# 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**: Talia Al-Mushadani

## Overall Project workflow

Found all the expected weekly directories in your parent directory.

Your Git repo size when I checked week 7 was about 1.11 MB — a good size, 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 the way lost parts of your git history — but I won't check these possibilities!

You had a .gitignore throughout, with meaningful exclusions specific to certain weeks – great. There were many extra ones though, such as Sphinx file patterns. You will likely find this useful as well: `https://www.gitignore.io`.

You had an overall readme file (with nothing useful!), and then one within each week. The weekly Readmes were succinct and clear, including a description of each code file. You could have included the language and dependencies requirements. Also 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.

## WEEK 1

Found directories Data, Sandbox, Code

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

UnixPrac1.txt was fine. Each solution could have been described in a comment, including a breakdown of the key components of each unix command, but that's OK. Compare with the solutions, especially the last one.

csvtospace.sh was fine. You could have made it give (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: 100

## WEEK 2

Found the Code, Sandbox, Data, Results directories

Found 21 code files: lc2.py, boilerplate.py, basic_csv.py, cfexercises2.py, align_seqs_better.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, align_seqs_fasta.py, control_flow.py, test_control_flow.py

Found no extra files; great!

test_control_flow.py gave an error: -5pts

lc1.py, lc2.py, dictionary.py, tuple.py were all nicely done. They could have produced slightly better-formatted output – Compare with the solutions on the repo.

align_seqs.py was nicely done. 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.

You did align_seqs_fasta.py and align_seqs_better.py correctly — so +5 extra credit pts.

All other scripts were fine – great job.

Points for this week: 95

## WEEK 3

Found directories Practicals, Code, Data, Results

Found 35 code files: browse.R, PP_Regress.R, maps.R, Vectorize2.py, apply1.R, sample.R, run_get_TreeHeight.sh, get_TreeHeight.py, Vectorize.sh, replicate.R, boilerplate.R, TreeHeight.R, by1.R, PP_Lattice.R, next.R, Girko.R, Vectorize1.R, How_to_plot.R, tapply.R, SQLinR.R, break.R, plotLin.R, basic_io.R, Vectorize1.py, try.R, apply2.R, get_TreeHeight.R, TAutoCorr.R, Vectorize2.R, DataWrangTidy.R, preallocate.R, PP_Regress_loc.R, DataWrang.R, MyBars.R, control.R

Found a populated Results directory; ideally, this directory should be empty other than, perhaps, a Readme. -2pts

SQLinR.R gave an error: -5pts

Vectorize1.R was fine.

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

Vectiorize3-4.py were not what I had asked (the python versions of Vectorize1.Rand Vectorize2.R)!: but +2 pts for effort

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

TAutoCorr.R was OK, except you resampled with replacement (compare with solution)). The report: the interpretation did not quite make sense. The p-value, which would actually be somewhat wrong because of the resampling problem, indicates significance nonetheless! Glad you plotted the histogram of the permuted correlation coefficients and the correlation pattern itself. -5pts

Also, you could have kept the writeup in a completely separate directory

You did the Mapping (good description of the data and potential biases) and pp_regress_loc extra credits – +5 pts.

Points for this week: 95 pts

## WEEKS 4, 5 & 6

Not assessed.

## WEEK 7

Found directories Code, Data, and Results

Found a README

Found 18 code files: TestR.py, regexs.py, LV3.py, profileme2.py, timeitme.py, blackbirds.py, TestR.R, profileme.py, LV4.py, fmr.R, using_os.py, LV1.py, DrawFW.py, re4.py, Nets.R, run_fmr_R.py, LV2.py, run_LV.sh

MyFirstJupyterNb.ipynb missing: -10 pts

regexs.py, re4.py gave errors: -10pts

The two LV* scripts with profiling were fine. You also did the LV3-4 extra credits: +5pts . Also compare with at the solutions.

using_os.py provided no output to screen. -3pts

blackbirds.py was incomplete – it did not quite isolate the species names: - 5pts

Points for this week: 72

## Overall Assessment

You did a good job overall, including an attempt atleast at most all extra credit Qs.

Neat, clean project organization and code, but a few silly errors. I was impressed by your efforts to understand as many details of the programming languages and coding as possible. You clearly like coding!

Overall, as 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 very, very well! In particular, you seem to be well on your way to becoming comfortable with both Python and R – that's great!

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**: 83

*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