



**UNIVERSITY
OF LONDON**

Programme Regulations 2021–2022

Data Science

MSc

PGDip

PGCert

Individual modules

**Data Science and Financial
Technology**

MSc

PGDip

Individual modules

**Data Science and Artificial
Intelligence**

MSc

PGDip

Individual modules

Important document – please read
This document contains important
information that governs your
registration, assessment and
programme of study



Contents

Important information regarding the Programme Regulations	2
1 Structure of the programmes.....	5
2 Registration.....	8
3 Recognition of prior learning and credit transfer.....	9
4 Assessment for the programme.....	10
5 Transfer of Registration.....	16
6 Scheme of award.....	17
Appendix A – Structure of the programmes.....	19
MSc Data Science	19
PGDip Data Science	20
PGCert Data Science.....	21
MSc Data Science and Financial Technology	22
PGDip Data Science and Financial Technology.....	23
MSc Data Science and Artificial Intelligence	24
PGDip Data Science and Artificial Intelligence.....	25
Appendix B – Module Outlines.....	26
Appendix C – Assessment Criteria	32

Important information regarding the Programme Regulations

About this document

Last revised 12 April 2021

As a student registered with the University of London you are governed by the current General Regulations and Programme Regulations associated with your programme of study. These Programme Regulations are designed and developed by the Goldsmiths, which is responsible for the academic direction of the programme, in collaboration with the University of London. Further information about how to use the Programme Regulations and Programme Handbook can be found in the Student Guide.

In addition to Programme Regulations you will have to abide by the General Regulations. These regulations apply to all students registered for a programme of study with the University of London and provide the rules governing registration and assessment on all programmes; they also indicate what you may expect on completion of your programme of study and how you may pursue a complaint, should that be necessary. Programme Regulations should be read in conjunction with the General Regulations.

The relevant General Regulations and the Programme Regulations relating to your registration with us are for the current year and not the year in which you initially registered. On all matters where the regulations are to be interpreted, or are silent, our decision will be final.

Further information about your programme of study is outlined in the Programme Specification which is available on the relevant Courses page of the website. The Programme Specification gives a broad overview of the structure and content of the programme as well as the learning outcomes students will achieve as they progress.

If you have a query about any of the programme information provided please contact us. You should use the 'ask a question' button in the student portal <https://my.london.ac.uk/>

To note

Throughout the Regulations, 'we' 'us' and 'our' mean the University of London; 'you' and 'your' mean the student, or where applicable, all students.

Terminology

For the programmes offered under these regulations, you should note the following:

Module: Individual units of a programme are called modules. Each module is a self-contained, formally structured learning experience with a coherent and explicit set of learning outcomes and assessment criteria.

Core module: Core modules are central to the teaching and learning on the programme and often introduce concepts and ideas that appear in the optional modules. Core modules must be passed and cannot be compensated.

Compulsory module: Compulsory modules introduce concepts and ideas that appear in optional modules. Students must take these modules as part of their studies. They are subject to the rules for compensation.

Optional module: Optional modules are designed to extend the concepts and ideas introduced in core and compulsory modules and to introduce other relevant concepts and techniques. Students may select their optional modules from a list. They are subject to the rules for compensation.

Preparatory module: a non-credit bearing, optional module available to students to ensure readiness to start core and compulsory modules.

Significant changes made to the programme regulations 2021-2022

General changes

- Changes to maximum number of attempts at an element of assessment:

For students who first registered in 2019-20 and October 2020: The maximum number of attempts permitted for an element of coursework, written examination, or project report is **two**, excluding module [DSM020]. Students who first registered in 2019-20 and October 2020 will be permitted one additional attempt at [DSM020].

- **For students who first registered from April 2021:** The maximum number of attempts permitted for an element of coursework, written examination, or project report is **two**.

Changes to progression

- **Students will now be considered for progression to the following routes, prior to** final module results being ratified by the exam board and released: the **PGCert**, PGDip and MSc. This applies as follows:
- Students registered on the PGCert or PGDip intending to progress to the PGDip or MSc will be considered for progression once they have passed the first element of assessment on their modules, eg. Coursework. If the student wishes to transfer to a higher award before their final result is ratified and released, they must log an enquiry and notify us as soon as possible, in order to avoid delays to their progression.
- Students registered on individual module(s) such as DSM020, as an entry point to a higher award will be considered for progression once they have passed the first element of assessment on their module(s).

Students who first register for the April 2022 session may be offered a performance-based admission, with the requirement to study one or more modules and pass its first element of assessment before being permitted to attempt the rest of the modules on the programme.

New modules

DSM030 Statistics and Statistical Data Mining

Closed modules

DSM030 Mathematics and statistics for data science

Programme changes

1. *DSM030 Mathematics and statistics for data science* has now closed and is replaced by *DSM030 Statistics and Statistical Data Mining*. Students who first registered in 2019-20 who have passed DSM030 will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining towards their final award.
2. The topics and learning outcomes for *DSM020 Data programming in Python* have changed.
3. There are significant changes to the assessment of each Data Science module. The updated weightings for each module are reflected in summary table of assessment.

Introduction of preparatory course: *Foundations of Data Science: K-Means Clustering in Python*

Alternative Assessments during the Coronavirus (COVID-19) Outbreak

In line with our current General Regulations, the University may offer you alternative assessments where necessary. This includes holding online timed assessments in place of written examinations, which are usually held at examination centres. Please note that this statement replaces any published information relating to assessments or written examinations in any of our materials including the website. Previously published materials relating to examinations should therefore be read in conjunction with this statement.

The University of London continues to work towards supporting the academic progression of all its students. The University also continues to be mindful of the health and wellbeing of its students during this pandemic, whilst protecting the academic standards of its awards.

1 Structure of the programmes

Qualifications

1.1

The following qualifications are offered under these regulations:

- Master of Science in Data Science (MSc)
- Postgraduate Diploma in Data Science (PGDip)
- Postgraduate Certificate in Data Science (PGCert)
- Master of Science in Data Science and Financial Technology (MSc)
- Postgraduate Diploma in Data Science and Financial Technology (PGDip)
- Master of Science in Data Science and Artificial Intelligence (MSc)
- Postgraduate Diploma in Data Science and Artificial Intelligence (PGDip)

Qualification structures

For further information on the modules for each programme, see [Appendix A](#).

Data Science

1.2

For the award of **MSc Data Science** you must complete:

- four core modules (60 credits total)
- two compulsory modules (30 credits total)
- four optional modules (60 credits total)
- a Final Project (30 credits total)

1.3

For the award of a **PGDip Data Science** you must complete:

- four core modules (60 credits total)
- two compulsory modules (30 credits total)
- two optional modules (30 credits total)

1.4

For the award of a **PGCert Data Science** you must complete:

- two core modules chosen from a list (30 credits total)
- any two other modules (either core or optional modules) (30 credits total)

Data Science and Financial Technology

1.5

For the award of **MSc Data Science and Financial Technology** you must complete:

- four core modules (60 credits total)
- three compulsory modules (45 credits total)
- three optional modules (45 credits total)
- a Final Project (30 credits total)

1.6

For the award of **PGDip Data Science and Financial Technology**, you must complete:

- four core modules (60 credits total)
- three compulsory modules (45 credits total)
- one optional module (15 credits total)

Data Science and Artificial Intelligence

1.7

For the award of **MSc Data Science and Artificial Intelligence**, you must complete:

- four core modules (60 credits total)
- three compulsory modules (30 credits total)
- three optional modules (60 credits total)
- a Final Project (30 credits total)

1.8

For the award of **PGDip Data Science and Artificial Intelligence**, you must complete:

- four core modules (60 credits total)
- three compulsory modules (30 credits total)
- one optional module (30 credits total)

1.9

The maximum number of modules you can register for at any one time is six 15 credit modules, or four 15 credit modules and the 30 credit Final Project.

Over a 22 week session, a 15 credit module will typically require five to seven hours of work/effort per week, and a 30 credit module will typically require ten to 15 hours of work/effort per week.

Intermediate qualifications

1.10

If you are registered on the **PGDip or MSc Data Science**, you can apply for a PGCert or a PGDip Data Science as an intermediate qualification, providing you have successfully completed the required modules.

1.11

If you are registered on the **PGDip or the MSc for either of the specialisms**, you can request a PGCert Data Science or a PGDip relating to your specialism as an intermediate qualification, providing you have successfully completed the required modules.

See [Appendix A](#) for the award structures.

1.12

If you satisfy the necessary requirements and wish to receive a PGDip or PGCert as an intermediate qualification, you **must apply** to be granted those qualifications. A PGDip or PGCert will **not** be awarded automatically as an intermediate qualification.

1.13

If you do not apply for the PGDip or the PGCert as an intermediate qualification at, or prior to, your final examination entry for the programme you have registered for, you will **not** be awarded these qualifications at a later date.

1.14

Introduction of preparatory course: *Foundations of Data Science: K-Means Clustering in Python*

Individual modules taken on a stand-alone basis

1.15

There is provision for individual modules to be studied and assessed on a stand-alone basis without being registered for a related qualification. You may register for any number of core or optional modules on a stand-alone basis (subject to module availability), with the exception of the Final Project. However, a maximum of eight modules (120 credits) may be counted as credit towards an MSc qualification, a maximum of four modules (60 credits) may be counted towards a PGDip qualification and a maximum of two modules (30 credits) may be counted towards a PGCert qualification.

1.16

Credit for the stand-alone individual modules will be considered provided that the application for a related qualification is made within three years of the completion of the relevant modules. Neither progression nor credit is automatic.

1.17

The following Data Science modules are available on a stand-alone basis:

DSM010	Big data analysis
DSM020	Data programming in Python
DSM030	Statistics and statistical data mining
DSM040	Machine learning

DSM050	Data visualisation
DSM060	Data science research topics
DSM070	Blockchain programming ♠
DSM080	Mathematics of financial markets ♠
DSM100	Artificial intelligence ♦
DSM110	R for data science ♠
DSM120	Financial data modelling ♠
DSM140	Natural language processing ♦
DSM150	Neural networks ♦
DSM160	Social networks and graph analysis ♦

♠ *Modules available in April 2021*; ♦ *Modules available in October 2021*

2 Registration

Full details of the Entrance Requirements for the programme are listed on the requirements tab of the [website](#)

Effective date of registration

2.1

Your effective date of registration will be either:

- 1 October if you first register before the September registration deadline
- 1 April if you first register before the March registration deadline

Date of first examinations

2.2

If your date of registration is:

- 1 October, you will take your first examinations in March of the following year
- 1 April, you will take your first examination in September of the same year

Period of registration

The minimum and maximum periods of registration to complete the programme are given in the [Programme Specification](#).

2.3

If you progress from the PGCert or PGDip to the PGDip or MSc respectively, your maximum period of registration will continue from the date you initially registered under these regulations.

2.4

If you progress from individual modules to the PGCert, PGDip or MSc, you will be given a new period of registration beginning from when you register on the PGCert, PGDip or MSc.

2.5

If you transfer between the Data Science programmes and a specialism, you will not be granted a new period of registration, it will continue from the date you initially registered under these regulations.

Order in which to take the modules:

2.6

For all programmes, you can attempt the modules in any order, subject to module availability.

Whilst you can attempt modules in any order, you are advised to have attempted the assessments for all other modules before registering for the Final Project.

Progression within the programme

2.7

Upon successful completion of the PGCert or a PGDip offered under these regulations you may progress to a PGDip or a related MSc programme respectively. This also applies to students offered an individual module e.g. DSM020, as an entry point, permitting them to transfer with full credit on to any of the following routes: PG Certificate in Data Science, PGDip in Data Science or MSc in Data Science.

2.8

You will be considered for progression to the following routes, when you have passed the first element of assessment for your module(s): PGCert, PGDip and MSc. This applies as follows:

- Students registered on the PGCert or PGDip intending to progress to the PGDip or MSc will be considered for progression once they have passed the first element of assessment on their modules, eg. Coursework. If the student wishes to transfer to a higher award before their final result is ratified and released, they must log an enquiry and notify us as soon as possible, in order to avoid delays to their progression.
- Students registered on individual module(s) such as DSM020, as an entry point to a higher award will be considered for progression once they have passed the first element of assessment on their module(s).
- Students who first register for the April 2022 session may be offered a performance-based admission, with the requirement to study one or more modules and pass its first element of assessment before being permitted to attempt the rest of the modules on the programme.

3 Recognition of prior learning and credit transfer

The rules for recognition of prior learning are set out in the General Regulations.

3.1

If you are registered for an MSc programme, you may apply for recognition of prior learning (RPL) mapped against modules up to a total of 120 UK credits.

3.2

Recognition of prior learning will not be considered for the Final Project.

3.3

If you are registered for a PGDip programme, you may apply for recognition of prior learning mapped against modules up to a total of 60 UK credits.

3.4

If you are registered for the PGCert Data Science, you may apply for recognition of prior learning mapped against modules up to a total of 30 UK credits.

4 Assessment for the programme

In line with our current General Regulations, the University may offer you alternative assessments where necessary. This includes holding online timed assessments in place of written examinations, which are usually held at examination centres. Please refer to Programme Specifications for details on equipment that may be required for alternative assessments.

4.1

Modules are assessed by either a combination of coursework elements or a combination of written examination element and coursework element. The grade awarded for each module will be based on all the elements of assessment.

4.2

The pass mark for any element of assessment is 50%.

4.3

You will be required to pass each element of the assessment in order to pass a module overall, subject to the application of the rules for compensation.

Rules for compensation

4.4

We will allow compensation for an assessment element within **optional and compulsory** modules if both:

- the mark awarded for the assessment element is between 45%-49% *and*
- the mark for the other assessment element is sufficient to produce an overall combined weighted pass mark for the module.

4.5

We will **not** allow compensation for an assessment element within core modules or the Final Project.

4.6

Summary table of assessment:

Programme structure			
Modules		Assessment Element 1	Assessment Element 2
Core modules	[DSM010] - Big data analysis	Coursework item 1: 30%	Coursework item 2: 70%
	[DSM020] - Data programming in Python	Coursework item 1: 30%	Coursework item 2: 70%
	[DSM030] - Statistics and statistical data mining	Coursework: 50%	Written examination: 50%
	[DSM040] - Machine learning	Coursework item 1: 50%	Coursework item 2: 50%
Compulsory modules	[DSM050] - Data visualisation	Coursework item 1: 30%	Coursework item 2: 70%
	[DSM060] - Data science research topics	Coursework item 1: 50%	Coursework item 2: 50%
Optional modules	[DSM070] – Blockchain programming	Coursework 80% (overall) split as follows: Item 1: 30% Item 2: 50%	Written examination: 20%

Programme structure		
Modules	Assessment Element 1	Assessment Element 2
[DSM080] - Mathematics of financial markets	Coursework: 50%	Written examination 50%
[DSM100] - Artificial intelligence	Coursework item 30% (overall) split as follows: Item 1: 15% Item 2: 15%	Written examination: 70%
[DSM110] - R for data science	Coursework item 1: 30%	Coursework item 2: 70%
[DSM120] - Financial data modelling	Coursework: 50%	Written examination: 50%
[DSM140] - Natural language processing	Coursework item 30% (overall) split as follows: Item 1: 15% Item 2: 15%	Written examination: 70%

Programme structure			
Modules		Assessment Element 1	Assessment Element 2
	[DSM150] - Neural networks	Coursework item 30% (overall) split as follows: Item 1: 15% Item 2: 15%	Written examination: 70%
	[DSM160] - Social networks and graph analysis	Coursework item 30% (overall) split as follows: Item 1: 15% Item 2: 15%	Written examination: 70%
Final project	[DSM500] - A final project in data science	Final project coursework: 70% (overall) split as follows: Project proposal: 15% Final project: 55%	Written examination: 30%

Taking assessments

Please refer to the rules on assessment and assessment offences in the [General Regulations](#).

4.7

When you register for a module, you must take the assessments at the first available opportunity.

4.8

Written examinations take place on two occasions each year in September and March.

Deferring an assessment

The deadlines and process for deferring a written examination are given on the [VLE](#).

If you defer taking a written examination you have to pay a module continuation fee when you register for the session in which you wish to take the examination. You do not have to register for the next available session; there are two sessions each year and you have five years to complete your studies.

4.9

You can defer taking the written examination element of a module if you notify us by the deadline. You may only defer **once** per module. It is not possible to defer coursework submissions.

4.10

If you do not notify us of your deferral and do not attend the written examination you will be given a mark of zero for the assessment element and it will count as an attempt at the written examination.

4.11

If you defer the written examination, the mark for the coursework element of the module will be carried forward if it is 50% or more.

4.12

If you defer the written examination and your mark for the coursework element is below 50%, the [rules for compensation](#) will not apply and you must resit the coursework element.

Deadlines for items of assessment and submission guidance

See the [VLE](#) for full details of all the assessment points, deadlines and submission guidance.

4.13

All assessments must be submitted by the prescribed deadlines.

4.14

For coursework and project items, you should not exceed the maximum word limits by more than 10%. If the word count is between 10% to 20% above the maximum word limit, five marks will be deducted. If the word count exceeds the maximum word limit by more than 20%, you will receive a mark of zero for your work.

The content within the main body of text comprises the overall word count, including in-text citations, references, quotes, heading and sub-headings. The cover page, reference list and any appendices do **not** count towards the overall word count. Full submission instructions are included in the VLE with coursework submission forms.

Number of attempts permitted at an element of assessment

If you register to resit one or more elements of assessment for a module, you will be required to pay a module continuation fee. You do not have to take the assessment at the next available session; there are two sessions each year and you have five years to complete your studies.

4.15

The maximum number of attempts permitted for an element of coursework, written examination, or project report is **two**, except for DSM020, where students first registered on **2019-20 and October 2020**.

4.16

You will not be permitted to resit any element which you have passed.

4.17

If you fail a core module or the Final Project at the second attempt your registration on the programme will cease.

4.18

If you fail a compulsory module at the second attempt, and your mark cannot be compensated, your registration on the programme will cease.

4.19

If you fail an optional module at the second attempt, you may register for an alternative module, where possible.

Swapping a module

4.20

You will be permitted to swap an optional module for an alternative, where available, on a maximum of **two** occasions provided you are within your maximum period of registration. If you do so:

- you must withdraw from your current module
- you will have two attempts to pass the new module(s)
- your mark for the new module will be used for the purposes of classification.

4.21

You will not be permitted to swap any module which you have passed.

Students are permitted to withdraw from a module within **14 days of the module start date** with no financial penalty. You will not be permitted to register for an alternative until the next session.

If you choose to withdraw from a module later than 14 days after module start date you will need to pay a new module fee should you decide to swap for an available alternative.

Mitigating circumstances

See the mitigating circumstances page on the [Student Portal](#) for details of what mitigating circumstances are and how to submit a request to us.

4.22

If you enter one or more examinations and experience difficulties immediately before or during the examination period through illness or other adequate cause, you should immediately inform the University of London

University of London, and provide a supporting medical certificate or other official documentation for the modules affected.

The documentation must be received within three weeks of the last written examination, or before the final submission date for coursework (see 'Mitigating circumstances' in Section 12 of the [General Regulations](#)).

5 Transfer of Registration

Transfer between programmes offered under these regulations

5.1

All requests to transfer between programmes offered under these regulations will be considered by the Programme Director on a case-by-case basis.

5.2

You may apply to transfer between equivalent programmes offered under these regulations provided that

- you have selected, or are still able to select, the core and compulsory modules on the programme to which you wish to transfer;
- you are still within your maximum period of registration AND
- you are not yet eligible for the award for the programme on which you are currently registered

5.3

If you transfer between MSc programmes we will credit you with any modules that you have already passed and any RPL that we previously awarded you provided they form part of the structure of your new programme.

5.4

If you have been awarded credit for a module, we will not allow you to resit it upon transfer.

5.5

Any failed attempts made will be carried forward and will be counted towards the number of attempts permitted at the same modules following transfer.

5.6

If you are permitted to transfer between MSc programmes offered under these regulations, all modules studied will be displayed on your final transcript when you receive your award. This includes modules which are discarded upon transfer.

6 Scheme of award

Mark scheme

6.1

The following mark scheme is used for the MSc, PGDip and PGCert:

Mark range	Outcome
70% and over	Distinction
60% – 69%	Merit
50% – 59%	Pass
0% – 49%	Fail

6.2

To be granted the award with Distinction, your mean average mark for modules must be 70% or above.

6.3

To be granted the award with Merit, your mean average mark for modules must be between 60% and 69%.

6.4

To be granted the award with Pass, your mean average mark for modules must be between 50% and 59%

Weighting of the assessments

6.5

The overall mark for modules is calculated by weighting the marks achieved for the coursework element and the written examination or two coursework elements. The weightings for each module can be found in [Assessment for the programme](#)

6.6

The overall mark for the Final Project is calculated by weighting the marks achieved for the coursework element and the examination element in a ratio of 70:30.

Date of award

6.7

The date of your award will correspond to the year that the requirements for the award were satisfied.

Qualification requirements

6.8

To be awarded an MSc, you must

- achieve a mark of at least 50% in each of the modules. The mark may be achieved through the application of the [rules for compensation](#), where appropriate.

6.9

To be awarded a PGDip, you must

- achieve a mark of at least 50% in each of the modules. The mark may be achieved through the application of the [rules for compensation](#), where appropriate.

6.10

To be awarded the PGCert Data Science, you must

- achieve a mark of at least 50% in each of the modules. The mark may be achieved through the application of the [rules for compensation](#), where appropriate.

Exit qualifications

6.11

If you are registered on an MSc or PGDip programme and are unable to complete your programme of study, an exit qualification (i.e. a PGCert or PGDip) may be granted for the successful completion of 60 or 120 credits respectively.

6.12

All exit qualifications are granted at the discretion of the exam board.

6.13

If you are registered on a specialist MSc programme and cannot complete your programme of study, you may be awarded a PGDip in line with your named specialism, providing you have completed the modules listed in Appendix A. If you have not completed the modules listed but have obtained enough credits, you will be awarded a PGDip Data Science.

6.14

If you complete 60 credits, you may only be awarded a PGCert Data Science.

6.15

The exit qualification will be in the year in which you satisfied the requirements for that award.

6.16

If you accept a PGCert or PGDip as an exit qualification under these regulations we will not permit you to register or re-register for the related MSc at a later date.

6.17

If you have transferred from stand-alone modules to a MSc and you are unable to complete your programme of study, you will only be permitted to use a maximum of four modules (60 credits) gained through stand-alone module studies towards the award of the PGDip.

6.18

If you have transferred from stand-alone modules to a PGDip and you are unable to complete your programme of study, you will only be permitted to use a maximum of two modules (30 credits) gained through stand-alone module studies towards the award of the PGCert.

Appendix A – Structure of the programmes

Foundations of Data Science: K-Means Clustering in Python

The programmes across the Data Science suite assume rudimentary experience with programming in Python. There is provision for students who do not have prior Python experience to take an online course as a primer on Python programming. This is a non-credit bearing, optional module available to students to ensure readiness to start core and compulsory modules and should take no more than 30 hours. It is highly recommended that students who have not had much experience in Python programming to complete this. [Please visit the course website on Coursera to review enrolment options.](#)

MSc Data Science

For the award of an **MSc Data Science** you must complete:

Four core modules (60 credits total):

- Big data analysis [DSM010]
- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹
- Machine learning [DSM040]

+

Two compulsory modules (30 credits total):

- Data visualisation [DSM050]
- Data science research topics [DSM060]

+

Four optional modules chosen from the below list (60 credits total):

- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]
- Artificial intelligence [DSM100]
- R for data science [DSM110]
- Financial data modelling [DSM120]
- Natural language processing [DSM140]
- Neural networks [DSM150]
- Social networks and graph analysis [DSM160]

+

A final project in data science [DSM500] (30 credits total)

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

PGDip Data Science

For the award of a **PGDip Data Science** you must complete:

Four core modules (60 credits total):

- Big data analysis [DSM010]
- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹
- Machine learning [DSM040]

+

Two compulsory modules (30 credits total):

- Data visualisation [DSM050]
- Data science research topics [DSM060]

+

Two optional modules chosen from the below list (30 credits total):

- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]
- Artificial intelligence [DSM100]
- R for data science [DSM110]
- Financial data modelling [DSM120]
- Natural language processing [DSM140]
- Neural networks [DSM150]
- Social networks and graph analysis [DSM160]

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

PGCert Data Science

For the award of a **PGCert Data Science** you must complete:

Two core modules chosen from the below list (30 credits total):

- Big data analysis [DSM010]
- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹
- Machine learning [DSM040]

+

Two other modules chosen from the below list (30 credits total):

- Data visualisation [DSM050]
- Data science research topics [DSM060]
- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]
- Artificial intelligence [DSM100]
- R for data science [DSM110]
- Financial data modelling [DSM120]
- Natural language processing [DSM140]
- Neural networks [DSM150]
- Social networks and graph analysis [DSM160]
- Any core modules not already taken from the list above

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

MSc Data Science and Financial Technology

For the award of an **MSc Data Science and Financial Technology** you must complete:

Four core modules (60 credits total):

- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹

Machine learning [DSM040]

- Financial data modelling [DSM120]

+

Three compulsory modules (45 credits total):

- Big data analysis [DSM010]
- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]

+

Three optional modules chosen from the below list (45 credits total):

- Data science research topics [DSM060]
- Data visualisation [DSM050]
- Artificial intelligence [DSM100]
- R for data science [DSM110]
- Natural language processing [DSM140]
- Neural networks [DSM150]
- Social networks and graph analysis [DSM160]

+

A final project in data science and financial technologies [DSM500] (30 credits total)

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

PGDip Data Science and Financial Technology

For the award of a **PGDip Data Science and Financial Technology** you must complete:

Four core modules (60 credits total):

- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹
- Machine learning [DSM040]
- Financial data modelling [DSM120]

+

Three compulsory modules (45 credits total):

- Big data analysis [DSM010]
- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]

+

One optional module chosen from the below list (15 credits total):

- Data science research topics [DSM060]
- Data visualisation [DSM050]
- Artificial intelligence [DSM100]
- R for data science [DSM110]
- Natural language processing [DSM140]
- Neural networks [DSM150]
- Social networks and graph analysis [DSM160]

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

MSc Data Science and Artificial Intelligence

For the award of an **MSc Data Science and Artificial Intelligence** you must complete:

Four core modules (60 credits total):

+

Three compulsory modules (45 credits total):

- Big data analysis [DSM010]
- Data science research topics [DSM060]
- Neural networks [DSM150]

+

Three optional modules chosen from the list below (45 credits total):

- Data visualisation [DSM050]
- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]
- R for data science [DSM110]
- Financial data modelling [DSM120]
- Natural language processing [DSM140]
- Social networks and graph analysis [DSM160]

+

A final project in data science and artificial intelligence [DSM500] (30 credits total)

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

PGDip Data Science and Artificial Intelligence

For the award of a PGDip Data Science and Artificial Intelligence you must complete:

- Four core modules (60 credits total):
- Data programming in Python [DSM020]
- Statistics and Statistical Data Mining [DSM030] ¹
- Machine learning [DSM040]
- Artificial intelligence [DSM100]

+

Three compulsory modules (45 credits total):

- Big data analysis [DSM010]
- Data science research topics [DSM060]
- Neural networks [DSM150]

+

One optional module chosen from the list below (15 credits total):

- Data visualisation [DSM050]
- Blockchain programming [DSM070]
- Mathematics of financial markets [DSM080]
- R for data science [DSM110]
- Financial data modelling [DSM120]
- Natural language processing [DSM140]
- Social networks and graph analysis [DSM160]

Notes

All students

¹ If you first registered for this programme in 2019-20 or the October session of 2020-21 and have passed DSM030 Mathematics and statistics for data science, you will be able to use DSM030 in place of DSM030 Statistics and Statistical Data Mining.

Appendix B – Module Outlines

Topics covered in each module may be revised to ensure currency and relevance. Students will be advised of any changes in advance of their study. See module descriptors under the Structure tab on the [website](#) for further information and any updates.

Foundations of Data Science: K- Means Clustering in Python

The programmes across the Data Science suite assume rudimentary experience with programming in Python. There is provision for students who do not have prior Python experience to take an online course as a primer on Python programming. This is a non-credit bearing, optional module available to students to ensure readiness to start core and compulsory modules and should take no more than 30 hours. It is highly recommended that students who have not had much experience in Python programming to complete this. [Please visit the course website on Coursera to review enrolment options.](#)

Big data analysis [DSM010]

This module covers the topic of Big Data which is a key element of contemporary applications of Data Science. It provides practical skills related to working with Big Data computing resources as well as the conceptual context of how Big Data relates to methods and technologies in statistics and computing. The module is complementary to Machine Learning and other modules.

By taking this module, you will gain an in-depth understanding of the technology and methods used for Big Data analysis and how they relate to concepts in statistics and computing more generally. The technologies will include distributed file systems, SQL and NoSQL databases. You will learn what the key challenges are in Big Data analysis, how they relate to privacy and security and how these are addressed with current technologies. You will work in Python, a modern language widely used in Big Data analysis. You will use querying to extract data, then design data processing and analysis pipelines to analyse the data. You will learn how to apply these techniques to data in business and scientific applications.

Topics covered in this module:

- Introduction to Big Data
- Environment for Big Data programming
- Distributed file systems and MapReduce
- Examples of Big Data applications
- Hadoop YARN
- Big Data Pipelines using Spark
- Functional Programming in Spark
- Spark Transactions and Actions
- Spark MLlib
- Recommender Systems

Data programming in python [DSM020]

This module aims to provide you with the programming skills you will need to carry out programming tasks you will encounter in the other modules in this programme. You will learn about general programming techniques such as variables, functions and control flow. You will learn how to work with different types of data structures such as arrays and dictionaries. You will develop data processing pipelines, which allow you to convert raw data into data that you can analyse. You will apply mathematical and statistical procedures to data. You will learn how to plot graphs of various types. You will also familiarise yourself with an industry standard data science programming environment which you can use throughout the programme

Topics covered in this module:

- Data structures
- Reading and writing data on the filesystem
- Retrieving data from the web
- Retrieving data from databases using query languages
- Cleaning and restructuring data, part 1
- Cleaning and restructuring data, part 2
- Data plotting
- Version control systems
- Unit tests
- Data processing pipelines

Statistics and Statistical Data Mining [DSM030]

Statistics is one of the cornerstones of reasoning in Data Science. This module aims to provide you with an understanding of the aspects of statistical analyses which are key to correctly interpreting and developing the results of the data analyses you will carry out in this module and elsewhere in the programme. It will also introduce you to some of the principal techniques for using statistics in data science. Topics covered in this module:

- Exploratory Data Analysis and Descriptive Statistics
- Data and Sampling Distributions
- Statistical Significance (t-test, ANOVA and so on)
- Regression and Prediction
- Classification and Classifier Evaluation Metrics
- Conditional Probability, Bayes' Theorem and the NaiveBayes Classifier
- K-Nearest Neighbours for Regression and Classification
- Tree Models for Regression and Classification
- Bagging, Random Forest and Boosting Techniques

Machine learning [DSM040]

Machine Learning covers a wide range of techniques that data scientists should be aware of, and able to use. This module aims to provide you with deep theoretical and practical knowledge of these techniques.

This module provides a broad view of Machine Learning and statistical pattern recognition. You will learn several techniques, including supervised learning (e.g. generative and discriminative learning, parametric and non-parametric learning), unsupervised learning (e.g. clustering), and theoretical aspects of Machine Learning (e.g. bias, variance). The module will also focus on recent applications of Machine Learning to areas of interest to data scientists.

Topics covered in this module:

- Introduction to Machine learning
- Classification
- Regression
- Model improvement
- Unsupervised learning
- Ensemble methods
- Neural networks and deep learning
- Working with time-series
- Probabilistic modelling
- Ethics and sustainability

Data visualisation [DSM050]

This module will equip you with data visualisation skills which you can use to generate compelling, interactive representations of data. Effectively presenting data in a manner in which it can be easily interpreted is a useful skill for a data scientist, especially when working with non-specialists. The module is complementary to other modules in the programme.

This module aims to equip you with a data visualisation skillset. By taking this module, you will be able to gain access to a wide range of data sources, and then to transform this data into compelling visualisations and plots. You will learn how to convert a data source into a visual story, wherein you will reveal and present the structure of the data. You will also learn how to publish your data visualisations on the web, such that your work is available to others. You will use a range of data visualisation tools, including interactive and animated graphing libraries.

Topics covered in this module:

- Introduction to data visualisation
- Data visualisation tools and methods
- Quantitative relationships, variables and types
- Descriptive statistics 1: Univariate analysis
- Descriptive statistics 2: Multivariate analysis
- Time-series data

- Geospatial data
- Connections and networks
- Large and high-dimensional data
- Scientific data visualisation

Data science research topics [DSM060]

This module will develop your awareness and abilities regarding the practice of professional data scientists. It will also expose you to a range of active research topics in Data Science, both in industry and academia. This will help you develop your Final Project ideas.

The module introduces you to research topics related to Data Science, for example: modern applications of Machine Learning, Data Mining, Computational Statistics, Computational Social Sciences, Cognitive computing, Business analytics and its applications.

In particular, you will have an opportunity to learn from Data Science and Computing industry professionals and academics whose mono-disciplinary or interdisciplinary research in areas of Computer Science, Sociology, Psychology, Bioinformatics, Biomedical Statistics and other disciplines is based or involves data analysis. The module will guide your work in exploring a research theme of interest, as a preliminary phase preparing you for the Final Project.

The module will also expose you to learning techniques of conducting research and writing scientific reports. Moreover, you will be taught approaches for devising and optimising your strategies adapted to your skills in seeking and securing an employment, whether in research or the industry sector. The module will also expose you to a discussion of the ethical and legal issues related to the area of Data Science and its applications.

Topics covered in this module:

Data science research methods

- Academic writing and presenting research results
- Ethics in data analytics, intellectual property
- Professional development of a data scientist
- Reproducibility of research
- Final project ideas
- Research topic 1: Artificial Intelligence (on the example of the Robot Scientist project)
- Research topic 2: Text mining
- Research topic 3: The Process of Knowledge Discovery in Databases (KDD)
- Research topic 4: Business data analytics

Data Science research topics in industry and academia; indicative topics (see more examples of the topics above):

- Artificial intelligence
- Text mining
- Business data analytics and its applications
- Data data-driven synthetic chemistry

- Data mining

Blockchain programming [DSM070] ♦

Blockchain is the technology that underpins Bitcoin and other cryptocurrencies. Blockchain promises to become a dominant technology in financial and other transactions, whether cryptocurrencies thrive or die. This module will give you a practical and theoretical knowledge of: how blockchains work, security issues with blockchains, where blockchains come from, how to analyse competing notions for blockchains (i.e. proof of stake vs proof of work; and bitcoin vs ethereum), what applications there are on the horizon.

This is both a practical and a theoretical module. It is practical in the sense that by running through the exercises you will have designed and implemented your own blockchain by the end. It is theoretical in the sense that you will have learned the theory behind security issues and the claims for competitive paths to blockchain issues. Your implementation is meant to reflect your thinking on the theoretical issues.

Mathematics of financial markets [DSM080] ♦

This module will provide you with a foundational knowledge of modern financial instruments. Such knowledge is crucial for people working in financial technology, who will have to create and understand software tools for optimising investment behaviour.

This module will give you the mathematical and qualitative tools to analyse modern financial markets, helping you develop strategies for market participants. It will introduce you to a range of derivatives and market behaviour. We take a distinctly mathematical and computational approach, leading you to understand how financial markets work well enough to analyse, evaluate and implement investment decisions involving financial instruments.

Artificial intelligence [DSM100] ♦

Through this module you will learn a wide range of Artificial Intelligence techniques. The module complements the Machine Learning module by focusing on Artificial Intelligence related topics it does not cover, such as symbolic representations, modelling, task learning and game playing. Through this module, you will widen and deepen your knowledge of Artificial Intelligence techniques, which should provide you with an additional set of methods with which you can address Data Science problems.

This module is focused on Artificial Intelligence techniques. You will understand the historical development of Artificial Intelligence including search, vision and planning. You will be familiar with the foundations of agent-based approach to software design, decision making and problem solving including under uncertainty. You will have an opportunity to apply Artificial Intelligence techniques.

R for data science [DSM110] ♦

This module will provide you with a wide range of applied data analysis techniques in R. With its strong focus on practical application of the techniques, and learning by doing, you will gain confidence in your ability to select appropriate techniques for attacking a wide variety of data sources. This module is complementary to other modules in the programme.

Financial data modelling [DSM120] ♦

Through this module you will learn nonlinear models and corresponding learning algorithms used in financial data mining of time series, a key contemporary application area for data scientists. The focus of the module is applications of Data Science techniques to real-world financial data. You will have an opportunity to further master and apply Data Science skills you gained in other modules of the programme to financial data.

This module aims to provide you with the ability to analyse financial, time series data using Machine Learning techniques. Analysis of financial data is an exciting application area for Data Science. You will study linear, non-linear and density models of sequential data, using Neural Networks. You will use gradient descent techniques and apply Kalman filters to enable proper dynamic treatments. You will investigate the concepts of overfitting, generalisation and performance evaluation. You will see several practical applications of Neural Networks and extended Kalman filters from the FinTech area of Financial Data Analytics. These include value-at-risk estimation, option pricing, portfolio estimation and automated, algorithmic trading.

Natural language processing [DSM140] ♦

Machine processing of natural language is a key target for the application of Data Science techniques. It has a range of specialised techniques that are being developed in a large and growing research field of Natural Language Processing (NLP). By taking this module you will gain a solid grasp and practical experience of those techniques.

This module is intended to provide you with a grounding in both rule-based and statistical approaches to NLP, and combines theoretical study with hands-on work employing widely used software packages. The module focuses on text processing, and does not deal with speech or multi-modal communication.

Neural networks [DSM150] ♦

Neural Networks are widely used techniques for modelling and classifying data. They are used in industry for data analysis applications such as image classification, speech analysis and regression tasks. This module will provide you with specialised theoretical and practical knowledge of a range of Neural Network architectures that are appropriate for data oriented tasks. This module is complementary to the Machine Learning and Artificial Intelligence modules in the programme, focusing on the area of neural computation.

This module introduces the theory and practice of neural computation. It offers the principles of neuro-computing with Neural Networks widely used for addressing real-world problems such as regression, pattern recognition and time-series prediction.

Social networks and graph analysis [DSM160] ♦

The analysis of graph-based data, a key example of which is social network data, is an important theoretical application area for Data Science, with considerable engagement from industry. This module aims to familiarise you with this conceptually and technically challenging area.

By taking this module, you will have an opportunity to master the theory and practice of graph analysis and social data analysis. You will be shown how to gather datasets from public social networking platforms. You will learn how to convert social network data into graph representations and then apply typical algorithms to analyse the structure of the graph. You will visualise graphs and assess data flow and influence between nodes in the graph.

Final project [DSM500] ♦

Overview

In this module you will undertake a substantial independent research project that will allow you to demonstrate a wide range of skills: project planning, management, research, and written presentation. You will integrate the knowledge gained throughout the programme and use skills acquired in other modules in the implementation of your Final Project. The work will consist of a combination of research and highly applicative elements in various proportions. For your project work you can make use of methodologies from various components of Data Science as instruments of research.

Final Project Assessment

1. Coursework

The project will be submitted as a sequence of deliverables. These will include, but not be limited to:

- Project proposal
- Preliminary project report
- Weekly progress logs, including source code versioning logs where appropriate
- Final Project report, including source code and data sets where appropriate
- Project presentation and demonstration where appropriate

2. Examination

The project will also be assessed through a two-hour unseen written examination, based on general questions about data science projects and specific questions about the project undertaken by the student.

Appendix C – Assessment Criteria

This is an indicative description of expectations at each grade level. Overall grades comprise qualitative and quantitative elements. The setting of questions, tasks and requirements and the accompanying marking scheme should take account of the criteria below.

Distinction (80+%)

An answer falling into the mark range 80+% demonstrates:

- very significant ability to evaluate critically existing methodologies and suggest new approaches to current research or professional practice;
- very high levels of creativity, originality and independent thought;
- very significant ability to plan, organise and execute independently a research project, coursework assignment or examination question;
- very significant ability to evaluate literature and theory critically and make informed judgements;
- very significant ability to analyse data critically;
- outstanding levels of accuracy, technical competence, organisation and expression.

Distinction (70-79%)

An answer falling into the mark range 70 to 79% demonstrates:

- a capacity to develop a sophisticated and intelligent argument;
- clear evidence of wide and relevant reading, referencing and an engagement with the conceptual issues;
- original thinking and a willingness to take risks;
- a significant ability to plan, organise and execute independently a research project, coursework assignment or examination question;
- rigorous use and a sophisticated understanding of relevant source materials, balancing appropriately between factual detail and key theoretical issues. Materials are evaluated directly and their assumptions and arguments challenged and/or appraised;
- significant ability to analyse data critically;
- correct referencing.

Merit (60 to 69%)

An answer falling into the mark range 60 to 69% demonstrates:

- a detailed understanding of the major factual and/or theoretical issues and directly engages with the relevant literature on the topic;
- strong evidence of critical insight and thinking;
- a capacity to develop a focused and clear argument and articulate clearly and convincingly a sustained train of logical thought;
- an ability to plan, organise and independently execute a research project, coursework assignment or examination question;
- an ability to analyse data critically;
- clear evidence of planning and appropriate choice of sources and methodology with correct referencing.

Pass (50 to 59%)

An answer falling into the mark range 50 to 59% demonstrates:

- a reasonable understanding of the major factual and/or theoretical issues involved;
- evidence of some knowledge of the literature with correct referencing;
- examples of a clear train of thought or argument;
- an ability to plan, organise and independently execute a research project, coursework assignment or examination question;
- how to introduce and appropriately conclude a text.

Fail (40 to 49%)

Fails to meet the minimum requirements of the assessment criteria. An answer falling into the mark range 40 to 49% demonstrates:

- some awareness and understanding of the literature and of factual or theoretical issues, but with little development and/or irrelevant/unrelated material or arguments included;
- a limited ability to present a clear and coherent argument;
- a limited ability to plan, organise and execute a research project, coursework assignment or examination question;
- a limited ability to analyse data;
- incomplete referencing.

Fail (20 to 39%)

Fails to meet the minimum requirements of the assessment criteria. An answer falling into the mark range 20 to 39% demonstrates:

- clear conceptual or factual errors or misunderstandings;
- only fragmentary evidence of critical thought or data analysis;
- very limited ability to plan, organise and execute a research project, coursework assignment or answer an examination question;
- fails to develop a coherent argument that relates to the research project or assignment or fails to answer the question or to develop an argument that relates to the question set;
- a lack of engagement with the relevant literature or fails to demonstrate a knowledge of the key issues;
- incomplete referencing.

Fail (0 to 19%)

Fails to meet the minimum requirements of the assessment criteria. An answer falling into the mark range 0 to 19% demonstrates:

- little or no knowledge or understanding;
- little or no knowledge of the relevant literature;
- no demonstrable ability to plan, organise and execute a research project, coursework assignment or examination question;
- no evidence of critical thought or data analysis;
- short answers and incoherent argument;
- major errors in referencing.