**Open Source Bioinformatics Project**

Grading: The complete project is worth 20% of your final grade and is graded out of a total of 300 points.

**Part 1: Literature Selection (30pts)**

Get paper approved by October 25, 2019

Search for new bioinformatics literature on the [bioRxiv](https://www.biorxiv.org/collection/bioinformatics) Bioinformatics feed from any time in the last year that you find interesting. Read the preprint and locate any associated software packages or code (usually on GitHub).

Get approval for your selection by emailing the preprint and link(s) to software to [rharbert@stonehill.edu](mailto:rharbert@stonehill.edu) by 4:00PM on October 25, 2019.

Ask questions and share your preliminary choices on the Slack [#bioRxiv](https://introbioinformatics.slack.com/messages/CPEJ65N5P) channel to avoid overlapping paper selections.

**Part 2: Project Proposal (70pts)**

PROPOSAL DUE: November 1, 2019

The project proposal will consist of a ~2 page summary of the manuscript and bioinformatics method you have selected. The proposal will give relevant details on the question or problem being addressed in the manuscript, the methodological approach taken to answer/solve that, and a discussion of the availability of the associated software/code that is supposed to help. The summary will conclude with a preliminary assessment of reproducibility of the study based on how much of the code and original data you are able to locate in freely available repositories (GitHub, Dryad, NCBI, EMBL, or others).

**Part 5: Instructor Consultations (50 pts)**

Between November 1 and Thanksgiving Break meet with your instructor at least twice to work on your project.

**Part 4: Paper (50 pts)**

PAPER DUE: December 10, 2019 (Last day of classes)

Prepare a 2-4 page paper in the style of a scientific research article (Abstract, Introduction, Methods, Results, Discussion) that uses your chosen software to analyze some data. Give background on the tool and the data, lay out the implementation of the analysis, present results (figures), and discuss meaning. The scope of this paper may be scaled to match how challenging the tool is to use.

**Part 5: Demonstration (200pts)**

Prepare a <10 minute demonstration of the bioinformatics method(s) you have selected. You will find the code and attempt to show us how it works using a small data set. You can do this “cooking show” style where you show us the setup with “live-coding” and then move straight to the output and/or data visualization instead of waiting for the program to run if that will take longer than a minute or two.

A peer review day will be held on December 3rd where you will present your demonstration to another student for feedback.

Demonstrations to the class will be conducted December 4th and 9th.