DATA 1010 SYLLABUS
FALL 2020
BROWN UNIVERSITY
SAMUEL S. WATSON

EMAIL sswatson@brown.edu

WEBSITE https://data1010.github.io

CLASS MEETING Monday, Wednesday, Friday from 10:00 to 11:50, on Zoom and Prismia.

OFFICE HOURS Tuesday and Thursday 15:00 to 17:00.

COURSE DESCRIPTION An introduction to the mathematical methods of data science through a combination of computational exploration, visualization, and theory. Students will learn programming basics, topics in numerical linear algebra and scientific computing, mathematical probability (probability spaces, Bayes theorem, and the central limit theorem), statistics (point estimation, confidence intervals, hypothesis testing, maximum likelihood estimation, density estimation, bootstrapping, and cross-validation), and machine learning (regression, classification, and dimensionality reduction, including linear regression, decision trees, support vector machines, neural networks, principal component analysis, t-SNE, Bayes nets, MCMC, Bayesian methods and probabilistic programming).

TEXTBOOK The course content will has been made available as a sequence of free *Mathigon* courses at https://mathigon.org/data-gymnasia.

For those who would like a traditional textbook with more examples and problems, I recommend *A first course in probability* by Sheldon Ross, *All of Statistics* by Larry Wasserman, and *Pattern Recognition and Machine Learning* by Christopher Bishop.

STUDENT SATISFACTION & INCLUSION My top priority is for you to have an excellent learning experience in this course. I intend to set clear learning objectives and equip you with the right tools to achieve them. I invite comments, criticisms, concerns, and suggestions at any time. If you perceive that you are not doing as well as you'd like, please see me right away. I can help with math/programming/stats concepts, of course, but I am also happy to help you troubleshoot your approach to studying, problem solving, etc. I will grant you the assumption that you are doing your best to learn, and I have zero interest in making you feel judged for where you are in the learning process. *You should have every expectation that you can translate sustained hard work into a high level of course success*.

VIDEO LECTURES In advance of each class session, students will watch lecture videos (recorded during last year's run of the course), answer accompanying questions, and complete assigned Data Gymnasia sections.

CLASS SESSIONS

Each class will be conducted as a guided problem solving session. We will discuss follow-up questions about the prepared videos, recap the main ideas, and solve new exercises. The format will incorporate individual work, group work, and whole-group instruction.

Class will be mediated by Zoom (for audio and video) and Prismia (for chat `https://prisima.chat`), both during the remote period in September and then also when the class goes in-person in October.

Prismia data on student participation will be used to generate a participation score. Students who answer all or almost all of the questions asked on Prismia will receive full participation credit. Others will receive points in proportion the proportion of question answered.

Students who are not able to participate in live class sessions should request accommodation.

GRADESCOPE ASSIGNMENTS

Written assignments will be due at 11 PM on each Friday. In recognition of the extra workload associated with reading, watching videos, and answering questions before class each day, the homework will involve a *small* number of computational explorations and expositional exercises. These assignments will be viewed as writing exercises and assessed based on clarity.

Solutions will be submitted as Jupyter notebooks on the grade submission platform *Gradescope*.

There will be a 2-hour grace period to help you out in the event of technical difficulties. There is no score penalty, but your work will be marred with a red "late" indicator. No submissions are accepted after the grace period.

To sign up for Gradescope, visit www.gradescope.com and use the entry code **M87N7B**. Be sure to use your Brown email address to sign up (the primary one, with an underscore, not any email alias you might use), and also enter your Banner ID (the one that starts with B) in the student ID blank.

HOMEWORK POLICY

There are no dropped homework assignments. Dates are coordinated with DATA 1030 and 1050, so curricular conflicts should not be an issue. In event of medical or family emergency, contact me for accommodation.

ARTICLES

Students will write two in-depth Medium articles developing and explaining an extension of an idea from the course.

GRADING

Class preparation	20%
Participation	20%
Homework	40%
Articles	20%

PREREQUISITES

Prerequisites to this course include problem solving facility with sets and functions, basic programming skills, linear algebra, and some topics in multivariable calculus. We have prepared Data Gymnasia courses, aside from the ones we're using the develop the course content, to help you fill in or refresh any of this background.

COLLABORATION AND ACADEMIC INTEGRITY

You may collaborate on solving homework problems, but you must write your solutions entirely by yourself, and you may not do so with reference to notes taken while working in a group. Writing solutions based on notes which represent the ideas of others short-circuits the exercise and impedes your learning. On each homework submission, write a statement on the front page either listing collaborators or confirming that you did not collaborate. Using the internet to look up solutions to homework problems is not acceptable, although of course you may freely use any available resources to learn the material more generally.

COURSE-RELATED WORK EXPECTATIONS

Students will meet 2 hours per day in class (86 hours total), and readings for class will take about two hours per class day (86 hours total). Weekly written assignments will take about 10 hours per week (140 hours total). In addition, the two Medium articles are expected to require 24 hours each.

DISABILITY SUPPORT

Please inform me if you have a disability or other condition that might require modification of these procedures. I am happy to accommodate your learning needs. You should also contact the Student and Employee Accessibility Services at 401-863-9588 or SEAS@brown.edu.