

CMEE Masters: Computing Coursework Assessment

Note that:

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

In the weekly feedback/assessments, please compare with the solutions 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: Jake Curry

Overall Project workflow

Found all the expected weekly directories in your parent directory.

You had a .gitignore, with many many exclusions! It could have been more specific to certain weeks. There are many in there that you would likely never need. You will likely find this useful: <https://www.gitignore.io>.

You had a readme file with a description, and then one for each week. The weekly Readmes were seasonably succinct, including a description of what the overall project structure is. You could have stated what the language and dependencies requirements are in each week. Check out this resource: <https://github.com/jehna/readme-best-practices>. As you become a seasoned programmer, you will learn to make the readme file descriptions even more informative yet succinct.

Your Git repo size when I checked week 7 was a reasonable 48.03 MB, suggesting you did not keep unnecessary binary files under VC, and that you did not commit excessively. It could also mean that you did not commit enough, and/or somehow along the way lost parts of your git history — but I won't check these possibilities!

WEEK 1

Found directories Data, Sandbox, Code

Found 12 code files: ConcatenateTwoFiles.sh, CountLines.sh, variables.sh, tiff2png.sh, csvtospace.sh, MyExampleScript.sh, UnixPrac1.txt, tabtocsv.sh, boilerplate.sh, CompileLaTeX.sh, FirstExample.tex, FirstBiblio.bib

FirstExample.tex, FirstBiblio.bib were missing: -20 pts

UnixPrac1.txt was fine. Each solution was described in a comment. You also broke the description down into the key components of the unix command, great. Compare with the solutions to see some more compact solutions.

csvtospace.sh was fine. You could have made it 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 other shell scripts.

Points for this week: 80

WEEK 2

Found the Code, Sandbox, Data, Results directories

Found 19 code files: lc2.py, boilerplate.py, basic_csv.py, cfexercises2.py, dictionary.py, debugme.py, scope.py, cfexercises1.py, tuple.py, basic_io.py, lc1.py, oaks_debugme.py, oaks.py, loops.py, using_name.py, align_seqs.py, sysargv.py, control_flow.py, test_control_flow.py

Found no extra files, great!

lc1.py and lc2.py gave an error: -10pts

They were OK otherwise, but check the solutions for some additional things you could have done.

dictionary.py, tuple.py were fine. They could have produced slightly better-formatted output – Compare with the solutions on the repo.

align_seqs.py was nicely implemented. You could have written it as a self-sufficient script that could also take external inputs optionally (though I did not ask for it specifically). Compare with the solution.

All other scripts were fine – good job.

Points for this week: 89

WEEK 3

Found directories Practicals, Code, Data, Results

Found 27 code files: browse.R, PP_Regress.R, maps.R, apply1.R, sample.R, run_get_TreeHeight.sh, boilerplate.R, TreeHeight.R, PP_Lattice.R, next.R, Ricker.R, Girko.R, Vectorize1.R, break.R, plotLin.R, basic_io.R, nlls_modelling.R, try.R, apply2.R, get_TreeHeight.R, TAutoCorr.R, Vectorize2.R, DataWrangTidy.R, preallocate.R, DataWrang.R, MyBars.R, control.R

Found a populated Results directory; ideally, this directory should be empty other than, perhaps, a Readme.

nlls_modelling.R was an unexpected script - that's OK, but it gave a path error: -2pts

MyBars.R gave an error: -5pts

Vectorize1.R was fine.

Vectorize2.R was fine, nice job — compare with the solution.

PP_Regress.R: fine. Also have a look at my solution.

TAutoCorr.R was fine. However, compare with the solution for a different approach.

The report: nicely implemented, but you could have kept the writeup in a completely separate directory. Glad you plotted the histogram of the permuted correlation coefficients as well. You could also have plotted the correlation pattern itself. Some more interpretation of the results would have been nice. -2pts

You did the Mapping (good description of the data and potential biases) extra credit – +2.5 pts.

Points for this week: 92.5 pts

WEEKS 4, 5 & 6

Not assessed.

WEEK 7

Found directories Code, Data, and Results

Found a README

Found 16 code files: TestR.py, regexs.py, profileme2.py, timeitme.py, blackbirds.py, TestR.R, profileme.py, runLV.sh, fmr.R, using_os.py, LV1.py, MyFirstJupyterNB.ipynb, DrawFW.py, Nets.R, run_fmr_R.py, LV2.py

LV2.py, regexs.py, and using_os.py gave errors: -15 pts

Other than that, The two LV* scripts with profiling were fine. You would ideally have run the LV2.py simulation to equilibrium.

using_os.py is OK other than the error, but compare with the solution. The script could have provided some more meaningful output to screen. -2pts

blackbirds.py was fine, including nicely-formatted output .

Points for this week: 83

Overall Assessment

You did an good job overall. Quite a clean project organization and code, but some errors, most easily fixable.

I would have liked to see you attempt a few more extra credit Qs.

I can see that you like R much more than Python – that’s fine!

Overall, as this is the first time you have done programming in a heady mix of UNIX, Python, & R with a sprinkling of \LaTeX and git, you did very well!

It was a tough set of weeks, but I believe your hard work in them has given you a great start towards further training, a quantitative masters dissertation, and ultimately a career in quantitative biology!

Provisional Mark: 74

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 Assessment Appendix of the online TheMulQuaBio notes and git repository.

We will discuss where you 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. This may require you to compare your code with the solution code in many cases.

Signed: Samraat Pawar

January 18, 2019