CS 4950 Spring 2016
Computing for the Sciences

Term Project

Goal:
The goal of the project is for each team to find something in their (non-cs) discipline that would be useful for them to learn how to do, now or in a future job or graduate school. The project needs to be able to be accomplished with the skills we are learning this semester.

Topic:
To figure out a good project you might want to ask faculty, or people who work in the career you are interested in, what programming tasks they do (or wish they could do).

Here is a list of project proposals that might give you some ideas:
http://www.stats.ox.ac.uk/research/genome/projects/completed_projects
Most of these will be too ambitious for this course, but there might be pieces of them that could work. If you scroll to the bottom, the "High School" projects might have a more reasonable scope.

Another option might be to work through the computations of a scientific paper or report to see if you can replicate or extend them. Here is an example of a geography-based report:
http://eig.org/wp-content/uploads/2016/02/2016-Distressed-Communities-Index-Report.pdf

As a default, I have some leukemia data that I use for a data mining class. I could modify the tasks to fit what we are doing in this class.

General Plan:
I envision that you will need to:

- Form a team
- Write a proposal
- Get data (this may involve interfacing with or creating a database)
- Pre-process, clean, reformat the data
- Do something with the data
    - compute statistics
    - create graphs
    - match sequences
    - find substrings
- Output as appropriate
- Write a report explaining your project and findings

Getting Started:
The first step is to form teams of four people. If you want to work in a smaller team you will need to get permission from me. I am setting up initial teams of 4, trying to have two CS students – one more advanced and one less advanced – and two biology or other majors. It is important that the group work together, so that the computer science students understand the biology (or other science) behind the project and the biology (or other science) students learn the programming.

Next step is to choose a topic (see discussion above).

Then write a topic proposal. I will meet with each group to discuss the proposal to make sure that it is appropriate for the course.

Topic Proposal:
The proposal should include:

- Introduction: Background and Motivation
- The project:
- Goals and Objectives
- The data
- Work plan
    - Reading
    - Design
    - Programming
    - Testing
    - Writing summary
- References - in appropriate format

Note that the more of this you nail down now, the more you will be able to paste into your final report. Also, the feed back that I get from employers includes the desire to have students get more team experience.

There will be a written report required and probably 15-20 minute group presentations at the final exam time.