03-701: Practical Computing for Biologists

Fall 2022

**Instructor**: Joel McManus

 *Office*: MI 255

 *Phone:* 268 9407

 *Email*: mcmanus@andrew.cmu.edu

**Class time**: Mon and Fri, 12:30 PM – 1:50 PM, Posner 146.

**Course objectives:** Modern biology research require s the use of computers, and many research applications involve using (usually free) software written for the linux/macOS command line interface. The goal of this course is to provide biology students with practical computing skills for biological research, with a focus on genomics. Topics covered include using linux, installing and using public software, basic programming with python, and combining tools to make pipelines. This is an introductory survey class designed to provide a basic introduction to multiple tools and platforms, without providing mastery in any particular platform.

**Course work:** Students will be graded on three components:

1. Participation – classes will involve practical exercises. Class attendance and participation contributes 25% of the course grade. Assessed through personal reflections (PR).
2. Weekly homework assignments to reinforce and practice concepts introduced in class will comprise 50% of the course grade. Usually due Thursdays at 10 PM (see schedule).
3. A final project combining public tools to evaluate and analyze public data. The final project will be due during the final day of the course in the form of a short report and accompanying code (25% of the course grade).

Policies

**Excused absences:** Class attendance is required as part of the grade. One excused absence is provided by default. Additional absences can be excused for extenuating circumstances. Absences for religious observances must be submitted by email to the instructor during the ﬁrst two weeks of the course. Zoom is available for health absences (email me >=12 h before class).

**Academic integrity:** All class work should be done independently unless explicitly indicated on the assignment handout. The university’s policy on cheating and plagiarism can be found here: http://www.cmu.edu/policies/student-and-student-life/academic-integrity.html. In part it reads “…students at Carnegie Mellon are expected to produce their own original academic work. Collaboration or assistance on academic work to be graded is not permitted unless explicitly authorized by the course instructor(s). The citation of all sources is required. When collaboration or assistance is permitted by the course instructor(s), the acknowledgement of any collaboration or source of assistance is likewise required. Failure to do so is dishonest and is the basis for a charge of cheating, plagiarism, or unauthorized assistance. Such charges are subject to disciplinary action.” Disciplinary actions are detailed here <http://www.cmu.edu/academic-integrity/headernav/policies.html>. You should be familiar with these policies in their entirety.

**Required Textbook:** This course uses “Computing Skills for Biologists”, available on Amazon.com (<https://www.amazon.com/Computing-Skills-Biologists-Stefano-Allesina/dp/0691182752>). eTextbook and paperback options are available.

**Homework**: Homework assignments include exercises from the textbook ~ 50% and instructor provided exercises ~ 50%. Homework will be submitted online.

**Textbook exercises:** Answer keys for textbook exercises are available at <http://computingskillsforbiologists.com/> under “Downloads”. **For your benefit**, youare expected **to work through textbook exercises on your own** before checking your answers.

1. **Instructor exercises:** Answer keys are not provided in advance for these exercises. You are expected to solve these **independently**.
2. **Be sure to comment any code you write.**

**Late Homework:** You get one "free" late (up to 5 days) homework assignment. Additional late submissions will incur the following grade penalties:

48 hours of the deadline - 10% reduction in the homework assignment grade.

2-5 days after the deadline - 50% reduction in the homework assignment grade.

> 5 days after the deadline will receive a zero grade.

Schedule

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Date** | **Topic** | **References** | **Pre-class** | **Assignments due** |
| Jan 19 | Intro, with bash | CSB Ch 1.1-1.6 | Install Linux, Setup bash | PR – 08/30 10 pm |
| Jan 22 | More bash / zsh | CSB Ch 1.6-1.9 and pdf | Read pdf / Ch 1 intro | HW – 09/08 10 pm |
| Jan 26 | File formats / Public data | NA | Install IGV, watch video | PR 09/10 10 pm |
| Jan 29 | Pub.software / conda | See canvas  | Install anaconda | HW – 09/15 10 pm |
| Feb 2 | Python basics 1 | CSB Ch 3.1-3.6 | Read Ch 3 intro | PR – 09/17 10 pm |
| Feb 5 | Python basics 2 | CSB Ch 3.7, Ch 4 | Read Ch 4 intro | HW – 09/22 10 pm |
| Feb 9 | Regular expressions 1 | CSB Ch 5.1-5.4 | Read Ch 5 intro | PR – 09/24 10 pm |
| Feb 12 | Regular expressions 2 | CSB Ch 5.5-5.9 | Review Ch 5 | HW – 09/29/10 pm |
| Feb 16 | Python cont’d Biopython | CSB Ch 6.4 | Visit biopython.org | PR – 10/01 10 pm |
| Feb 19 | Final Project planning | NA | NA | HW – 10/06 10 pm |
| Feb 23 | R basics 1 | CSB Ch 8.1-.12 | Read Ch 8 intro | PR – 10/08 10 pm |
| Feb 26 | R basics 2 | CSB Ch 8.18-21 | Review Ch 8 | HW – 10/13 10 pm |
| Mar 1 | Last Day of Class – Projects due | NA | NA | Final Projects due at 10 pm |

**AI-assistance:**

Given the goal of the course, AI-assistance programs (e.g. ChatGPT, etc) are acceptable, **when cited**. Please use comments to note code you obtain from such tools, as well as any corrections you make to AI-suggested code.

**Support and Resources for Dealing with Stress:**

Take care of yourself.  Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at <http://www.cmu.edu/counseling/>. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:

CaPS: 412-268-2922

Re:solve Crisis Network: 888-796-8226

If the situation is life threatening, call the police:

 On campus: CMU Police: 412-268-2323

 Off campus: 911

If you have questions about this or your coursework, please let us know.

**Diversity and Inclusion**

Carnegie Mellon University is a diverse community of students, faculty, staff, and researchers. Our diversity comes from our own personal identities, and we must not discriminate against others based on these personal identities. This course is intended to serve students from all diverse backgrounds and perspectives.

At CMU, we work to advance and promote diversity and inclusion because it is just and fair to do so. In this class, we are all expected to be respectful to each other and listen with an open mind to each other’s opinions. This is especially important for teamwork, as everyone has something valuable to contribute, and all must be respected.

Unfortunately, bias and discrimination occur, both intentionally and unintentionally. If you feel someone has been subjected to unfair bias based on their identity, you can share your concerns with me or with the Center for Student Diversity and Inclusion at csdi@andrew.cmu.edu, (412) 268-2150. Incidents can also be anonymously reported at [reportit.net](http://www.reportit.net/) username**:** tartans password**:** plaid.