

CMEE Masters: Miniproject Assessment

February 14, 2022

Assignment Objectives: To address on a model-fitting problem using computational methods, and produce a written report, all in a coherent, reproducible, modular workflow under version control.

Student's Name: Chalita Chomkatekaew

Overall Miniproject Mark: 80%

Overall Project Organization

All your directories are in place and uncluttered. You have included a comprehensive readme file with details of the project structure, key files and their functions, and even a list of programming tools (although you only list versions for Python and your OS) and their required packages. This is excellent programming practise, very well done! One suggestion is to pick one naming convention and stick to it – you currently employ a mix of CamelCase and snake_case for your scripts, log files and output files.

You could have put the writeup L^AT_EX source files and pdf in a separate directory – this is what you should aim to do for your final dissertation. Similarly, logs can go in a dedicated logs subdirectory (potentially within the results subdirectory) rather than being mixed in with your actual results.

Overall an excellently organised project, clean and well documented. The potential improvements on are fairly minor and will come with practice. Very well done!

The Code

Your choice of coding tools is appropriate. You have chosen R for the bulk of the actual computation within the project, and python for the master script. Having a preferred language for the computation is perfectly OK, though we advise you to remain open to Python or C for more computationally intensive work in future. Your choice of packages is reasonable and not excessive, which is good - overuse of packages limits your development as a programmer and can pose problems for reproducibility.

Your code is generally well-commented, if perhaps somewhat sparsely so in Data_visualisation.R. Generally speaking it is clear what each part of the code is intended to achieve. Your project is very well partitioned into scripts with specific tasks (wrangling, fitting, analysis, visualisation) and you have even separated out functions and execution to a certain degree, by creating a separate script containing the functions for each of the different models you compare. Excellent!

Your master script ran without error, of which you should be proud! You successfully fit 4 models (quadratic, cubic, Gompertz and logistic) to your data, and compare them using AIC and BIC. However, we note that you chose to log all the population sizes even for the polynomial models. This is an unusual choice, and technically means you have chosen to investigate whether the log of the population follows a polynomial relationship w.r.t time, rather than the population itself.

A better option might have been to fit the polynomial models to non-logged data, the logistic and Gompertz models to logged data, and to manually calculate non-logged residuals for these so that you can still perform model comparison using AIC/BIC.

Recall that you should write into your workflow commands that will delete all existing output files every time the workflow is run (they should be re-generated afresh). You currently only do this for some of the temporary Latex files. Your workflow also generates a stray pdf plot in the code subdirectory.

Your workflow prints clear progress updates to the terminal, this is very helpful for determining what is happening while running your master script. We note that you have chosen to suppress all warnings in your tryCatch blocks using the silent=TRUE flag. This is potentially dangerous since warnings can indicate genuine problems with your code or data. If you have checked the warnings yourself and found them to be unimportant, and just wish not to clutter the terminal output with warnings for the user, this is worth mentioning in the readme and/or in comments.

Your project ran quickly (60s), with most time spent fitting models and optimising start values. Well done for producing efficient yet effective code.

Overall the standard of coding in this project is excellent. The project is cleanly organised and the code runs without errors. Well done indeed!

Marks for the project and computational workflow: 80%

The Report

You understand the methods and tools very well, and have engaged with the literature to a commendable degree. The clarity of your writing occasionally lets you down a little, and you ought to have included a proper discussion of limitations/shortcomings.

Title: Descriptive without being too vague, though does not convey the main finding.

Abstract: Quite acceptable. Brief background to begin, then description of the methods and research question and the main results. Take-home message is a little vague, a more confident final statement would have been good! (68%)

Intro: Thorough and well-presented background with strong links to the literature. Ends with a description of the paper's aims and (somewhat vague) hypotheses. (85%)

Methods: Each section of the workflow is described succinctly and accurately. Plenty of extra credit stuff (random sampling of init parameters, model comparison beyond just AIC, temperature- dependence analysis). Computing tools section is present. (85%)

Results: Comprehensive, though somewhat awkwardly worded in places. Subsections clearly divide the results into distinct findings (although second subsection heading is a little confusing). Figures are included and clearly referenced (though one mistaken ref to Fig 3 rather than Fig 2). (68%)

Discussion/Conclusion: States key findings and competently situates them with respect to the wider literature. However the analysis is occasionally muddled by awkward or unclear language, which is a bit of a shame! Credit for including the temperature analysis and for efforts at interpreting it. Only thing missing really is a proper paragraph discussing shortcomings and

how to overcome them in future. Conclusion is succinct and accurate, although the final sentence is a little vague to me. (72%)

(Some specific feedback is in the attached pdf, and we can also discuss more aspects of your write-up in our 1:1 feedback meeting)

Marks for the Report: 80%

Signed: Samraat Pawar & Alexander Kier Christensen

February 14, 2022

Notes on Assessment :

- This written feedback will be discussed in a 1:1 session scheduled after this assessment has been given to you.
- The coursework marking criteria (included in this feedback at bottom) were used for both the computing and report components of the Miniproject Assessment. *In contrast*, Your final dissertation project marks are going to be based pretty much exclusively on the written report and viva (not code). Expect your final dissertation report to be marked more stringently, using the dissertation marking criteria (also included in this report).
- In the written feedback, the markers may have contrasted what you have done with what you should do in your actual dissertation. *This does not mean that you were penalized* — one of the main goals of the miniproject is to provide feedback useful for your main dissertation. However, there may be cases where what you have done is just really bad practise (for example missing line numbers or abstract), irrespective of whether it is a mini- or main- project report – you will be penalized in that case.
- The markers for this assessment are playing the role of somebody trying to understand and use your project organization and workflow from scratch. So it will seem like the feedback is particularly pedantic in places — please take it in the right spirit!