

# CMEE Masters: Miniproject Assessment

**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 name:** PokMan Ho

## Overall Project Organization

All the directories were in place, you had some graphs in Results already and it looks like you had some processed data in Data as well as the raw data. I suspect you just forgot to remove these before a git push.

You had a `readme` file listing the key files. You also listed packages needed / dependencies (and mentioned what these packages are for), and the R/Python versions used for code development. You included clear instructions for running each script individually as well as the project overall, excellent well done! Overall, good.

You could have put the writeup  $\text{\LaTeX}$  source files in a separate directory – this is what you should aim to do for your final dissertation.

You may want to consider naming your scripts more explicitly in the future (having most of them name `Log_X` is a little confusing). Overall the project organization and documentation was clean and logical — good job!

## The Code

Your choice of coding tools and packages was more or less appropriate. You used a minimal number of packages which is good for your skills and reproducibility.

Your code was clean, R is clearly your favoured language but that is totally OK as your Python was neat and used well. You have compartmentalized your R well, and hard coded mostly only where absolutely necessary – good job. You have demonstrated that you know bash very well, however you may have over-complicated things for yourself (I’ll come to this in a moment).

Running the project gave no errors, however both ‘null device’ and ‘1’ are printed to the terminal after each echoed statement that loads. Equally, not all statements you have included do print to the terminal, which suggests something is going wrong silently! Combined with the fact that the report doesn’t generate if the `Log_report.pdf` is deleted from the results directory tells me that something is going wrong in the pipeline, but without an error message that is quite difficult to track down. If the report is present then there are no errors during  $\text{\LaTeX}$  compilation but the pipeline appears to fail before it reaches this point. The error is to do with executability and how you call your `.sh` scripts – try calling them with ‘`bash`’ as opposed to ‘`./`’ as your scripts weren’t executable on my machine. Once I’d made them executable your pipeline runs without error (barring the printing mentioned earlier).

You should add error messages and checks which abort the pipeline if an error is encountered to help diagnose the issue. You should also add commands to delete any previously generated output so that all the output is entirely fresh.

You printed progress messages to the console which were clearly delineated, this is good!

You were ambitious in coding and it mostly paid off, however there are clearly a few bugs to still work out. Overall, good.

## **The Report**

Good effort. Some technically challenging elements well executed, mixed with some questionable choices of analyses. The writing is inconsistent, with logical clarity a clear area for improvement.

In your main project report, to the extent possible, consider using a results-focused title for your main dissertation — basically, the main finding(s) of the paper can be indicated in the title. Obviously, if you have a lot of findings, you can't do that. But then, it means you have too many 'storylines' in your study.

In your main project report, make sure your Introduction is more explicit about your objectives and hypotheses, and provided clear answers to them at the end, in the Discussion.

Have a look at this for further guidelines on figures for your main dissertation: [http://abacus.bates.edu/~ganderso/biology/resources/writing/HTW\\_Guide\\_Table-Figures\\_9-30-08.pdf](http://abacus.bates.edu/~ganderso/biology/resources/writing/HTW_Guide_Table-Figures_9-30-08.pdf).

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

## Overall Assessment

Marks for the project and computational workflow: 80

Marks for the Report: 65

Overall Marks (*50:50 Computing:Report weighting*): 72.5

**Signed:** Samraat Pawar

June 2, 2020

## 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).
- If there were technical errors made in the model fitting and selection, the points have been deducted from the report (not the computing) component's marks.
- In many cases, the marker would 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!
- Ultimately, keep in mind that this mini-project was partly an exercise in reproducible workflow development — you may need to trade-off some computational elegance (but hopefully not reproducibility!), such as having everything run with one `run_project` command, in favor of a good written report — that's what matters most in the end. In this context, the main thing to keep in mind is that one or both of your markers will likely not be particularly quantitative, so you will need keep the explanations simple (but not patronizingly so!). In general, this advice holds while writing papers for more general (not narrow-subject focused) journals as well — keep it succinct and simple. Therefore, please also consider the report component mark separately from the computing component mark.

## MARKING CRITERIA for EXAMS and ESSAYS and COURSEWORK

The following criteria are the basis on which the Department assesses both exam answers and coursework.

Literal Grade	Criteria (Problem type answers are marked on a semi-absolute scale)
A*	<b>Exceptional</b> Answer is an exceptionally well presented exposition of the subject, showing: (i) command of the relevant concepts and facts, (ii) a high critical or analytical ability**, (iii) originality, and (iv) evidence of substantial outside reading (where applicable). Numeric marks available 100, 95, 90, 85.
A	<b>Excellent</b> Answer is a very well presented exposition of the subject, showing many of the above features, but falling short in one or two of them. Numeric marks available 80, 76, 72.
B	<b>Very Good to Good</b> Answer (i) shows a clear grasp of the relevant concepts and facts, (ii) gives an accurate account of the relevant taught material ( <i>as exemplified in the model answer</i> ), and (iii) shows evidence of some outside reading or of critical or analytical ability**. Numeric marks available 68, 65, 62.
C	<b>Adequate</b> Answer: (i) shows a grasp of the basic concepts and facts, (ii) gives a mainly accurate account of at least half of the relevant taught material ( <i>as exemplified in the model answer</i> ), and (iii) does not go beyond that, or goes beyond that but is marred by significant errors. Numeric marks available 58, 55, 52.
F	<b>Unsatisfactory</b> Answer:  1.shows only a weak grasp of the basic concepts and facts, and is marred by major errors or brevity; numeric marks available 48, 45, 42; 2.shows a confused understanding of the question; is too inaccurate, too irrelevant, or too brief to indicate more than a vague understanding of the question; 35, 30, 25; 3.includes at most one to four sentences or facts that are correct and relevant to the question; numeric marks available 20, 15, 10, 5; 4.contains nothing correct that is relevant to the question; numeric mark 0.

\*\* *Analytical* = assessing a hypothesis or statement by breaking it down into its elements and examining their inter-relationships and contribution to the whole; cf. *Critical* = judging a hypothesis or conclusion by examining the validity of the evidence adduced for it.

**MSc & MRes PROJECT ASSESSMENT – THESIS (OR MID-PROJECT REPORT)  
MARKING CRITERIA**

Literal Grade	% Grade	Criteria (Please give leeway if it is a mid-project report)
<b>A*</b>	100 95 90 85	<b>Exceptional.</b> Work is of a publishable standard**. It is an exceptionally well presented exposition of the project, showing: (i) command of the relevant concepts and facts, (ii) a high level of analysis, (iii) originality in thought and experimental or modelling design, and (iv) mastery of the relevant literature.
<b>A</b>	80 76 72	<b>Excellent.</b> Thesis is written to a publishable standard** with minor revision. It is a very well presented exposition of the project, showing most of the above features, but falling short in one of them.
<b>B</b>	68 65 62	<b>Very Good to Good.</b> Thesis contains potentially publishable material**, but needs revision of the text and further research. It is otherwise a well presented exposition of the project, showing: (i) a clear grasp of the relevant concepts and facts, (ii) appropriate, though not highly sophisticated analysis, and (iii) a sound knowledge of the relevant literature.
<b>C</b>	58 55 52	<b>Adequate.</b> Thesis is not written to a publishable standard and requires major revision and substantially more research. It is an adequately presented exposition of the project, showing: (i) a grasp of the basic concepts and facts, (ii) an adequate use of statistics in its analyses, and (iii) sufficient knowledge of the relevant literature to set its results in a scientific context.
<b>D</b>	48 45 42	<b>Unsatisfactory.</b> Thesis is an incomplete presentation of the project and is marred by major errors or gaps, missing analysis, lack of references, misconceptions, excessive brevity, etc, at most showing a weak grasp of the basic concepts and facts.
	35 30 25	Thesis as above, but presentation extremely poor and overall impression indicates a very weak grasp of the basic concepts and facts.
	20 15 10	Thesis as above, and in addition no real attempt to analyse data or present results in a scientific manner.
	5	Thesis as above but incomplete and lacking understanding in all areas.
	0	Thesis not produced.

\*\* This publishability implies that the data or theory is *per se* worth publishing.