Applying Computational & Data Science

Handbook 2023

2023-11-16

Dept. Earth Science & Engineering

Authored by: Gareth Collins

Imperial College London

Table of Contents

1 Overview

1.1 Synopsis

The aim of this module is to simulate applied computational and data science in the real world. The module comprises three group projects that synthesise different elements of the MSc program. Each project will span five days of full-time effort, simulating a working week. On Monday morning, groups will meet with the project "client". The client will outline the problem to be solved and be available for one-to-one meetings with the groups to specify the project scope. On Friday afternoon, the groups will deliver their work to the client for assessment. Each project will be assessed based on the group presentation; the software produced (e.g., functionality, robustness, ease-of-use, documentation); the software's performance in the specified tasks; and the ability of the group to work well as a team.

> "Applying Computational/Data Science will give you a real-world experience of high-productivity problem solving."

1.2 Objectives

- To simulate applied computational/data science in the real world.
- To synthesise knowledge from the taught modules
- To apply techniques learned in the course to real problems
- To develop collaborative programming skills
- To reinforce best practise for software development

1.3 Learning Outcomes

On successful completion of this module, you should be able to:

- Plan and produce software collaboratively.
- Collaboratively solve problems using software.
- Summarise work using collaborative presentations.

1.4 Module coordinators

You are welcome to contact the module coordinators Gareth Collins (ACSE) and James Percival (EDSML & GEMS) with any questions about this module.

2 Projects and Schedule

2.1 Group project format

Each project will run for one working week, Mon-Fri. The project will begin on Monday morning with a client briefing from an academic staff member or industry partner acting as the client. This briefing will specify the problem to be solved, provide the necessary scientific background to understand the problem and explain the project-specific deliverables and assessment criteria.

After the briefing, including questions, you will break out into teams and be given time to plan how you will work together. During this period, each group will be given the opportunity to ask the client questions to clarify the project scope, deliverables and assessment criteria.

The deadlines for the software/presentation will be different for each project, but usually some time on Friday.

From the end of the briefing on Monday until the final deadline on Friday, you are free to work together as you and your team see most appropriate. A space in the department (or at ScaleSpace) will be available for you to use for a period during the day, but you are welcome to work elsewhere if preferred. GTAs/staff will be available to answer questions at certain times during the day on Tue-Thu. You may also contact the module coordinator/project leader at any point with a query.

Each project has been designed such that the primary task is achievable in the time available, allowing for reasonable working hours. You are not expected to work long hours into the evening. Wednesday afternoon is sports afternoon and you are not required to work during this time. Each project will include challenging elements. The purpose of these is to keep you stimulated throughout the week if you make quick progress on the primary task. It is perfectly possible to achieve an A grade without completing 100% of the assignment. Equal effort should be expended to

develop software that is high quality and sustainable, as well as software that seems to solve the problem at hand.

Group members will be required to complete a short, anonymous peer-evaluation of your group colleagues, which will help inform the teamwork element of your assessment.

2.2 Schedule

There will be three group projects spread throughout the year.

1. Project 1: 20th-24th November, 2023
2. Project 2: 29th January-2nd February, 2024
3. Project 3: 20th-24th May, 2024

3 Working as a Team

An important aspect of this module is to provide experience of developing software, solving problems and presenting solutions collaboratively. Teamwork is a key skill that will be developed during the projects that comprise this module and a component of the assessment will reflect how well you work together.

Over the course of the year, you will work on three group projects. In each case, the composition of the groups will be chosen for you. You will work with different people for each project. We hope that you will all endeavour to work harmoniously and in good spirits with enthusiasm and application. If you have any issues with your group, please raise these with the module coordinator.

It will be up to you, as a group, how you choose to work together. You may wish to elect a group leader to oversee the project, delegate tasks and ensure good timekeeping. Or you may prefer to adopt a consensual approach with regular planning meetings. Similarly, it may be useful to discuss what each team member thinks that they can bring to the project and allocate roles based on individual strengths. Or you may prefer to work collaboratively on each aspect of the project. After the first group project, you should reflect on your previous experience and endeavour to challenge yourself to address weaknesses, for example by nominating yourself for roles where you are forced out of your comfort zone.

One of our goals in establishing this MSc programme is to encourage the computational and data science community to be more diverse: whoever you are, and whatever your background, we welcome you. During this MSc course, and in particular during this module, we hope that everyone finds their experience welcoming, encouraging and rewarding. We want to foster a community based on mutual respect, tolerance, and encouragement and we kindly ask that you respect these principles. If you experience or witness any unwelcome behaviour, we encourage you to challenge the behaviour or report it, in confidence, to the module coordinator or the course director.

Among the most common causes of disfunction in a team are poor communication and a mismatch of expectations. You may find it useful, therefore, to discuss in your groups how you wish to work together, as well as what you want to achieve. This might involve agreeing a team code of conduct or contract. An example team contract is provided in Appendix D.

At the end of each project you may be asked, as a group, to self-assess your "Group Process." In other words, how well you think you worked as a team. The purpose of this self-assessment is entirely formative: it should help you identify your individual strengths and weaknesses in teamwork and strategies for enhancing collaborative work in future. An example of a Group Process Self-Assessment form is given in Appendix B.

You will also each be asked to evaluate the relative contribution of each group member to the team project, including yourself. This score will be used to determine the "Team contribution" component of your project assessment. This is weighted as 10% of your overall individual mark for each project. An example of the peer-assessment form is given in Appendix C.

4 Assessment

The group projects in this module will each be assessed in the same way. There are three components to the assessment:

1. Software (70%)
2. Technical report / presentation (20%)
3. Teamwork (10%)

Aspects of the assessment criteria for each project will differ to reflect the specific requirements of each project. Details of specific assessment criteria will be included in the problem specification for each project. This handbook summarises the general principles of assessment, focusing on the assessment criteria that will be common to all four projects.

4.1 Software

Your software will be assessed based on:

- Functionality
- Performance
- Sustainability

Functionality means how well the software performs in the assigned tasks. The specifics of how functionality will be assessed will differ from project to project.

Performance means how fast or how efficiently the software performs the specified tasks. The importance of efficient of fast software will vary from project to project and will be assessed accordingly.

Throughout the course, best practice in software **sustainability** is advocated. This means that every software project that you undertake should strive to be: licensed, version-controlled, documented and tested.

The group projects in this module are an excellent opportunity to put what you have learned into practice. An important component of the mark

awarded for your group project will be allocated on the basis of software quality. Appendix A provides a table of questions that assessors will be asking of each software project.

For each group project in this module a GitHub repository will be created for you to host your software. It is up to you how you develop and populate this repository, but you may wish to consider the following:

1. What license should you choose for your software project?
2. How will you manage your project and issues that occur during development?
3. What code review policy will you use?
4. What form of testing is sensible and manageable in the time available?
5. How will you ensure that your code is usable?
6. How will you ensure that your code can be developed and improved by others?

4.2 Presentations

In addition to the software, each project will require submission of a Technical Report and/or delivery of an Oral or Video Presentation. Specific details of the content and assessment criteria of the Technical Report or Presentation will be given in the problem specification for each project. This handbook describes the general assessment criteria that will apply to all Technical Reports / Presentations in this module.

The Technical Report will normally take the form of a short report that provides answers to each of the Project subtasks. Subtasks will be clearly explained during the client briefing. They will normally comprise some technical writing (for example, a brief description of the solution method or algorithm or a brief description of how to use the software) as well as examples of the software output (for example, a test demonstration of the software or some graphical display of a simulation output or dataset).

If required, the Technical Report should be uploaded (committed) to the Project repository prior to the submission deadline. Written feedback will be provided on the Technical Report.

Alternatively, Projects may involve the delivery of an in-person or video presentation to the “client” and staff on Friday afternoon. Specific details of the content and assessment criteria of the presentation will be given in the problem specification for the project where relevant. This handbook describes the general assessment criteria that will apply to all presentations in this module.

The presentation will typically take the form of a very short presentation, possibly followed by a question and answer session if delivered in person. You may choose to present as a group, with different speakers for different parts of the presentation, or you may nominate a smaller team to speak on behalf of the whole group. For in-person presentations the whole group must be available and prepared to answer questions.

After an in-person presentation you will be asked questions by course staff and the “client” that might address any aspect of the project, including technical questions about your solution algorithm, implementation, documentation and testing or questions about how your group managed the project, divided workload, worked collaboratively, etc.

Alternatively, you may be asked to record a short video presentation to address some of the project deliverables. In this case, clear instructions will be given about what should be included in the presentation as there will be no opportunity for questions to be asked after the presentation.

4.3 Teamwork

The projects in this module are group projects. Teamwork is therefore an essential skill that will be developed in this module.

Your contribution to the team will be assessed both implicitly and explicitly. Your ability to work well as a team will be reflected implicitly in the success of your project and 90% of your individual mark for each project will be awarded based on your group's product (70% software, 20% report / presentation).

You will also receive an individual mark for your contribution to the group project that will constitute 10% of your total mark for that project. This will be determined by an anonymous peer evaluation at the end of each project. An example of the anonymous peer evaluation is provided in Appendix C.

5 Feedback

You will receive several forms of feedback for this module. Each group will receive written feedback on the group software project, technical report and presentations. This will highlight strengths and weaknesses of each element of the project. You will also receive an individual Provisional Grade (A-F) for the project as a whole, which includes the group project mark plus your teamwork score. You will receive this feedback within four weeks of the end of the project.

You will also be asked as a group to assess your group performance as a formative self-assessment. The purpose of this exercise is to identify strengths and weaknesses in your group process and highlight areas that you intend to work on in future group projects. An example of a Group Process Self-Assessment form is given in Appendix B. This exercise is entirely formative and does not form part of the assessment.

6 Mitigating circumstances

If you are unable to participate, partially or fully, in any of the group projects, for whatever reason, it is important that you inform the Module Coordinator as early as possible. You should also consider informing the rest of your group. It is not possible to reschedule or rerun the group projects, however, any absence owing to circumstances out of your control can be taken into account.

7 Academic Integrity and Misconduct

Academic integrity means that your work acknowledges the ideas and results of others, that it is conducted in an ethical way and that it is free from plagiarism. Academic integrity is fundamental to learning, teaching and research and it is important to understand what it means to you and the international community of research that you are joining.

Academic misconduct is the attempt to gain an academic advantage, whether intentionally or unintentionally, in any piece of assessment submitted to the College. This includes plagiarism, collusion, or other dishonest practice. Even though you are working collaboratively, these conducts must be avoided in the group projects as with any other coursework.

Plagiarism

Plagiarism is the presentation of any thoughts, words, images, research or diagrams that were created by someone else, or generated by an AI tool, as though they were your own. In your group projects, this means any other person **outside your group**. Plagiarism may be intentional, by deliberately trying to use another person's work by disguising it or not citing the source, or unintentional where citation and/or referencing is incorrect.

Plagiarism must be avoided. You should take particular care when making use of software written by someone outside of your group. Each project will provide specific guidance about which aspects of the software must be written from scratch and which aspects can make use of packages or code written by others. However, if you use code sourced externally from your group you must include clear and proper attribution of credit. This includes code generated by AI tools such as ChatGPT or CoPilot. You can support your understanding of proper referencing and citation by using the resources available from the College such as the Library learning support webpages at:

www.imperial.ac.uk/admin-services/library/learning-support/plagiarism-awareness/

Where plagiarism is detected in group work, members of that group may be deemed to have collective responsibility for the integrity of work submitted by that group and may be liable for any penalty imposed, proportionate to their contribution.

Collusion

Collusion is the term used for work that has been conducted by more than one individual, in contravention of the assessment brief.

You should note that whilst the College encourages students to support each other in their studies you should be careful to ensure that you do not exceed any assessment brief with regards to group work. For group projects, you are actively encouraged to work collaboratively as a team—sharing ideas and code within your team is **exactly** what you should be doing and is **not** collusion. In group projects, collusion constitutes the sharing of between groups. You are therefore strongly discouraged from discussing the group project with peers outside your project group. As a guiding principle: always acknowledge the contributions of others in your work, and do not leave yourself open to allegations that you have supplied answers to enable another student or group to commit academic misconduct.

Where it is alleged that there has been collusion, all parties will be investigated under the Academic Misconduct procedure.

Appendix A: Software Quality Checklist

Question	Y/N
Does the software have a licence?	
Is there a high-level description of what/who the software is for?	
Is there a high-level description of what the software does?	
Is there a high-level description of how the software works?	
Is there an example of how to use the software?	
Is there specific documentation for users?	
Are there clear instructions for how to install the software?	
Are there clear instructions for how to use the software?	
Are any of the features of the software undocumented?	
Is there specific documentation for developers?	
Is the algorithm behind the software clearly explained?	
Is there an API where appropriate?	
Is the code clearly commented?	
Is the code clear? Does it conform to PEP8?	
Is the code version-controlled with code review?	
Does the code-base include a test suite?	
Are there any unit tests?	
Are there any integration tests?	
Are there any regression/validation tests?	
Is the test-suite comprehensive?	
Is the test-suite automatically run?	

Appendix B: Example Group Process Self-Assessment Form

Sample Group Process Evaluation Form						
Individually, reflect on your group's dynamics and – anonymously – rate them according to each of the following variables (using a scale from 1 to 5). As a group, discuss the results and brainstorm concrete ways to improve your group processes.						
Goals						
Goals are unclear or poorly understood, resulting in little commitment to them.	1	2	3	4	5	Goals are clear, understood, and have the full commitment of team members.
Openness						
Members are guarded or cautious in discussions.	1	2	3	4	5	Members express thoughts, feelings, and ideas freely.
Mutual Trust						
Members are suspicious of one another's motives.	1	2	3	4	5	Members trust one another and do not fear ridicule or reprisal.
Attitudes Toward Difference						
Members smooth over differences and suppress or avoid conflict.	1	2	3	4	5	Members feel free to voice differences and work through them.
Support						
Members are reluctant to ask for or give help.	1	2	3	4	5	Members are comfortable giving and receiving help.
Participation						
Discussion is generally dominated by a few members.	1	2	3	4	5	All members are involved in discussion.
Decision-making						
Decisions are made by only a few members.	1	2	3	4	5	All members are involved in decision-making.
Flexibility						
The group is locked into established rules and procedures that members find difficult to change.	1	2	3	4	5	Members readily change procedures in response to new situations.
Use of Member Resources						
Individuals' abilities, knowledge and experience is not well utilized.	1	2	3	4	5	Each member's abilities, knowledge, and experience are fully utilized.

Adapted from Russ Christianson, CoopZone.

Appendix C: Example Group Project Peer-Evaluation Instructions

Sample Numerical Peer Evaluation

Please use this form to evaluate the contributions of each team member to the group effort. Consider attendance and participation in team meetings, individual contributions to idea generation and research, communication within the group, etc. These evaluations are completely confidential and will never be shown to your team members. Please respond as honestly as possible.

1. Please allocate a total of 100 percentage points among your team members, including yourself, with higher percentages going to those members who contributed most. In the case of equal contribution, points should be divided equally among team members.

   Your name: _____

   Group name: _____

	Name	% Points
Yourself	_____	_____
Member 1	_____	_____
Member 2	_____	_____
Member 3	_____	_____
Member 4	_____	_____
Member 5	_____	_____
	Total	100 %

2. Explain any particularly high or low allocations, providing concrete examples to illustrate your reasoning.

We hope you enjoyed the group project. As explained in the module briefing, your mark for the project includes a mark for teamwork (weighted 10%) based on *peer evaluation*.

Please use the process described below to evaluate the contributions of each team member to the group effort. Consider attendance and participation in team meetings and chat, individual contributions to idea generation and research, communication within the group, etc. Please keep in mind the current circumstances when awarding relative contributions. These evaluations are completely confidential and will never be shown to your team members. Please respond as honestly as possible.

Submission process

In this repository, you can find a `points.csv` file and my-contribution.txt. Please ***modify the points column*** in `points.csv` to assign a contribution score (an integer) between 0 and 40 to each team member. A greater contribution should receive a higher score. If you wish to give an equal score (20 points) to each team member, you can leave the `points.csv` file unchanged.

Please assign points to each member of your team following two rules:

1. The sum of all points in `points.csv` must equal the number of people in `points.csv` times 20. For example, if there are 7 rows with people in `points.csv`, the total sum must be equal to 7 × 20 = 140.
2. If you give to anyone less than 5 or more than 30 points, you must write a comment in the comment row (format: `Comment,"your comment in quotes"`) to justify your decision.

Contribution Score	Meaning
31-40	I believe that this person made an essential contribution to the success of this project, and that it could not have been completed without their work.
25-30	I believe this person made a significant contribution to the success of this project

25-21	I believe this person made an above average contribution to the success of this project.
20	I believe this person made an equal contribution to the work on this project.
19-15	I believe this person made a below average contribution to the success of this project
9-14	I believe this person made a reduced contribution to the success of this project compared to other team members
1-5	I believe this person made essentially no contribution to the project. They were absent or uncontactable without reason for much of the time we worked on it.
0	I believe this person made no contribution to the project. They were absent or uncontactable without reason for virtually the entire time.

Secondly, in my-contribution.txt, with less than 300 words, explain what parts of the group project you worked on and what your contribution was.

After you are happy with your points, please commit (and push) `points.csv` file, and **ensure that the tests in GitHub Actions workflow pass**.

If you have any questions or difficulties, don't hesitate to get in touch with Marijan (m.beg@imperial.ac.uk).

Appendix D: Example Team Contract

Sample Team Contract

Team Name: _____ Date:_____

GOALS: What are our team goals for this project? What do we want to accomplish? What skills do we want to develop or refine?
EXPECTATIONS: What do we expect of one another in regard to attendance at meetings, participation, frequency of communication, the quality of work, etc.?
POLICIES & PROCEDURES: What rules can we agree on to help us meet our goals and expectations?
CONSEQUENCES: How will we address non-performance in regard to these goals, expectations, policies and procedures?

We share these goals and expectations, and agree to these policies, procedures, and consequences.

Team member name	Team member name
Team member name	Team member name
Team member name	Team member name