MSc in Data Science

Course Handbook 2020-21

Moodle web-board

Programme contact email

datasciteaching@lancaster.ac.uk

Contents

Programme, Study and Support Information

1. Welcome to Data Science at Lancaster

1.1 Introduction

Welcome on board the MSc Data Science! The MSc Data Science programme has been running since 2014-15, and over this period both the programme and the field have expanded to meet the increased demand within industrial, commercial and research sectors to explore, and extract meaningful information from, large data sets. We hope that your time studying with us will prove to be a challenging, yet rewarding, experience that prepares you for your future career as a Data Scientist.

This handbook provides an overview of the programmes structure and the modules offered. We are incredibly excited about this coming year, the curriculum, the available pathways and the opportunities that your studies will provide.

In formulating the degree scheme, we have drawn on our world-leading research and teaching expertise in Statistics, Computer Science, Data Mining, Business Analytics, Health Informatics and Environmental Science to provide you with a cutting-edge curriculum. This will include the use of state-of-the-art methodologies and big data technologies together with data from a variety of domains (e.g. social media, health, ecology, business etc.). When designing the curriculum, we also sought guidance from businesses and research organisations (so you will be exposed to real-world data science problems) several of whom will provide students placements for Masters students.

You will have the opportunity to learn from expertise drawn from departments across the university: the School of Computing and Communications (SCC); the Department of Mathematics and Statistics; the Lancaster Environment Centre (LEC); Lancaster University Management School (LUMS) and the CHICAS group in the Faculty of Health and Medicine (FHM). We will make extensive use of virtual learning environments (Moodle and Teams) through which you can network with your peers and ask questions of your tutors and lecturers.

1.2 Single point of contact

Your single point of contact for all programme and placement related enquiries is the Data Science programme email box:

datasciteaching@lancaster.ac.uk

For enquiries about specific modules you should contact the Student Office in the relevant department:

Mathematics and Statistics email address: mathsteaching@lancaster.ac.uk

SCC email address: scc-pg-office@lancaster.ac.uk

1.3 Departmental Support

All students will be allocated an Academic Advisor who will support you with any issues related to your studies and direct you to relevant central University services as appropriate. In addition to Academic Advisors, the MSc Data Science Team are here to ensure the successful delivery of the programme and to support your academic development throughout the coming year. The Team consists of the Programme Directors, the Head Tutor, Engagement Manager and Teaching Co-ordinator.

Programme Directors:
Prof Chris Edwards (email address: c.edwards@lancaster.ac.uk)
Dr Adam Sykulski (email address: a.sykulski@lancaster.ac.uk)

Head Tutor:
Dr Clement Lee (email address: clement.lee@lancaster.ac.uk)

Engagement Manager:
Simon Tomlinson (email address: s.tomlinson2@lancaster.ac.uk)

Teaching Coordinator: Katie Park Walford

The goal of the academic advisor system is to provide each student with a single advisor for the duration of the year, who

- holds a meeting at the beginning of each term;
- provides help with module choices;
- monitors the student's progress;
- supports the student's career planning;
- signposts the student to services available elsewhere in the university.

Once a term, your academic advisor will invite you for a brief meeting, usually no more than 10 minutes in length. This will give you the chance to discuss any issues that may have arisen, and to ask questions about your studies.

Your advisor will be able to advise you on various matters, and point you in the right direction for others. They are a good person to ask to provide a reference for a potential employer.

It is intended you will build a positive and lasting relationship with your academic advisor. However, it's always possible that difficulties may arise, and the Head Tutor and Programme Directors are available to help in this event.

1.4 Term Dates

Dates for the three University terms (Michaelmas, Lent and Summer).

2. Remote Learning

The 2020-21 year will be different to any other as we continue to navigate the changes brought about by the coronavirus (Covid-19) pandemic. In everything it does, Lancaster University will be following the advice given by the UK government to ensure that all students and staff can continue to study and work in a safe environment. Full details of the Lancaster Promise to students, including information on living on campus and flexible learning.

Frequently Asked Questions about living on campus, including available services and facilities, as well as teaching and assessment. These FAQ will be updated on a regular basis, so this site is a good first point of call whenever you have a Covid-19 related question.

The Student Services Directory provides information on all services available on campus, and any current changes to their availability.

As part of our flexible learning approach, the first term of the MSc Data Science will be taught using a blend of synchronous (real-time) and asynchronous (pre-recorded) online digital teaching sessions, as well as digital written materials. All material will be made available via Moodle (the Virtual Learning Environment used across the University) and Microsoft Teams (e.g. for live streaming, video calls, webinars, remote labs). Further information on both Moodle and Teams is given in Section 3 of this handbook.

For students on campus we will deliver some teaching sessions in person. These will be predominantly computer lab sessions. These smaller group sessions will give you the opportunity to discuss problems and check your understanding with one of our experienced staff members.

We realise that some students will not be able to arrive at Lancaster for the start of term. If this applies to you, then you will still have access to all online materials and teaching sessions (both synchronous and asynchronous) and we also plan to run equivalent remote sessions to all in-person sessions offered on campus.

All students will be taught on campus from January 2021. Further information will be provided nearer the time on the format of teaching sessions in the Lent Term, as this will depend largely on the guidance being given by the UK Government at that time, and also on the central approaches to teaching adopted by the University. All programme-level information will be conveyed to students via Moodle announcements on the programme web-board. Module-level information will be provided via the individual module pages on Moodle.

8

3. Core Information for Lancaster University PG Students

3.1 Central Support Services

This course handbook is your first point of call for all programme level information whilst you are studying on the MSc Data Science programme. Lancaster University has a wide range of advice and support services for students; these are provided by Student and Education Services and can be found on the main web page.

Key websites include:

- The Base - home for all student-based services:
- The University's Immigration and Visa Team for international students:
- On-campus health services
- Well-being and Mental Health
- Student finance including fees and living costs
- International student support

Key email addresses:

- The Base: thebase@lancaster.ac.uk
- Immigration and Visa Team: visa-advice@lancaster.ac.uk
- Counselling and Mental Health: counselling@lancaster.ac.uk

For those students based on campus, the Learning Zone is located centrally on Alexandra Square and is accessible 24-7. It provides relaxed surroundings for students to work within and bookable 'pods' for meetings, presentations and group work.

3.2 IT support

IT support is provided by Information Systems Services (ISS)

Specific information on services for postgraduate students

ISS contact information

Some specific services that you may wish to have support in are

Student Portal

This provides an online gateway to your modules (see Moodle below), university email account, student news and much more.

Moodle

This is the virtual learning environment that all lecturers will use to disseminate teaching materials, coursework assessments. It is also where you will submit coursework online. Each module has its own Moodle page, which you have access to once you have been enrolled on it.

Email

You are expected to check your Lancaster University email address at least once a day, as this is the way in which lecturers and other members of staff will contact you. Email can be accessed using Outlook Online or the mobile Outlook App.

The iLancaster App

ILancaster provides easily accessible information on services on and around campus, as well as access to your timetable and teaching information. There are versions for Android, iPhone/iPad and a Desktop version.**Further help and information.**

Teams

Microsoft Teams will be where many of your live online teaching sessions take place, including online workshops, webinars and examples classes. Teams can be used for both voice and video calls, and also has facilities for storing files and can be used as an area for backing up your work.

Apps Anywhere

Is used on Windows devices to access many widely used pieces of software.

3.3 Library and Learning Development

The Faculty of Science and Technology has a learning development team who support students to develop effective study practices and improve their writing skills. Further information and contact details can be found on their website at

A dedicated Moodle page for learning development

You can access support for literature searching and using library resources via the Faculty Librarian team. Further information and contact details can be found by self-enrolling on the Library skills for Science and Technology Moodle or the website

To find out more about the range of library resources available for students studying Data Science at the University, you can access library subject guides

Mathematics and Statistics subject guide
Computing and Communications subject guide

3.4 English Language Support

Dr Helen Hargreaves is the Learning Developer for English for Academic Purposes across the university. Her provision includes:

- Regular workshops on different aspects of English for academic study
- English language one-to-ones
- Opportunities to develop spoken English
- Online resources

For up-to-date information and to access resources, enrol on the English Language Development Moodle

3.5 Graduate College

All postgraduate students are members of Graduate College regardless of whether they are living in on-campus accommodation. The website for Graduate College contains a lot of information on study, social activities, support services and accommodation.

Specific details of college-based support systems

The College has a dedicated well-being officer.

For queries about accommodation in Graduate College, contact the Accommodation Team.

Full set of contact details for Graduate College.

3.6 Disabilities Service

The Disabilities Service provides advice and services for a whole range of disabilities. Where necessary, they will share information with departments to ensure that teaching and learning activities and assessments are made accessible to individuals with specific needs. You can find out more about the support services available on their website.

Sharing of information with the department is via an Inclusive Learning and Support Plan (ILSP) which is drawn up by the Disabilities Service and agreed with you. It is important that you are fully aware of any additional support that should be made available to you as part of your ILSP.

If you feel that you will require support at any level, then it is important that you register any disability

This link also provides guidance on what is classed as a disability and what information is required for registration.

If you feel that you are struggling to adjust to life as a student, either on or off-campus, the Transitions Team can provide support and advice.

The Transitions Team provides support to students all students, but particularly to those whose circumstances may make it more difficult for them to settle into University life.

3.7 Careers Advice

The University's Careers Service offers an extensive service tailored to your needs.

Their professional staff includes specialists in careers information, employer liaison, event management and careers guidance. They work closely with other staff within the university, the Students Union, professional bodies and a broad range of national and international employers to provide a variety of opportunities to help you progress your career goals. The Careers Service is in the Base, just off Alexander Square.

TARGETconnect is an online system administered by Careers and provides students with access to student and graduate vacancies, details of careers events, an appointment booking system to see a careers adviser and the online careers query system.

The Careers Service also runs bespoke careers workshops for Masters students.

3.8 The Students' Union & Student Support Office

Lancaster University Students Union (LUSU) is a body that represents all student views to the University, providing professional, academic and other advice for students. Students registering at Lancaster automatically become members of the Students' Union. There are no financial obligations associated with membership, though you can withdraw from the union if you wish, by completing an opt-out form. Further information on LUSU, including Societies, Services, Sport and Advice, is available on their website.

Contact details for LUSU.

For those based on campus, the LUSU Purple Card is a student discount card.

You will be able to pick up your card from the LUSU Welcome Desk in Bowland College, between University House and the Learning Zone.

4. Overview of the MSc Data Science programme

The role of a data scientist involves a diverse range of activities including research question formation, gathering and synthesis of data, data mining and visualisation, statistical-modelling, - inference and -validation, forecasting and prediction and dissemination of findings. This process is sequential, often cyclic in nature, and flows to form the data science pipeline:

- Research question formulation / provision
- Data identification and gathering
- Data processing and integration
- Model induction
- Model application and testing
- Results analysis
- Research question assessment

The image below shows the above points in a process flow diagram:
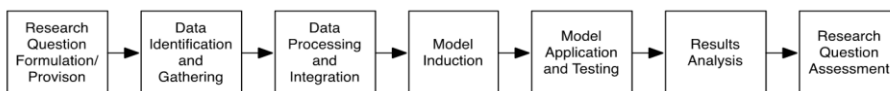
The wide variety of different skills and tools that a Data Scientist will work with is reflected both in the structure of the programme and the interdisciplinary curriculum.

The MSc programme consists of two terms of taught modules, a Summer Exam and assessment period, and a placement. At Lancaster teaching terms are ten weeks long: Michaelmas (weeks 1-10), Lent (weeks 11-20) and Summer (21-30). There is little formal teaching in the Summer Term; during the first half of the term you will be undertaking exams and/or projects for your taught modules. Then, from June to early September, you will undertake either a placement within an external organisation or a university-based research dissertation.

During the taught component of the programme, you will build on a common 'core' set of modules to develop a subject specialism in either Computing or Statistical Inference. Within the Statistical Inference specialism, there is a further opportunity to specialise further onto one of five pathways.

4.1 Specialisms

Computing specialism: aimed at students with a background in computer science who want to develop strong quantitative, data handling and analytical skills, along with skills that support the engineering of systems to support data science.

14

Statistical Inference specialism: aimed at students with a background in mathematics and statistics and who want to develop their statistical, computing and analytical skills for the extraction, synthesis, processing and analysis of large and complex data.

4.2 Pathways

For students on the Statistical Inference specialism, the Lent Term offers a wide variety of elective modules which cover advanced statistical modelling topics. The modules are grouped according to application domain. The groupings, known as pathways, are **Societal, Population Health**, **Environmental**, **Bioinformatics** and Business **Intelligence** pathways. Detailed course details, by specialism, are provided in Section 5 and Section 6. All module descriptions are provided in Appendix A.

5. MSc Data Science: Computing Specialism

Key Contact: Dr Chris Edwards (email address: c.edwards@lancaster.ac.uk), School of Computing and Communications

5.1 Computing specialism: overview

Underpinning the data scientist role are the technologies that enable the processing of data at large-scale, often using parallel processing paradigms.

The Computing specialism provides the training to understand how these technologies function and how they are implemented within both enterprise and research environments. Students will get hands-on experience of building, from scratch, large-scale systems that enable data science questions to be answered, using technologies.

The Computing specialism consists of a series of **taught modules** (120 credits) followed by the completion a **dissertation** (60 credits). The taught course component consists of 8 modules and can be decomposed as follows:

Five compulsory core modules **(75 credits)** spanning modern data science fundamentals: SCC460 Data Science Fundamentals (Michaelmas term & Lent term)

- SCC403 Data Mining (Michaelmas term)
- SCC461 Programming for Data Scientists (Michaelmas term)
- CFAS440 Statistical Methods and Modelling (Michaelmas term)
- CFAS406 Statistical Inference (Michaelmas term)

Three specialism modules **(45 credits):**

- SCC462 Distributed Artificial Intelligence (Lent)
- SCC413 Applied Data Mining (Lent)
- SCC411 Building Big Data Systems (Lent)

All module descriptions are, in order of mnemonic, contained in Appendix A. Modes of assessment and credit weightings for all modules follow in Tables 1 and 2 below.

The dissertation component **(60 credits)** consists of a substantial project applying data scientific methods to address a substantive research question. This will be undertaken during the summer term and will typically incorporate a ten-week industrial or research organisation placement.

5.2 Computing Specialism: Course Structure; Weightings and Assessment Strategy

Course Scheduling

The taught courses run during weeks 1 to 20. Core modules are taught in weeks 1 to 10. Specialism specific modules are taught in weeks 11 to 20 (Lent term). Open-book tests for CFAS440 and CFAS406 are held in May/June 2021. **Social distancing regulations may prevent these tests taking place on campus; contingency plans will be made for this and students given appropriate notice of any changes to assessment plans.**

Assessment

Computing Specialism core modules are assessed by either 100% coursework or a 50:50 split between coursework and an open-book test (Appendix A). All marks are provisional until they have been ratified by the Board of Examiners which meets in October. Credit for a module is given if the overall module mark is 50% or more.

Term 1 module weightings and mode of assessment. Please note: Exams are open-book for this specialism,

- Module number SCC460 (Data Science Fundamentals), running weeks 1-10 – coursework is 100% - no exam – credit weighting 15
- Module number SCC403 (Data Mining), running weeks 1-10 - coursework is 100% - no exam – credit weighting 15
- Module number SCC461 (Programming for Data Scientists), running weeks 1-10 - coursework is 100% - no exam – credit weighting 15
- Module number CFAS440 (Statistical Methods and Modelling), running weeks 1-5 - coursework is 50% - exam accounts for 50% – credit weighting 15
- Module number CFAS406 (Statistical Inference), running weeks 6-10 - coursework is 50% - exam accounts for 50% – credit weighting 15

Term 2 module weightings and mode of assessment

- Module number SCC462 (Distributed Artificial Intelligence), running weeks 11-20 – coursework is 100% - no exam – credit weighting 15
- Module number SCC413 (Applied Data Mining), running weeks 14-20 – coursework is 100% - no exam – credit weighting 15
- Module number SCC411 (Building Big Data Systems), running weeks 14-20 – coursework is 100% - no exam – credit weighting 15

Dissertation

For the academic year 20-21, **the deadline for submission of the dissertation is Friday 3rd September 2021**. Each dissertation will be double-marked, and a provisional mark agreed between the two markers and you will be asked to **attend a viva in person or via Teams** to defend your work. A copy of each dissertation and a brief report (including a provisional mark) agreed between the two internal markers will then be sent to the external examiner in advance of the final meeting of the Board of Examiners in October.

You will also be required to attend a poster conference, presenting work and findings of your dissertation, on Friday 10th September 2021. This conference is a final opportunity to get together as a group and celebrate learning achievements, not just over the dissertation period, but over the whole year.

Commented [LC1]: via Teams? also in stat inf specialism

Commented [E(2R1]: Yes, changed

Commented [E(3R1]: in both places

6. MSc Data Science: Statistical Inference Specialism

Entry Requirements: Whilst this specialism is open to all, some core and optional modules require a 2:1 in Mathematics and/or Statistics.
Key Contact: Dr Adam Sykulski (email address: a.sykulski@lancaster.ac.uk), Mathematics and Statistics

6.1 Statistical Inference Specialism: overview

The Statistical Inference specialism allows students to develop their statistical, computing and analytical skills for the extraction, synthesis, processing and analysis of large and complex data. The programme encompasses data science fundamentals but with an additional focus upon statistical modelling and inference of large and complex data structures and the underpinning theories. More specifically, the course provides a thorough 'core' training in statistical theory; data analysis and computing via a distinctive blend of leading-edge methodology and practical techniques (including Bayesian and computational methods and data mining) and 'optional' modules spanning, for example, longitudinal data analyses, time-to-event data, extreme value analysis, spatial data analyses and forecasting. The modules reflect inter-departmental research expertise and prepare students for career options in areas with a growing demand for data scientists. The elective modules form designated **Societal**, **Population Health, Environmental, Bioinformatics and Business Intelligence** pathways (Section 6).

This specialism consists of a series of **taught modules** (120 credits) followed by the completion a **dissertation** (60 credits). The taught course component consists of nine modules decomposed as follows:

1. A '**core**' set of five compulsory modules **(75 credits)** spanning modern data science fundamentals: SCC460 Data Science Fundamentals (Michaelmas & Lent term)
   - SCC403 Data Mining (Michaelmas term)
   - SCC461 Programming for Data Scientists (Michaelmas term)

   Students with a Mathematics and Statistics undergraduate degree will take

   - MATH552 Generalised Linear Modelling (Michaelmas term)
   - MATH551 Likelihood Inference (Michaelmas term)

   Students whose highest mathematical qualification is A-level (or equivalent) will take

   - CFAS440 Statistical Methods and Modelling
   - CFAS406 Statistical Inference

2. A compulsory specialism-specific module (**15 credits**): either MATH555 Bayesian Inference for Data Science or CFAS420 Statistical Learning.

3. A set of optional modules (**30 credits**) chosen from a number of electives delivered by a number of different departments: from the Department of Mathematics and Statistics, the CHICAS unit within the Lancaster Medical School, Lancaster University Management School and the Lancaster Environment Centre. All module descriptions are, in order of mnemonic, contained in Appendix A. Modes of assessment and credit weightings follow below.

4. A **dissertation** component (**60 credits**) consisting of a substantial project applying data scientific methods to address a substantive research question. This will be undertaken during the summer term and will typically incorporate a ten-week industrial or research organisation placement.

6.2 Statistical Inference: Course Structure; Weightings and Assessment Strategy

Course Scheduling

The taught courses run during weeks 1 to 20. Core modules are taught in weeks 1 to 10. Optional and specialism specific modules are taught in weeks 11 to 20 (Lent term). Exams are held in May/June 2021. **Social distancing regulations may prevent exams taking place on campus; contingency plans will be made for this and students given appropriate notice of any changes to assessment plans.**
Assessment

You will undertake a range of assessment methods. The credit weighting and balance between coursework and examination varies between modules. **Marks are provisional until ratified by the Board of Examiners which meets in June and October.** Final results will be sent to students in November 2021. Credit for a module is given if the overall module mark is 50% or more.

Dissertation Assessment

For the academic year 20-21, **the deadline for submission of the dissertation is Friday 3rd September 2021**. Each dissertation will be double-marked, and a provisional mark agreed between the two markers and you will be asked to **attend a viva in person or via Teams** to defend your work. A copy of each dissertation and a brief report (including a provisional mark) agreed

between the two internal markers will then be sent to the external examiner in advance of the final meeting of the Board of Examiners in October.

You will also be required to attend a poster conference, presenting work and findings of your dissertation, on Friday 10th September 2021. This conference is a final opportunity to get together as a group and celebrate learning achievements, not just over the dissertation period, but over the whole year.

Module weightings and mode of assessment; all students take SCC460, SCC403 and SCC461. Students either take CFAS440 and CFAS406 or MATH551 and MATH552 depending on academic background.[1][2]

- SCC460 (Data Science Fundamentals), running weeks 1-10 – coursework is 100% - no exam – credit weighting 15
- SCC403 (Data Mining), running weeks 1-10 – coursework is 100% - no exam – credit weighting 15
- SCC461 (Programming for Data Scientists), running weeks 1-10 – coursework is 100% - no exam – credit weighting 15
- MATH551 (Likelihood Inference), running weeks 1-5 – coursework is 40% - exam accounts for 60% – credit weighting 15
- MATH552 (Generalised Linear Models), running weeks 1-5 – coursework is 40% - exam accounts for 60% – credit weighting 15
- CFAS440 (Statistical Methods and Modelling), running weeks 1-5 – coursework is 50% - exam accounts for 50% – credit weighting 15
- CFAS406 (Statistical Inference), running weeks 6-10 – coursework is 50% - exam accounts for 50% – credit weighting 15
- MATH555[3] (Bayesian inference for Data Science), running weeks 11-20 – coursework is 30% - exam accounts for 70% – credit weighting 15
- CFAS420 (Statistical Learning), running weeks 16-19 – coursework is 100% - no exam – credit weighting 15
- MATH562 (Extreme Value Theory), running weeks 11-12 – coursework is 50% - exam accounts for 50% – credit weighting 10
- MATH563 (Clinical Trials), running weeks 11-12 – coursework is 50% - exam accounts for 50% – credit weighting 10
- MATH564 (Principles of Epidemiology), running weeks 13-14 – coursework is 50% - exam accounts for 50% – credit weighting 10
- MATH566 (Longitudinal Data Analysis), running weeks 15-16 – coursework is 50% - exam accounts for 50% – credit weighting 10
- MATH573 (Survival and Event History Analysis), running weeks 17-18 – coursework is 50% - exam accounts for 50% – credit weighting 10

---

[1] Students with a Mathematics and Statistics undergraduate degree should select MATH551 and MATH552; all other students will take CFAS440 and CFAS406.
[2] Note that timings of Lent Term modules are subject to change until the end of the Michaelmas Term
[3] Requires students to have taken MATH551 and MATH552 in the first term.

- CHIC563 (Geostatistical Modelling), running weeks 19-20 – coursework is 90% - exam accounts for 10% (test) – credit weighting 10
- CHIC571 (Infectious Disease Modelling), running weeks 14-15 – coursework is 100% - no exam– credit weighting 10
- BIOL445 (Bioinformatics), running weeks 17-18 – coursework is 50% - exam accounts for 50% – credit weighting 10
- MSCI523 (Forecasting), running weeks 11-20 – coursework is 100% - no exam– credit weighting 10
- MSCI526 (Introduction to Intelligent Data Analysis), running weeks 11-20 – coursework is 100% - no exam - credit weighting 10
- MSCI535 (Optimisation & Heuristics), running weeks 11-20 – coursework is 30% - exam accounts for 70% – credit weighting 10
- CFAS414 (Methods for Missing Data), running weeks 12-13 – coursework is 100% - no exam – credit weighting 10
- CFAS411 (Multilevel Modelling), running weeks 15-16 – coursework is 50% - exam accounts for 50% (in class test) – credit weighting 10
- LEC402 (Geoinformatics), running weeks 11-20 – coursework is 100% - no exam – credit weighting 15
- LEC468 (Modelling Environmental Processes), running weeks 11-15 – coursework is 50% - exam accounts for 50% – credit weighting 15

7. Designing your Programme of Study

The optional modules reflect both university research expertise and areas in which there is a high demand for data scientists. Elective modules are chosen to follow designated **Business Intelligence, Societal, Population Health, Bioinformatics and Environmental pathways**. The computing specialism also forms a designated pathway, with a systems and technologies focus. Table 3 groups Lent Term specialism-specific modules by pathway.

Subject to module pre-requisites, timetabling and credit constraints, modules may be selected across specialisms and pathways to form a bespoke programme of study, according to interests and aspirations. Any non-standard selection will need to be approved by or the Programme Director.

Designated Pathways: students choose modules to the value of 30 credits.

1. **Business intelligence**
   Forecasting (10 credits)
   Optimisation & Heuristics (10 credits)
   Introduction to Intelligent Data Analysis (10 credits)

2. **Societal**
   Multilevel Modelling (10 credits)
   Methods for Missing Data (10 credits)
   Principles of Epidemiology (10 credits)

3. **Population Health***
   Principles of Epidemiology (10 credits)
   Longitudinal Data Analysis (10 credits)
   Infectious Disease Modelling (10 credits)
   Survival and Event History Analysis (10 credits)
   Geostatistical Modelling (10 credits)

4. **Bioinformatics***
   Clinical Trials (10 credits)
   Principles of Epidemiology (10 credits)
   Bioinformatics (10 credits)

5. **Environment***
   Geoinformatics (15 credits)
   Modelling Environmental Processes (15 credits)
   Extreme Value Theory (10 credits)
   Geostatistical Modelling (10 credits)

Longitudinal Data Analysis (10 credits)

6. **Computing**
   Distributed Artificial Intelligence (15 credits)
   Building Big Data Systems (15 credits**)**

* to follow this designated pathway: choose modules totalling 30 credits from those available

Note that Extreme Value, Longitudinal Data Analysis, Survival and Event History Analysis and Environmental Epidemiology require MATH551 and MATH552. This subset form a set of study modules for those students wanting an increased mathematical focus.

8. Mode of Attendance

8.1 Full-time: One Year

The academic year runs from October to September and consists of Michaelmas, Lent and summer terms.   Lectures take place during the Michaelmas (October to December: weeks 1 - 10) and Lent (January to March: weeks 11 – 20) University terms. **Exams are held each year in May and June,** after which the dissertation period commences (June to September).

Modules delivered during Michaelmas term (weeks 1 to 10) are typically presented over a five- or ten-week period. Modules delivered in Lent term run variously, with many running intensively over two weeks comprising of an intensive period of lectures and labs (typically 20 hours over four days) followed by a period for a module specific project.

Full-time students are expected to be available for attendance at the University throughout the year**.**

8.2 Part-time: Two Years/Three Years

Students can study part-time over two or three years. The two-year arrangement is designed for students who can attend classes at the University on a regular basis. Such students will typically follow half of the taught modules that are available during each of the Michaelmas (weeks 1-10) and Lent (weeks 11-21) terms during each of their two years.   Alternatively, students may study full time during Michaelmas term of their first year and full time in the Lent term of their second year. Assessment is undertaken during the respective year of study alongside full-time students studying the module.   The three-year arrangement is to allow for increased flexibility. Subject to timetabling and module pre-requisites, a suitable split of the modules will be designed for individual students, recognising their other time constraints and previous knowledge.

Students will typically undertake the dissertation in the final year of study. Subject to skills and knowledge students may begin their dissertation earlier subject to approval with the Course Directors.

8.3 Attendance for Examinations, Dissertation Submission Dates, Assessment and Awards

Students must be present at the University for all written examinations that take place in May/June each year.   Part-time students undertake examinations for modules studied during the same examination period as full-time students. They

also submit the dissertation by the same deadline stipulated for full-time students in their final year of study. The scheme of credits, assessment and awards apply to both full- and part-time students.

All students are required to attend the university in September for the Dissertation Poster Session and to defend their work in a viva voce. The viva may be conducted remotely via Skype, for example.

Any resit assessments (coursework and examinations) take place during September after submission of the dissertation.

The Graduation Ceremony takes place mid-December each year.

9. Submission of Coursework and Feedback

Modules are offered across various departments. Each department has its own procedure for handling coursework and the module lecturer will provide information to you regarding the submission of coursework. All coursework is submitted online through the relevant module page on Moodle. Coursework will be checked electronically for plagiarism using the Turnitin software. Please ensure that you check and are aware of the coursework deadlines and submission procedures.

9.1 Plagiarism

Plagiarism occurs when you copy the work of another person and submit this as your own. Examples of plagiarism are copying model solutions from a previous year, copying work from another student, commissioning someone else (internal or external to the university) to complete a coursework for you or copying from online sources. Plagiarism is an extremely serious academic offence. If it is suspected that plagiarism has occurred, then the matter will be referred to the Departmental Academic Officer for investigation. If the Academic Officer believes that there is a case to answer then you be required to attend a hearing with the Academic Officer and the module lecturer. If it is found that plagiarism has occurred then a range of penalties can be applied, including a mark of zero for the coursework. A note of the offence will also be placed in your Student Record.

Further information on plagiarism can be found on the Student Educational Services website and in the Academic Malpractice Regulations and Procedures section of MARP (the University's Manual of Academic Regulations and Procedures). There is also information on appropriate referencing of sources. The Context section in the tutorial may be particularly helpful as it explains why we should reference.

9.2 Late submission of assignments and penalties

There are strict university-wide rules in place for the late submission of assignments. Full regulations can be found in MARP (see GR 2.3.5 and GR 2.3.6)

To summarise:

- Work submitted up to three days late without an agreed extension will receive a penalty of 10 percentage points (for example a mark of 62% would become 52%) and zero (non-submission) thereafter'.

- Saturdays and Sundays are included as days in this regulation. However, when the third day falls on a Saturday or Sunday, students will have until 10.00 a.m. on Monday to hand in work without receiving further penalty.

- For regular small components of work (e.g. weekly quizzes or lab worksheets) where it is necessary to provide feedback quickly, the three-day period mentioned above may be shortened. When this is the case, the module lecturer will make this clear.

Extension requests

To request a short extension on a small component of work e.g. a weekly quiz or problem sheet, please contact the module lecturer directly. Once feedback has been released, submissions will not be permitted, therefore it will not be possible to grant extensions in retrospect for small weekly work which have a short turn-around time between submission and feedback.

To request an extension on a project or longer piece of coursework on an SCC, MATH and CFAS modules, please email the Data Science Teaching email address. For MSCI, LEC, CHICAS and BIOL modules, lecturers will provide relevant contact details and extension procedures.

Be aware that for longer extensions, you will be asked to provide appropriate evidence, eg. A note from your GP or another medical/well-being professional.

Work submitted more than three days late without an agreed extension will be considered a non-submission, awarded a zero and treated according to the standard procedures for failed work.

9.3 Exceptional Circumstances

Exceptional circumstances refer to actions or events outside your control which result in any circumstances which can reasonable be shown to have affected your academic performance. Examples of such events might be a severe episode of poor mental or physical health, unexpected increase in caring duties or illness/death of a close relative or friend. Examples of possible impact are missing a coursework deadline (extended or otherwise), meeting a deadline but under-performing compared to usual standards or missing an exam.

It is your responsibility to make those circumstances known to the department by emailing either your Academic Tutor or the Data Science Teaching email address.

You should make every effort to inform the department within 48 hours of the missed deadline unless prevented from doing so by acceptable circumstances, in which case you should inform the department as soon as possible. You will need to provide the department with evidence of the exceptional circumstances as soon as they are able.

When informed of Exceptional Circumstances in reasonable time, the department will try to mitigate for these at the time. Occasionally it is not possible to mitigate fully. In this case, you may submit an Exceptional Circumstances case to the Exceptional Circumstances Committee which meets in Week 29 and again in before the final Exam Board in October. Details on Exceptional Circumstances, and the relevant form, can be found on the Lancaster University website.

9.4 Feedback and Notification of Marks

Feedback on assessed work will typically be provided within four weeks of submission (excluding vacations and unforeseen staff absences). Once coursework has been marked and the marks recorded by the Student Office, both marks and feedback will be released via Moodle.

Exam papers are marked in line with university timescales. It is University policy not to return examination scripts to students.

It should be remembered that until the External Exam Board has met, (October each year) any marks given to students are **provisional and may be subject to change.** The External Exam Board does not usually meet until around 6 weeks after the end of the programme.

As per the University regulations there is no appeal against academic judgement.

10.   Student Feedback Mechanisms

10.1 Module Evaluation

Module evaluation takes place via the Lancaster University Module Evaluation System (LUMES). You will be contacted by email at the end of each module and asked to complete a module evaluation form online. Evaluation is more than feedback of the good and bad elements of a module. It provides continual information for us to improve the content and delivery of the modules we offer. Consequently it is very important that all students actively participate in the evaluation process. We ask that you respond to the request to submit your feedback as soon as possible. **Please note that your feedback is anonymous.**

For 2020/21, student feedback will be particularly vital to us as we aim to adapt and improve our blended learning delivery and assessment methods. If you have any comments or concerns then please contact the Programme Directors on the Data Science Teaching email address

10.2 Academic Representatives

Academic Representatives are invited to sit on the departmental Staff-Student Consultative Committees in both the Department of Mathematics and Statistics and SCC. This is the official mechanism for communication between students and academic staff. The committees are formed in accordance with University Regulations, which require staff-student consultation prior to programme changes. They also provide a way in which all students on the programme can feed back to the departments with either programme- or module-level concerns.

The committees meet at least once per term as a consultative body to discuss current programmes and future proposed changes. Dates for the Committees and the minutes of each meeting will be made available on Moodle by the departmental Student Offices.

For further details on the roles and responsibilities of Academic Representatives see the Lancaster University Students Union (LUSU) web-site

A request for Academic Representatives will be circulated during Welcome Week. All volunteers will be required to undertake a short training session with LUSU.

11.   Programme Rules and Requirements for Awards

11.1 Summary

Criteria for taught postgraduate awards (MSc, PG Diploma and PG Certificate) are given in Section 3 of the Postgraduate Taught Assessment Regulations

Section 5 of the above outlines the rules for classification of each award using the Fail, Pass, Merit and Distinction scale. The key points are highlighted below.

Programme Rules:

The award of MSc requires 120 credits worth of taught courses and 60 credits from a dissertation

The award of PG Diploma requires 120 credits of taught courses.

The award of PG Certificate requires 60 credits of taught courses.

Module Rules:

The pass mark for each module contributing to a taught MSc, PG Diploma or PG Certificate is 50%. Credit from a module can only contribute towards an award if the overall mark for the module is 50% or greater. The mark for each module is given by a stipulated combination of written examination and/or coursework.

Awards:

The overall mark is the weighted average of the marks gained in the taught modules and the dissertation for the award of MSc in Data Science, and in the taught courses only for the PG Diploma and the PG Certificate.

11.2 MSc in Data Science

To qualify for the Masters' degree in Data Science you must achieve 180 credits: 120 credits from the taught course component and 60 credits from the Dissertation. Credit for a module is given only if the overall module mark is 50% or more.

Condonation:

Notwithstanding this requirement, candidates shall be eligible for an award by compensation/condonation in respect of up to a maximum of 45 credits of a taught Masters programme provided that:

a) no single module mark falls below 40%;
b) the candidate's weighted mean is 50% or greater.

Higher awards:

MSc in Data Science with 'Merit'
In addition to the requirements for the MSc in Data Science, an 'MSc in Data Science with Merit' shall be awarded when the weighted average mark over the programme is 60% or more.

MSc in Data Science with 'Distinction'

In addition to the requirements for the award of the MSc in Data Science, an `MSc in Data Science with Distinction' is awarded when the weighted average mark over the programme is 70% or more.

Borderline Cases

Subject to the standard condonation rules above, where the overall weighted average falls within two percentage points of the Merit / Distinction degree award range (i.e. 58% / 68%, respectively) the following rules for degree awards will apply:

58.0% to 59.9%:

If more than 50% of module credits (i.e. > 90 credits) are at `Merit level' (i.e. are at 60% or above) the MSc in Data Science will be awarded with Merit.

68.0% to 69.9%:

If more than 50% of module credits (i.e. > 90 credits) are at Distinction level (i.e. at 70% or above) the MSc in Data Science will be awarded with Distinction.

11.3 Postgraduate Diploma in Data Science

To qualify for the Postgraduate Diploma in Data Science you must achieve a total of 120 from the taught courses. Credit for a module is given only if the overall module mark is 50% or more.

Condonation

Notwithstanding this requirement, candidates shall be eligible for an award by compensation/ condonation in respect of up to a maximum of 30 credits provided that:

   a) no single module mark falls below 40%;
   b) the candidate's weighted mean is 50% or greater.

Higher awards:

Postgraduate Diploma in Data Science with `Merit'
In addition to the requirements for the Postgraduate Diploma in Data Science, a 'Postgraduate Diploma in Data Science with Merit' shall be awarded when the weighted average mark over the programme is 60% or more.

Postgraduate Diploma in Data Science with `Distinction'

In addition to the requirements for the award of the Postgraduate Diploma in Data Science, a `Postgraduate Diploma in Data Science with Distinction' shall be awarded when the weighted average mark over the programme is 70% or more.

11.4 Postgraduate Certificate in Data Science

To qualify for the Postgraduate Diploma in Data Science you must achieve a total of 60 from the taught courses. Credit for a module is given only if the overall module mark is 50% or more.

Condonation

Notwithstanding this requirement, candidates shall be eligible for an award by compensation/ condonation in respect of up to a maximum of 20 credits provided that:

   1. no single module mark falls below 40%;
   2. the candidate's weighted mean is 50% or greater.

Higher awards:

Postgraduate Certificate in Data Science with `Merit'
In addition to the requirements for the Postgraduate Certificate in Data Science, a 'Postgraduate Certificate in Data Science with Merit' shall be awarded when the weighted average mark over the programme is 60% or more.

Postgraduate Certificate in Data Science with `Distinction'

In addition to the requirements for the award of the Postgraduate Diploma in Data Science, a `Postgraduate Certificate in Data Science with Distinction' shall be awarded when the weighted average mark over the programme is 70% or more.

11.5 Re-sits

A student who fails to achieve a mark of 50% for a module is entitled to one opportunity for reassessment in each failed module. Reassessment is compulsory for any module mark below 40%. Whilst reassessment is optional for module marks within the condonable range (at least 40% but below 50%), it may be necessary for some or all modules in this range to be reassessed if either the overall average over all modules lies below 50% or more than the stipulated number of credits for the desired award lie in this range e.g. for an MSc award this would be more than 45 credits.

A mark of not more than 50% can be given for modules re-taken (this is referred to as a capped re-sit).

The form of the reassessment is at the absolute discretion of the Examination Board, save that the form of reassessment must allow the student a realistic chance of achieving 50% in the re-sit.

Resit assessments, both coursework and examinations, take place in September each year. Students undertaking re-sit exams will need to be present at the university.

11.6 Notification of Final Degree Marks

The External Exam Board meets in late October to recommend awards. Final marks are released to students as soon as possible, thereafter. **Please note that the University Regulations state that written confirmation of results, provisional and final, may not be released to students who are in debt to the University.**

The Board of Examiners

The Board of Examiners consists of the:

- External Examiner for the MSc.
- Heads of Department;
- Course Directors;
- Course Tutor;
- Placement Officer;

- Assessment Officers;
- Postgraduate Teaching Coordinators.

The External Examiner for the period 2020-2021 is Professor Peter Triantafillou, Department of Computer Science, University of Warwick.

Graduation

The Postgraduate Graduation Ceremony will be in December 2021. Information regarding Graduation will be sent to you from the University Ceremonies Office.

Please note that it is essential that you keep your contact details address up-to-date in order to receive the relevant graduation mailings.

12.   Module Descriptions

Prerequisites for Lent Term modules are stated only where choice of Michaelmas Term statistics modules will affect access to that course, namely for those modules only available to students who have taken MATH551/MATH552. Part-time students should discuss with their Academic Tutor or the relevant module convenors to ascertain suitability of Lent Term modules for each year of study. Details of the learning outcomes and syllabus for each module can be found in the online programme and module catalogue which can be accessed via the Student Portal.

BIOL445 Bioinformatics

Module Convenor: Dr Derek Gatherer
Assessment: Exam (50%) and coursework (50%)
Credits: 15
Software used: MEGA, DNASp, BEAST, FigTree, Tracer, SimPlot, SwissModel, Galaxy, Artemis, EMBOSS

CFAS406 Statistical Inference

Module Convenor: Dr Clement Lee
Assessment: Coursework (50%; consisting of individual project 25% and group project 25%) and open-book class test (50%)
Credits: 15
Term: Michaelmas
Software used: R

CFAS411 Multi Level Models

Module Convenor: Dr Fang Wan
Assessment: Group project (50%) and open-book class test (50%)
Credits: 10
Term: Lent
Prerequisites: (MATH551 and MATH552) OR (CFAS406 and CFAS440))
Software used: R

CFAS414 Methods for Missing Data

Module Convenor: Dr Gareth Ridall
Assessment: Coursework (100%; consisting of group project 50% and individual project 50%)
Credits: 10

Term: Lent
Software used: R (packages: VIM, mice, amelia 2, miss MDA)

CFAS420 Statistical Learning

Module Convenor: Dr Alex Gibberd
Assessment: Coursework (100%; 30% group project and 70% individual project)
Credits: 15
Term: Lent
Software used: R, SPSS, Latent Gold

CFAS440 Statistical Methods and Modelling
Module Convenor: Dr Kanchan Mukherjee
Assessment: Coursework (50%; consisting of individual project 20% and group project 30%) and open-book class test (50%)
Credits: 15
Term: Michaelmas
Prerequisites:
Software used: R

CHIC563 Geostatistical Modelling

Module Convenor: Dr Emanuele Giorgi
Assessment: Coursework (90%) and Test (10%)
Credits: 10
Term: Lent
Software used: R

CHIC571 Modelling of infectious diseases
Module Convenor: Dr Chris Jewell
Assessment: Group presentation (20%) and project (80%)
Credits: 10
Term: Lent
Software used: R

LEC402 Geoinformatics

Module Convenor: Prof Alan Blackburn
Assessment: 100% Coursework (Literature Review 50%; Report 50%)
Credits: 15
Term: Lent
Software used: ArcGIS & ERDAS Imagine

LEC468 Modelling Environmental Processes

Module Convenor: Dr Wlodek Tych
Assessment: Coursework (50%) and Exam (50%)
Credits: 15
Term: Lent
Software used: Matlab/Simulink package

MATH551 Likelihood Inference

Module Convenor: Dr Chris Sherlock
Assessment: Moodle quizzes (20%), open-book test (20%) and Exam (60%)
Credits: 15
Term: Michaelmas
Prerequisites: UG Mathematics/Statistics (probability theory; calculus; matrices etc)
Software used: R

MATH552 Generalised Linear Modellling

Module Convenor: Dr Juhyun Park
Assessment: Moodle Quizzes (15%), Group Project (25%) and Exam (60%)
Credits: 15
Term: Michaelmas
Prerequisites: UG Mathematics/Statistics (probability theory; calculus; matrices etc)
Software used: R

MATH555 Bayesian Inference for Data Science

Module Convenor: Dr Marco Battison
Assessment: Coursework (30%) and Exam (70%)
Credits: 15
Term: Lent
Prerequisites: MATH551, MATH552
Software used: R

MATH562 Extreme Value Theory

Module Convenor: Dr Emma Eastoe
Assessment: Coursework (50%; weekly exercises 20% and project 30%) and Exam (50%)

Credits: 10
Term: Lent
Prerequisites: MATH551, MATH552
Software used: R

MATH563 Design and Analysis of Clinical Trials

Module Convenor: Dr Fang Wan
Assessment: Coursework (50%) and written exam (50%)
Credits: 10
Term: Lent
Software used: R

MATH564 Principles of Epidemiology
Module Convenor: Dr Simon Lunagomez
Assessment: Coursework (50%; 15% Moodle quizzes, 35% project) and written exam (50%)
Credits: 10
Term: Lent
Software used: R

MATH566 Longitudinal Data Analysis

Module Convenor: Dr Fang Wan
Assessment: Group project (40%) and Exam (60%)
Credits: 10
Term: Lent
Prerequisites: MATH551; MATH552
Software used: R

MATH573 Survival and Event History Analysis

Module Convenor: Dr Kanchan Mukherjee
Assessment: Coursework (50%) and written exam (50%)
Credits: 10
Term: Lent
Prerequisites: MATH551; MATH552
Software used: R

MSCI523 Forecasting

Module Convenor: Dr Sven Crone
Assessment: Coursework (100%)
Credits: 10

Term: Lent
Software used: TBC

MSCI526 Introduction to Intelligent Data Analysis

Module Convenor: Dr Nicos Pavlidis
Assessment: Coursework (100%)
Credits: 10
Term: Lent
Software used: R

MSCI535 Optimisation and Heuristics

Module Convenor: Prof Adam Letchford
Assessment: Coursework (30%) and Exam (70%)
Credits: 10
Term: Lent
Software used: MPL, LINDO and EXCEL SOLVER etc

SCC.462 Distributed Artificial Intelligence

Module Convenor: Dr Leandro Soriano Marcolino
Assessment: Coursework (100%)
Credits: 15
Term: Lent
Software used: Optional; Linux environment

SCC.403 Data Mining

Module Convenor: Prof Plamen Angelov
Assessment: Coursework (100%)
Credits: 15
Term: Michaelmas
Software used: Python

SCC.411 Building Big Data Systems

Module Convenor: Dr Yehia Elkhatib and Dr Peter Garragan
Assessment: Coursework (100%)
Credits: 15
Term: Lent
Software used:

SCC.413 Applied Data Mining

Module Convenor: Dr Alistair Baron
Assessment: Coursework (100%)
Credits: 15
Term: Lent
Software used: Python

SCC.460 Data Science Fundamentals

Module Convenor: Dr Ioannis Chatzigeorgiou, Dr Keith Cheverst
Assessment: Coursework (100%)
Credits: 15
Term: Michaelmas

SCC.461 Programming for Data Science

Module Convenor: Dr Clement Lee, Dr Leandro Soriano Marcolino
Assessment: Coursework (100%)
Credits: 15
Term: Michaelmas
Software used: R, Python, LaTeX