CMEE Masters: Computing Coursework Assessment

Note that:

All script/code errors and other info mentioned below are in the weekly assessment log files

In the weekly feedback/assessments, please compare with the solution whenever needed to see why I might have taken off points for a particular exercise/script or code file. We can then discuss these in your 1:1 post-assessment feedback session.

Assignment Objectives: To work on a series of computing/programming exercises and problems in a coherent, modular, reproducible workflow under version control.

Student's Name: Louie Adams

Overall Project workflow

Found all the expected weekly directories in your parent directory.

You had a .gitignore throughout, with meaningful exclusions specific to certain weeks, great. You could have more pattern exclusions – You will likely find this useful: `https://github.com/github/gitignore`

You had a readme file with a list of the weeks' content, and then within each week, a readme with a detailed file list for each week. OK, but you may not always want/need to do it in this detail. For example, no need to list files in Sandbox or Data. After all, the Code files would presumably contain sufficient info about what sorts of data are expected as input.

Of equal or more importance is a description of what the overall project structure is and what the language and dependencies requirements are, which would be useful for a new user trying to understand and run your workflow. As you become a seasoned programmer, you will learn to make the readme file descriptions more informative yet succinct.

Your Git repo size when I checked week 7 was 201.45 MB - a bit too large. This suggests you kept unnecessary binary files under VC, and that you might have committed excessively. You need to address this, if you haven't already!

WEEK 1

Found directories Data, Results Sandbox, Code

Found 11 code files: ConcatenateTwoFiles.sh, CountLines.sh, variables.sh, CompileLaTeX.sh, csvtospace.sh, FirstExample.tex, MyExampleScript.sh, FirstBiblio.bib, UnixPrac1.txt, tabtocsv.sh, boilerplate.sh

UnixPrac1.txt was fine. Each solution was described in a comment, great. You also broke the description down into the key components of the unix command a bit, which - good. Compare with the solutions, especially for the last problem.

csvtospace.sh was fine, but one addition you could have made to the script was to throw an error (with a message) if no input csv file was provided. In general, it is a good idea to add some input checks and return a meaningful message with error for utility files like this, especially in case somebody else uses it. Similar comment for ConcatenateTwoFiles.sh (running without two input files will not work), tabtocsv.sh and CompileLaTeX.sh, CountLines.sh, Variables.sh. But it's OK. No points deleted for this.

Points for this week: 100

WEEK 2

Found the Code, Sandbox, Data, Results directories

Found 18 code files: lc2.py, boilerplate.py, basic_csv.py, test_oaks.py, dictionary.py, debugme.py, scope.py, tuple.py, basic_io.py, lc1.py, oaks.py, loops.py, using_name.py, cfexercises.py, align_seqs.py, sysargv.py, control_flow.py, test_control_flow.py

Found no extra files, great.

using_name.py, cfexercises.py gave indentation errors: -10pts

lc1.py, lc2.py, dictionary.py, tuple.py were all OK. lc*.py did nit give any output to screen (-3pts). The other two could have given better-formatted output (-2pts) — compare with the solution on the repo.

align_seqs.py was fine. You could have written it as a self-sufficient script that would also take external inputs. Do compare with the solution.

All other scripts were fine.

Points for this week: 85

WEEK 3

Found directories Practicals, Code, Data, Results

Found 20 code files: browse.R, PP_Regress.R, apply1.R, sample.R, run_get_TreeHeight.sh, boilerplate.R, TreeHeight.R, PP_Lattice.R, next.R, Vectorize1.R, break.R, basic_io.R, try.R, apply2.R, get_TreeHeight.R, TAutoCorr.R, Vectorize2.R, DataWrang.R, TAutoCorr.tex, control.R

The "extra" Rplots.pdf file is a R quirk — I am not actually deleting points for it, as it only happens when a R script is run externally from unix using `RScript`, and there is no easy solution for this.

Vectorize1.R was fine.

Vectorize2.R was OK, but did not give the full results on screen — compare with the solution (-2pts).

PP_Regress.R: good — also have a look at my solution.

TAutoCorr.R was OK – compare with the solution though (the p-value is too low). The report was OK, but could have gone further. You could have shown the correlation, and plotted the histogram of the permuted correlation coefficients as well. The interpretation was fine, but you could have speculated a bit more!

Points for this week: 98 pts

WEEKS 4, 5 & 6

Not assessed, but happy you kept everything organized as much as possible.

WEEK 7

Found directories Code, Data, and Results

Found a README

Found 17 code files: using_os.py, Nets.py, regexs.py, timeitme.py, DrawFW.py, TestR.R, Nets.R, LV1.py, fmr.R, re4.py, run_LV.py, TestR.py, blackbirds.py, run_fmr_R.py, profileme.py, MyFirstJupyterNb.ipynb, LV2.py

Found the following exra file: MyFirstJupyterNb-checkpoint.ipynb (could have .gitignore-d it) ; -0.5pt

using_os.py worked, but compare with the solution. The code could have provided some meaningful output to screen. -1pt

Nets.py was fine - good job.

blackbirds.py was fine, but could have given nicer output to screen. Do look at the solution.

You did the LV* scripts with profiling. Good. Also look at the solution. LV2.py did not save the plot to results.

Other scripts fine, but some docstrings missing.

Points for this week: 90 pts

Overall Assessment

You did an OK job overall. I would have liked to see you at least attempt a few of those extra credit q's.

Very few errors, good!

Overall, if this is the first time you have done programming in a heady mix of UNIX, Python, & R with a sprinkling of LaTeXand git, you did well.

It was a tough set of weeks, but I hope it gave you an inkling of why and how you would/could use Python, R, UNIX, etc together or as and when required.

Provisional Grade

A*	
A	69
B	
C	
F	

The overall assessment will typically have significantly lesser marks than a simple weighted average of each week's points because the overall assessment is based on not just the "Computing Coursework Assessment Criteria", but also the the "Marking Criteria for Exams, Essays and Coursework".

Both sets of marking criteria are in the Appendix of the SilBioComp document.

We will discuss where gained or lost marks, and what you could have improved further in your 1:1 post-assessment feedback session. To the extent possible, please come with questions about specific scripts based upon the feedback you have received.

Signed: Samraat Pawar

January 31, 2018