will be available to all graduate students, postdocs, staff, and faculty, although only registered students will be evaluated. The course will count as a single module (0.25 credits) for CSB graduate students. All graduate students interested in taking the course for credit should enroll through ACORN.
Anyone wishing to audit the course should fill out the request form at https://bit.ly/2VBFble

Evaluation

Item	Note	% Mark
Homework Assignments	6 weekly assignments ranging from 15-20% each	100%

Pre-requisites**:** CSB1020 *Introduction to R* (or equivalent). Access to a computer.

Reference Material: *R for Data Science* (http://r4ds.had.co.nz/)

Syllabus

Class	Topic
1	**Re-Introduction to R, RStudio, and Jupyter Notebooks**: R and RStudio basics, setting up R for Jupyter Notebooks, installing R packages, best practices for producing graphs, best coding practices, functions and syntax, data types and structures, importing and exporting data, tidy data formatting, saving data and plots.
2	**The grammar of graphics with ggplot:** box-, violin-, beeswarm-, and jitter plots, combining layers in ggplot, kernel density plots, and parallel coordinate plots.
3	**Finishing touches for ggplot**: themes, aesthetics, color palettes, mathematic annotation with expression() and bquote(), scaling data, error bars, handling outliers, and multi-panel plots.
4	**Visualizing differential expression data**: heatmaps, volcano plots, side-by-side boxplots, dotplots, and Upset plots.
5	**Common visualization methods for data classification/partitioning:** clustering, principal component analysis, multidimensional scaling, and linear projection with t-SNE plots and UMAP.
6	**Simplifying Genes and genomes:** sequence logos, phylogenetic trees, network graphs, Manhattan plots, Gviz, GenomeGraphs, and other helpful packages.

Subject to change