

# MATH5872M - Dissertation in Data Science and Analytics Student Handbook 2016/17

Module Manager: Charles Taylor

April 21, 2017

## 1 Introduction

In addition to what is written in this document, Section 6 of the School of Mathematics Taught Postgraduate Student Handbook gives some guidance on your MSc Dissertation and Section 2.10 of that handbook for a reminder about plagiarism (what it is and why you *must not* commit it). The purpose of this document, which should be read in conjunction with the above references, is to give you further details.

Recall that the dissertation is a compulsory module that must be passed for the award of the MSc degree; it is worth 60 credits out of the total of 180 credits required to be studied for the degree. Thus, it is a significant part of your year's study.

The dissertation is a piece of work executed by the student; it will be assessed on its academic content and on the quality of its presentation. The dissertation is examined by the School of Mathematics, and then sent to the External Examiner for his/her assessment. Part of the assessment of the dissertation is by an oral presentation and an oral examination.

## 2 Timetable

The deadline for submission of completed dissertations is **11pm on Friday, 25 August, 2017**. Indeed, an electronic version must be submitted through the VLE by **11pm on Friday, 25 August, 2017**. An academic integrity form must also be handed in along with a paper copy of a title sheet and abstract to the Taught Student Office before the deadline. (No paper copy of the dissertation needs to be handed in.)

These deadlines will be firmly adhered to unless a student has written permission in advance to submit his/her dissertation at a later date. In accordance with University rules, late submission without prior approval results in a deduction of 5 marks (out of 100) for each day between 25 August and the date of submission; submissions more than 14 days late, without prior permission, will not be accepted, and you will fail the module, and hence the degree. Note that problems with submitting the dissertation due to the lack of internet access, issues with TurnItIn, etc., are not regarded as valid reasons for late submission. Students can make a case that 'special circumstances' apply to them. Information on these special circumstances can be found on the School of Mathematics website; any such cases related to the submission of the dissertation must be made by 24 August 2017.

The oral examinations and presentations will be held in the period from **Friday 8 to Friday 15 September 2017**; students will be advised of the exact date and time subsequently. **All students must be available throughout this period.** Students are advised to have a copy of the dissertation for their own records and to refer to in the oral exam.

### 3 Assessment

The written dissertation will be assessed by an internal assessor appointed by the Module Manager; this internal assessor will be a different person from any supervisor. The dissertation will also be assessed by the External Examiner, who must confirm the final grade to be awarded for the dissertation.

The oral presentation will normally be attended by at least the supervisor(s), the internal assessor, and the Module Manager. Each student will be allocated one hour for the oral exam. In this hour, the student's presentation will last for a maximum of 20 minutes, and this will be strictly enforced. There will be questions for approximately 30 minutes.

Provisional grades for the written dissertation and the presentation will be assigned by the assessor in consultation with other colleagues present. These grades will be moderated by the Module Manager to ensure consistency between different candidates and between different subject areas. All provisional marks are sent to the relevant External Examiner for his/her comments.

There will be a formal Examiners' Meeting in November 2017. It is necessary that the external examiners confirm these marks. When they have done this, marks will be available via the Student Portal.

### 4 Supervision

Your supervisor will provide guidance on your dissertation. This will include:

- giving some initial discussions on the general methodology;
- giving advice on programming;
- suggesting references to books and journal articles as sources;
- discussing in general terms the plan of the dissertation when you have made some progress, advising what it is realistic to include;
- helping you to notice deficiencies in your work;
- commenting on presentation, and organisation of the dissertation when shown draft versions;
- occasional advice on problems with  $\LaTeX$ .

However, you should not expect your supervisor to read every word in your drafts, nor to provide detailed comments on each chapter, nor to give detailed instructions on how to write in  $\LaTeX$ .

The frequency and length of meetings with your supervisor are somewhat flexible. In general, each student should agree a timetable of contact meetings with their supervisor(s). Note that, in the period from June to September, supervisors have their own research, conferences, family and other commitments besides their role as a supervisor, and so are not normally available for the whole of this period.

As specified in the catalogue, the supervisor(s) should be available for a total of 6 hours.

Should you have difficulties with your dissertation, please discuss these with your supervisor in the first instance. Where a problem is unresolved, please then contact the Module Manager. If the problem persists, please consult the Head of Department.

## 5 Advice on timing

To meet the deadline for the submission of your dissertation, you should schedule your work through a series of self-imposed targets. This means that you must plan ahead and keep closely to your plan. Let us recall that problems with L<sup>A</sup>T<sub>E</sub>X compiling, access to the internet, etc., are not regarded as a valid excuse, so you should ensure that you have allowed adequate time to circumvent such difficulties.

As part of the project, you may be required to write and run your own computer programme, and analyse your own data or data from other sources; please be aware that this can be time-consuming and so should not be left until late in the project.

Students are strongly advised to submit drafts of and plans for their dissertation to their supervisor; please ensure that he/she has sufficient time to make comments.

## 6 Written presentation and content

The objectives as specified in the Module catalogue are that  
*‘On completion of this module, students should be able to:*

- *be taught through research experience the planning, execution and maintenance of a data science project;*
- *present their project as a dissertation, which will be presented orally.*

It is not a requirement a dissertation for the MSc includes original methods for statistical data analysis. However a good dissertation will often improve and clarify a presentation of a topic from specified sources. A dissertation that does contain some correct, original work, such as variations on existing techniques, resolution of specified questions from the sources, or new examples, may be regarded as ‘outstanding’. Similarly, a dissertation which applies standard methods in a thoughtful way to a large and/or complex data set, yielding new understanding of the data, would be regarded as ‘outstanding’. The balance of literature review, mathematical theory, and data analysis will vary from topic to topic.

It is stressed that the the aim of the dissertation is to assess your ability to undertake independent work. Most dissertations are necessarily a compromise between lofty goals, and what can be achieved in reality, and this is recognised by the examiners. Try to make your dissertation interesting to read and reflect your own motivation.

The standard of English, style and overall presentation is your responsibility. Your supervisor is not responsible for proof-reading, or checking your grammar and spelling; however, he/she should comment in some detail on the clarity and the English and grammar of a sample of your draft work.

## 7 Format of the report

The final report must be written in  $\text{\LaTeX}$  and then the pdf file should be submitted via the VLE. A template .tex file will be available from the module area on the VLE.

As a guide, the final dissertation should contain 60–80 pages of text (not including the title pages, etc., and the bibliography). However, please note that this is only a guide: an excellent dissertation could have fewer than 60 pages; certain subjects might require additional tables, computer programs, appendices, etc., taking the total beyond 80 pages. If you think that the body of your dissertation (excluding title pages, table of contents, appendices, and bibliography) will exceed 80 pages, you must obtain written permission from the module manager.

The submission must be accompanied by an academic integrity form (this form can be found on the School of Mathematics website), and will be checked for plagiarism using the standard University software. If any plagiarism is suspected, the standard University procedures on this matter will be followed. Information on what constitutes plagiarism can be found in the University Taught Student Guide.

The final dissertation should be accompanied by a summary page on one sheet of paper describing the scope of the work and the main results, indicating the main sources used. This may be the same as the abstract page within the dissertation if you wish.

Pages and sections should be numbered for easy reference, and the title page should contain your name, student number, the module code and title, the title of the project, the name of supervisor(s), and the date of submission. The report should include a contents page and a full bibliography, indicating all the books, articles, and websites that are referred to.

Recall that, in the text, if a lemma or theorem is taken from some source, there must be a reference to the source; if it has been significantly expanded or adapted, then you should say this. Apart from short statements, you should not copy directly from sources, and everything must be expressed in your own words, except where explicitly stated. Students whose work is too closely based on their sources will have marks deducted.

### Summary

The submitted dissertation should contain the following:

1. Title page: Name, number, module, title, supervisor, date;
2. One-page summary of the dissertation and the sources used;
3. Contents page;
4. Numbered pages (and sections);
5. Bibliography of sources used.

## 8 Advice to students on tackling the project

You are expected to draw up detailed notes on the topic, appropriately illustrated with examples and counter-examples, and results of your data analysis as appropriate. The final format of the report will depend somewhat on the subject area and topic chosen. One common format is that the completed set of notes should be, as near as possible, the sort of notes and examples a student would be expected to collect during a lecture course at the level of the MSc. Another format is that the dissertation might be a complete report which could be presented to the person

whose data has been analysed. This can involve paraphrasing material from a book, and solving problems in textbooks and other sources to which the answers are not provided. Some projects may require the student to write extensive computer code for simulations or data analysis; it may be appropriate to include sample annotated code as an appendix.

### **How to write a good project**

1. The definitions and statements that you make should be correct and precise; it may well be that you and the assessor and the External Examiner know ‘what you mean to say’, but the assessors read what you write on the page, not what they guess might be in your head.
2. Recall that technical writing is not the same as casual conversation: you are expected to write with greater formality and care, in precise English, than when speaking to other people. Recall also that English is a flexible language, but that there ‘rules of grammar’, and you are expected to follow them.
3. The project should be well-organised, and follow a logical structure, often moving from the general to the more specific.
4. A lot of advice on writing mathematics can be found in Chapters 3 and 4 of Dr. Kevin Houston’s book, *How to Think Like a Mathematician*. For example, write in sentences, explain your notation, and, when you give a definition, give at least one example.
5. Give answers to challenging questions. Assessors are more likely to be impressed by your solution to harder exercises.
6. Give your own examples after definitions.
7. Generalise results.
8. Collect material from a variety of sources.
9. Use consistent notation. Different books use different notation; if you replace the notation in the correct places, then you demonstrate understanding.
10. If you have to take something verbatim from a source, then quote the source and demonstrate that you understand it by giving a pertinent discussion or a good example.
11. As a rule of thumb, a method which you have learned on the MSc programme need not be described in detail, but one which you have learned as part of the project should be explained as if the reader is unfamiliar with it.
12. Make sure the account is clear (e.g., technical terms are defined) and logical (e.g., the explanation of a method which you have had to research comes *before* its use in a data analysis).
13. Use a number of sources from journals.

A key point is that you should *demonstrate* that you understand what you have written.

## **9 Advice on the presentation**

Given the time constraints, it will not be possible for you to explain every detail from your report. Thus you should choose some aspect, e.g., a particular method, set of examples, or aspect of the data analysis, and use that to demonstrate that you understand the material in the report.

There are two main aims to the presentation:

1. students should demonstrate that they understand and have mastered the material;
2. students should demonstrate their presentation skills.

Students can make their presentation in any format they like, e.g., using a computer projector, OHP slides, black/white board, video, flip-chart, etc. A computer projector and black/white board will be available. Students should inform the Module Manager at least one week ahead of time if they have any other requirements for the presentation. It cannot be assumed that the room for the presentation will contain a computer with every possible computer program. Students are responsible for ensuring **in advance, and not at the time of the presentation**, that they can use the equipment that they require.

You should bear the following point in mind while preparing your presentation.

- The presentation will pass by quicker than you think, so plan carefully.
- Cover some part of the material in depth rather than all parts superficially.
- A good example is worth a thousand words.
- $\LaTeX$  reports can be quickly turned into a 'Powerpoint' type presentation by using the  $\LaTeX$  package Beamer. Google 'Latex beamer' to find this, or consult the School of Mathematics  $\LaTeX$  page <http://www1.maths.leeds.ac.uk/latex>.
- **Practise your presentation** in front of friends and/or other students. This will give you confidence and you will be less nervous. It also allows you to judge the timing. It is possible to book a room in which you can practise the presentation in front of some friends; to arrange this, please contact the Taught Student Office.
- Does it look professional?

## 10 Assessment Guidelines

The dissertation will be assessed under the following headings, and each heading carries a different weight. Some notes on how to interpret these headings is given below; in each case, the examiners will take the difficulty of the material in the project into account.

**Understanding (30%)** *Based on report, presentation and answers to questions, does the student understand the methods described and the work done? Are suitable analyses carried out and/or examples used to illustrate the theory? Are sound conclusions drawn from any analyses, simulations or examples?*

**Achievement (20%)** *Is the work done of the quantity and level that could reasonably be expected of a competent student in the time available? Is there a derivation of an original result, substantial analysis of a dataset or great effort spent programming?*

**Initiative (20%)** *Did the student exercise initiative; for example, by influencing the direction of the project or by locating their own sources of data and/or information?*

**Report (20%)** *Is the report laid out well, with good structure and use of figures, tables etc? Is there clarity in the exposition? Are results (theoretical, or of data analysis and simulations) precise and unambiguous? Are there few typographical errors and are any mathematical expressions clearly formatted?*

**Presentation Skills (10%)** *Was a suitable amount of material chosen for the presentation, and does it include the key points? Are the slides clear and engaging, and is the presentation given effectively?*

The following is meant as guidance to what constitutes a particular overall mark. Please note that these are indicative and not exhaustive examples of what may be considered typical work at a given level of performance.

### **Distinction, 70 – 100**

- Extremely well organised and presented.
- Project could serve as a basis for a course at the Master's level, or be presented as a professional analysis of challenging data.
- Excellent choice of examples and logical flow.
- Required little help from supervisor (relative to the difficulty of topic).
- Good evidence of originality and independent thinking.
- Mastery of material.
- Thoughtful and correct data analysis.

To achieve a Distinction students do not need to have achieved mastery or excellence in all the above. Marks will be given for originality and evidence of independent thinking.

As we have stated, in some cases a dissertation will contain original results in statistics. (However, this is not a requirement). The presence of some original work will indicate that the dissertation may be awarded a distinction, but other factors will also be considered.

The full range of marks from 70 to 100 will be utilised by the examiners, with the following criteria in mind.

**Distinction, 95 – 100** Excellent in all criteria, with original work of publishable quality.

**Distinction, 85 - 94** Excellent in most criteria and highly competent in others, shows mastery of the material. The dissertation could be used as a basis for a course in the material without many changes.

**Distinction, 75 - 84** Excellent in many criteria and competent in others, demonstrating a high degree of mastery of material.

**Distinction, 70 - 74** Excellent in many criteria and competent in others, demonstrating a high degree of mastery with some minor gaps.

### **Merit, 61 – 69**

- Well organised and presented.
- Good choice of examples and logical flow.
- Required reasonable help from supervisor (relative to the difficulty of the topic).
- Some evidence of independent thinking; follows standard texts sometimes.
- Sound understanding of the material.
- Project could with some significant corrections and/or extensions be used as a basis for a course on the material.
- Generally satisfactory data analysis which could be improved in small ways.

**Pass, 50 – 59**

- Adequately organised and presented.
- Reasonable choice of examples and logical flow.
- Required substantial help from supervisor (relative to the difficulty of the topic).
- Little evidence of independent thinking; tends to follow the sources closely.
- Includes some good understanding of material.
- Some reasonable data analysis, but which is incomplete or needs significant correction.

**Fail, 0–49**

- Poorly organised and presented.
- Poor choice of examples and logical flow.
- Required significant help from supervisor (relative to the difficulty of the topic).
- No evidence of independent thinking; very closely follows sources.
- Little understanding of the material.
- No new examples.
- Contains confused or illogical arguments and definite errors.
- Data analysis is largely wrong or misleading.
- Many errors in English that make the dissertation hard to read.