



Trinity College Dublin

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

School of Computer Science and Statistics

Postgraduate Certificate in Statistics & Data Science (online) 2023–2024



SEPTEMBER, 2023

Dear Participant,

Welcome to the Postgraduate Certificate in Statistics and Data Science.

This booklet contains some important information on various aspects of the course - dates of lecture terms, examination regulations and course outlines. So please take some time to read it.

This is the third year that the Certificate has been officially. All of the lectures, homework, reading material, live tutorial sessions and examinations will be made available online through the College's Blackboard system, accessible with your TCD username and password at <https://tcd.blackboard.com/>. This year we see the course title change from 'Statistics' to 'Statistics and Data Science', as we seek to broaden the course to take a brief look at some of the non-statistical aspects of data science.

If you have any queries or problems **in relation to the academic side** of the course then please contact the course director Professor Simon Wilson on swilson@tcd.ie. Natasha Blanchfield is the executive officer who administers the course; her number is +353 1 896 1787, email: pqcert.stats@tcd.ie. Administrative queries should be addressed to Natasha in the first instance.

Note that we will want to communicate with you by email from time to time. We will use the Blackboard system to do this, which sends messages to your College email address that you will be assigned (many of you will have these already in the form NAME@tcd.ie). Emails will NOT be sent to non-TCD email addresses. It is important, therefore, that you check your college mail regularly (and frequently!).

It is also important that you register at the appropriate times – if you do not, you may be charged a late registration fee.

I hope you enjoy the course.

Best wishes,

Simon Wilson

Course Director

Disclaimers:

The information contained in this document is intended to provide a guide to those seeking admission to the programme, and to the students on the course. Trinity College Dublin reserves the right to update or change syllabi, timetables, or other aspects of the programme at any time. Changes will be notified to current students by email.

Table of Contents for Course Handbook.

Postgraduate Certificate in Statistics and Data Science

INTRODUCTION	5
ACADEMIC STRUCTURE	5
ECTS – EUROPEAN CREDIT TRANSFER SYSTEM (ECTS)	5
REQUIREMENTS FOR POSTGRADUATE CERTIFICATE	5
KEY DATES:	6
STAFF	6
COURSE DIRECTOR	6
COURSE ADMINISTRATION.....	6
LECTURERS	6
EXTERNAL EXAMINER:	6
LOCATION OF LECTURES	6
TIMETABLE	7
EXAMINATIONS	8
MARKING SCALE	8
PREVIOUS EXAMINATION PAPERS	8
DEFERRING EXAMINATIONS.....	9
INDIVIDUAL WORK AND PLAGIARISM	9
STUDENT SUPPORTS AND SERVICES	9
THE POSTGRADUATE ADVISORY SERVICE	10
THE GRADUATE STUDENTS UNION	10
TRINITY CAREERS SERVICE	11
RESEARCH ETHICS	11
OTHER USEFUL LINKS AND COLLEGE CONTACTS	12
HEALTH & SAFETY	12
EMERGENCY PROCEDURE	12
APPEALS PROCESS FOR GRADUATE STUDENTS	13
DATA PROTECTION	13
FEEDBACK AND EVALUATION	13
BLACKBOARD	13
STUDENT INFORMATION SYSTEM (SITS)	13
APPENDIX A: MODULE DESCRIPTORS	14
MODULE 1 CODE	14
ST8001	14
MODULE 2 CODE	16
ST8002.....	16
MODULE 3 CODE	18

ST8003.....	18
MODULE 4 CODE	20
STP80080.....	20
APPENDIX B: PLAGIARISM CALENDAR ENTRY.....	22
APPENDIX C: COLLEGE APPEALS PROCEDURES.....	25
APPENDIX D: SCSS ASSESSMENT SUBMISSION SHEET.....	27

Alternative formats of the Handbook can be made on request.

Introduction

This handbook describes the Postgraduate Certificate in Statistics and Data Science. It details the required regulations and provides information on services and supports throughout College. In the event of any conflict between the General regulations published in the University Calendar and information contained in this handbook, the provision of the General Regulation in the Calendar will prevail. See <https://www.tcd.ie/calendar/graduate-studies-higher-degrees/> for the General College Regulations.

Academic Structure

The Postgraduate Certificate in Statistics and Data Science is run by the School of Computer Science and Statistics. The program is run online and consists of four modules: Introduction to statistical concepts and methods, Implementing statistical methods in R, Linear regression and Foundations of data science.

This course is now run entirely online and there is no need to come to Trinity College's campus for any aspect of it. The course will be delivered through Blackboard, the system that Trinity College uses for online learning; this includes lecture presentations, homework and other resources, live tutorial sessions and examinations. Blackboard can be accessed at <https://tcd.blackboard.com> using your Trinity College username and password.

The program is normally completed within one academic year. To facilitate students registered on the PG Certificate in Statistics and Data Science as their primary course who, due to work commitments, cannot do this, it is possible to take the course over two years. Note, though, that full fees are charged for the second year. Students who wish to take the Postgraduate Certificate course over two years must apply to do so before the end of the first semester. Students have to complete the course within two years.

The modules on statistical concepts and methods and on using R take place in the first semester, taking up 7 weeks and 4 weeks respectively. The other two modules on linear regression and foundations of data science take place in the second semester, taking 7 weeks and 4 weeks respectively. Detailed syllabi for each of the modules are given in Appendix A

ECTS – European Credit Transfer System (ECTS)

Each module is assigned an European Credit Transfer System (ECTS) rating. The modules on statistical concepts and methods, and on linear regression, are assigned an ECTS value of 10 with the other two modules assigned 5 ECTS each, for a total of 30 ECTS for the programme. As a rough guide, 1 credit represents 20-25 hours estimated input including class contact time, assessments and examinations.

Requirements for Postgraduate Certificate

To obtain the Postgraduate Certificate participants must pass all 4 modules. See below for further details.

Key dates:

Start of Semester 1: Monday 11th September 2023

Exams for Semester 1 modules: Monday 11th December to Friday 15th December 2022

Start of Semester 2: Monday 22nd January 2024

Exams for Semester 2 modules: Monday 29th April to Friday 3rd May 2024

Reassessment 2024: TBA (Contingency dates maybe added)

These dates are subject to change – all up to date information is available on the Academic Registry website. <https://www.tcd.ie/academicregistry/exams/>

Academic Year Structure <https://www.tcd.ie/calendar/academic-year-structure>

Staff

Course Director

Professor Simon Wilson Simon.Wilson@tcd.ie Tel +353 1 896 1759

Course Administration

Ms. Natasha Blanchfield (pgcert.stats@tcd.ie) Tel: +353 1 896 1787

Lecturers

Professor Mimi Zhang Mimi.Zhang@tcd.ie Tel +353 1 896 2726

Dr. John McDonagh MCDONAJ5@tcd.ie

Professor Emma Howard EMHOWARD@tcd.ie

Professor Andrea Patane apatane@tcd.ie

Professor Giovanni di Liberto gdilib@tcd.ie Tel +353 1 896 1261

Professor Simon Wilson Simon.Wilson@tcd.ie Tel +353 1 896 1759

External Examiner:

Professor Theodore Kypraios, University of Nottingham

Location of Lectures

Lectures will be taking place online.

Timetable

The online nature of the programme means that there is no strict timetable of lectures.

Each week, new content will be posted to the course's webpage within Blackboard. This will occur on the Friday before each week in the teaching semester. So the first material will appear on Blackboard on Friday 8th September 2023. On the Thursday of each week in the teaching semester, a live tutorial session will take place with the module lecturer or a demonstrator that will cover the material released the Friday before. Other live sessions, in the form of 'office hours' where students can go online and ask questions, may also be timetabled. Module lecturers will arrange the exact time for live tutorials and communicate that to you before the start of term.

The timetable for the release of content for the 4 modules follows the schedule listed below.

Semester 1 timetable

Week number	Module code and session number	Date of release of content	Date of live tutorial
1	ST8001 Session 1	08/09/2023	14/09/2023
2	ST8001 Session 2	15/09/2023	21/09/2023
3	ST8001 Session 3	22/09/2023	28/09/2023
4	ST8002 Session 1	29/09/2023	05/10/2023
5	ST8001 Session 4	06/10/2023	12/10/2023
6	ST8002 Session 2	13/10/2023	19/10/2023
7	Reading Week – no content released		
8	ST8001 Session 5	27/10/2023	02/11/2023
9	ST8002 Session 3	03/11/2023	09/11/2023
10	ST8001 Session 6	10/11/2023	16/11/2023
11	ST8001 Session 7	17/11/2023	23/11/2023
12	ST8002 Session 4	24/11/2023	30/11/2023

ST8001: Introduction to statistical concepts and methods

ST8002: Implementing statistical methods in R

Semester 2 timetable

Week number	Module code and session number	Date of release of content	Date of live tutorial
1	ST8003 Session 1	19/01/2024	25/01/2024
2	ST8003 Session 2	26/01/2024	01/02/2024
3	ST8003 Session 3	02/02/2024	08/02/2024
4	ST8003 Session 4	09/02/2024	15/02/2024
5	ST8003 Session 5	16/02/2024	22/02/2024
6	ST8003 Session 6	23/02/2024	29/02/2024
7	Reading Week – no content released		
8	ST8003 Session 7	08/03/2024	14/03/2024
9	STP80080 Session 1	15/03/2024	21/03/2024
10	STP80080 Session 2	22/03/2024	28/03/2024
11	STP80080 Session 3	29/03/2024	04/04/2024
12	STP80080 Session 4	05/04/2024	11/04/2024

ST8003: Linear regression

STP80080: Foundations of data science 1

Examinations

Each module will be examined separately. The form of the examination may vary from module to module and may include a combination of assignments, projects and written online examinations. Details will be given by the lecturers concerned.

As noted above, any online examination for the semester 1 modules takes place in the week of 11th December 2023. Any online examination for the semester 2 modules takes place in the week of 29th April 2024. Individual examination dates are set by the Examinations Office of the University and will not be available until later in the semester. Examination timetables will be posted to your my.tcd account:

No letters will be sent regarding this matter.

Subject to the recommendation of the court of examiners, students who are unsuccessful in the annual examinations may be allowed a supplemental examination. The supplemental examinations will normally be held in the week of 26th August 2024 (Contingency dates maybe added). Failure in the supplemental examination leads to exclusion from the course.

Marking Scale

Each module will be graded as follows: fail (< 50%), pass (50%+). To obtain the Postgraduate Certificate with distinction, participants must obtain an overall average of at least 70%. The average is weighted using the ECTS units for the various modules. Distinctions are not awarded at supplemental examinations.

Previous Examination Papers

Copies of previous examinations papers are published at:

http://www.tcd.ie/Local/Exam_Papers/index.html

Note that Supplemental Examination papers are not published. Please also note that all examinations from 2019 and before were 'face to face' examinations, taking place in a room and under invigilation and so less representative of the examinations of an online programme, which necessarily take place also online.

Deferring Examinations

In exceptional circumstances (e.g. illness, force majeure), permission may be granted by the Dean of Graduate Studies to defer examinations to the supplemental examination period. Contact the course administrator to organize the request to the Dean of Graduates Studies. Please note that inadequate preparation for the examinations is not a valid basis for a deferral and that force majeure does not include attendance at conferences, research visits or other work-related requirements.

If a student cannot sit an examination for medical reasons, **medical certificates must be submitted to the Course Administrator**; the College regulations require that this be done **“within three days of the beginning of the period of absence from the examination”**.

Unexplained absence from examinations results in exclusion from the course.

Remember that all students must complete the course within two years.

Individual Work and Plagiarism

It is important to highlight that all work submitted must be your own, and not taken directly from the internet or other sources. The College takes plagiarism seriously. The College regulations governing plagiarism are available in the college calendar and are copied in Appendix B. You are expected to be familiar with these rules and to understand what is considered plagiarism.

Before beginning your first assignment, you must complete the online tutorial on avoiding plagiarism ‘Ready, Steady, Write’, located at

<http://tcd-ie.libguides.com/plagiarism/ready-steady-write>

You are also encouraged to use the College Library’s repository of resources on plagiarism and its avoidance at

<http://tcd-ie.libguides.com/plagiarism>

In the case of group work, groups should establish some mechanism to ensure that no member engages in plagiarism. Do not sign the Group Assignment Declaration if you have not assured yourself that the whole assignment is original.

Student Supports and Services

Trinity College provides a number of supports for students. There is a handbook on the Student Services available at <http://www.tcd.ie/students/supports-services/>

The Postgraduate Advisory Service

The Postgraduate Advisory Service (PAS) is a free and confidential service available to all registered postgraduate students in Trinity College. PAS offers a comprehensive range of academic, pastoral and professional supports including one-to-one appointments, workshops and trainings, and emergency financial assistance.

PAS exists to ensure that all postgraduates students have a dedicated, specialist service independent of the School-system to whom they can turn for support and advice during their stay in College. Common concerns students present to PAS include stress; financial worries; queries about regulations or services available at Trinity; supervisor-relationship concerns; academic progression issues; academic appeals; and plagiarism hearings.

The Postgraduate Advisory Service is led by the Postgraduate Student Support Officers who provide frontline support for all Postgraduate students in Trinity. These Support Officers will act as your first point of contact and a source of support and guidance; they can also put you in touch with or recommend other services, depending on your needs.

For an appointment, please e-mail postgrad.support@tcd.ie

Website: https://www.tcd.ie/Senior_Tutor/postgraduateadvisory/

To keep up to date with the supports and events for postgraduate please check out the monthly PAS newsletter sent to all postgraduates via email or follow PAS on Instagram or Twitter: @TCDPGAdvisory

The Graduate Students Union

Located on the second floor of House Six, Trinity's Graduate Students' Union (commonly referred to as the GSU) is an independent body within College that represents postgraduate students throughout the University. Upon registration, all postgraduates are automatically members. The Union is managed by two full-time sabbatical officers the President and the Vice President. The President is responsible for strategy and policy formulation, whilst representing postgraduate students on a wide range of College Committees.

The Vice-president acts as the Union's Education and Welfare Officer and advises students on matters such as academic appeals and supervisor relationships; whilst also representing postgraduate students on a range of University Committees. The Vice president is also here to help on more personal matters, such as financial concerns, illness and bereavement. Any discussions about such concerns are treated with the strictest confidentiality.

Contact us at either: president@tcdgsu.ie, or vicepresident@tcdgsu.ie and please visit our website www.tcdgsu.ie for the latest postgraduate student updates. The GSU organise a range of social and research related events throughout the year and you'll hear about these from your GSU postgraduate representative, as well as via the GSU weekly email. Please get involved, either as a representative or at an individual level.

Trinity Careers Service

As a Trinity College Dublin student you have access to information, support and guidance from the professional team of Careers Consultants throughout your time at Trinity and for a year after you graduate. The support offered includes individual career guidance appointments, CV and LinkedIn profile clinics and practice interviews. The Trinity Careers Service and the School of Computer Science and Statistics also hold an annual Careers Fair in October which gives you the opportunity to find out about career prospects in a wide range of companies..

Visit <https://www.tcd.ie/Careers> for career and job search advice

Sign into MyCareer to book appointments, find information about vacancies and bursaries, and book your place on upcoming employer events.

Follow the service on Instagram for career news and advice @trinity.careers.service

Research Ethics

Any research project that involves human participation conducted through the courses (for example, a questionnaire or survey, or system user-evaluation, etc.) must have independent review by a Research Ethics Committee before its commencement.

A basic principle is that prospective participants should be fully informed about the research and its implications for them as participants, with time to reflect on the possibility for participation prior to being asked to sign an informed consent form.

For research associated with the School of Computer Science & Statistics, detailed information on the ethics application process can be found at <https://www.scss.tcd.ie/undergraduate/ethics>

It takes time to prepare an application for research ethics approval, to have the application considered, and to respond to feedback on the application where issues are raised. You should plan in your work for the time it takes to obtain research ethics approval.

Retrospective approval will not be granted.

Please also note, research conducted in the School of Computer Science and Statistics should be undertaken with cognisance of the TCD Guidelines for Good Research Practice; see <http://www.tcd.ie/about/policies/assets/pdf/TCDGoodResearchPractice.pdf>

Other Useful links and College Contacts

There are many other useful sites in TCD. Here are a number of them. If you find any other TCD links that you think would be useful for the class please e-mail the Course Administrator (pgcert.stats@tcd.ie).

Site	Address
Junior Dean and Registrar of Chambers Junior.Dean@tcd.ie	http://www.tcd.ie/Junior_Dean/
TCD Website	http://www.tcd.ie
Library	http://www.tcd.ie/library
IT Services	http://www.tcd.ie/itservices/
Graduate Studies	http://www.tcd.ie/Graduate_Studies/
Student Counselling	http://www.tcd.ie/Student_Counselling/
Computer Science and Statistics	http://www.scss.tcd.ie
Mature Student Office	http://www.tcd.ie/maturestudents/
Chaplaincy	http://www.tcd.ie/Chaplaincy/

Health & Safety

The Faculty of Engineering, Mathematics and Science Health and Safety Manual 2015/2016 is now available for download at the url below.

<http://www.ems.tcd.ie/faculty-health-safety.php>

There will be no printed Health & Safety booklets distributed this year and Declaration Forms A, B & C will now be forwarded to all incoming, registered students via email and in Word format to ease completion.

Emergency Procedure

In the event of an emergency, dial Security Services on extension 1999 (+353-1-8961999 from a mobile phone or an external landline). Security Services provide a 24-hour service to the college community, 365 days a year. They are the liaison to the Fire, Garda and Ambulance services and all staff and students are advised to always telephone extension 1999 (+353 1 896 1999) in case of an emergency. Should you require any emergency or rescue services on campus, you must contact Security Services. This includes chemical spills, personal injury or first aid assistance. It is recommended that all students save at least one emergency contact in their phone under ICE (In Case of Emergency).

University Policies

- The *Dignity and Respect Policy* supports a respectful work and study environment free from bullying and harassment: <https://www.tcd.ie/equality/policy/dignity-respect-policy/>
- The *Student Complaints Procedure* ensures that legitimate student complaints are investigated in a transparent, fair and timely manner: https://www.tcd.ie/about/policies/160722_Student%20Complaints%20Procedure_PUB.pdf
- A comprehensive listing of Academic Policies is available at: <http://www.tcd.ie/teaching-learning/academic-policies/>

Appeals process for Graduate students

The College has an extensive appeals process. See Appendix C for a copy of the procedure.

Data Protection

Trinity College Dublin uses personal data relating to students for a variety of purposes. We are careful to comply with our obligations under data protection laws and we have prepared a short guide (available at https://www.tcd.ie/info_compliance/data-protection/student-data/) to ensure you understand how we obtain, use and disclose student data in the course of performing University functions and services.

Feedback and Evaluation

At the end of the year the School of Computer Science and Statistics will conduct surveys of each module. Students are welcome to contact the Director of the Course if they have a complaint during the year

BlackBoard

Trinity College uses the online learning system Blackboard which is accessible via <http://mymodule.tcd.ie>

Student Information System (SITS)

ACCESS VIA <https://my.tcd.ie/>

All communications from College will be sent to you via your online portal which will give you access to an 'in tray' of your messages. You will also be able to view your timetables online, both for your teaching and for your examinations. All fee invoices/payments, student levies and commencement fees will be issued online and all payments will be carried out online. You will be able to view your personal details in the new system – some sections of which you will be able to edit yourself. Your results will also be communicated to you via the online portal. Future plans for the new system include online module registration and ongoing provision of module assessment results.

Enjoy the Course!

Appendix A: Module Descriptors

Module 1 Code	ST8001
Module 1 Title	Introduction to statistical concepts and methods
ECTS weighting	10
Term	Semester 1 Assessed
Contact Hours	7
Module Personnel	Professor Mimi Zhang
Module Learning Outcomes	<p>ML01.1 Demonstrate a systematic understanding of the fundamental inferential ideas which underpin statistical methods</p> <p>ML01.2 Demonstrate a broad understanding of the role of statistical ideas and methods covering both data collection and data analysis</p> <p>ML01.3 Demonstrate a competence in the use of basic statistical tools: data summaries, graphs, estimation and hypothesis testing</p>
Module Learning Aims	The module is designed to introduce students to the basic methods of statistical analysis
Module Content	<p>Specific topics addressed in this module include:</p> <p>Data summaries and graphs</p> <p>Statistical models</p> <p>Sampling distributions: confidence intervals and tests</p> <p>Comparative experiments: t-tests, confidence intervals, design issues</p>
Recommended Reading	<p>Core reading materials. The core reading material is the module study notes. These texts complement that material:</p> <p>Moore, McCabe and Craig, "Introduction to the Practice of Statistics", 9th edition, published by Freeman.</p> <p>Carlson and Winqvist, "An Introduction to Statistics: An Active Learning Approach", 2nd edition, published by Sage.</p> <p>Online reading materials. There are several free online texts in statistics that cover most of the material in this module. This one covers almost all of the module content:</p>

	<p>Shafer and Zhang, "Introductory Statistics", https://open.umn.edu/opentextbooks/textbooks/135</p> <p>Supplementary reading or similar materials:</p> <p>Grimmett and Welsh, "Probability: an Introduction", published by Oxford University Press.</p> <p>Hogg, McKean and Craig, "Mathematical Statistics: an Introduction", 8th edition.</p> <p>Spiegelhalter, "The Art of Statistics: How to Learn from Data Hardcover – Illustrated", published by Basic Books.</p> <p>James, Witten, Hastie and Tibshirani, "An Introduction to Statistical Learning: with Applications in R", published by Springer.</p> <p>Silver, "The Signal and the Noise", published by Penguin.</p> <p>Taleb, "The Black Swan", published by Allen Lane.</p>
<p>Assessment Details</p>	<p>The module is assessed by a realtime online examination (100%).</p>

Module 2 Code	ST8002
Module 2 Title	Implementing statistical methods in R
ECTS weighting	5
Term	Semester 1 Assessed
Contact Hours	4
Module Personnel	Professor Mimi Zhang
Module Learning Outcomes	<p>Students who complete this module should be able to:</p> <p>MLO2.1 Load and manipulate data for analysis in R</p> <p>MLO2.2 Implement statistical techniques that are taught in ST8001 in R</p> <p>MLO2.3 Interpret the output of these statistical techniques that R produces</p>
Module Learning Aims	The module is designed to introduce students to the R package and how it can be used to implement many of the methods described in ST8001.
Module Content	<p>Install and run R through the RStudio environment.</p> <p>Basic operations in R.</p> <p>Import data, format data, visualise data and save data.</p> <p>Implement the statistical methods of ST8001 in R.</p> <p>Interpret R outputs.</p>
Recommended Reading	<p>Core reading materials. The core reading material is of course the module study notes. These texts complement that material:</p> <p>Frost, "Regression Analysis".</p> <p>The R project website has an extensive online manual at https://cran.r-project.org/manuals.html</p> <p>Davies, "The Book of R: A First Course in Programming and Statistics", 1st edition. Published by No Starch Press.</p> <p>Supplementary reading or similar materials:</p> <p>James, Witten, Hastie and Tibshirani, "An Introduction to Statistical Learning: with Applications in R", published by Springer.</p> <p>Montgomery, Peck and Vining, "Introduction to Linear Regression Analysis", 5th edition, published by Wiley.</p>

Assessment Details	The module is assessed by a series of practical tasks (30%) and a final individual project (70%)
Teaching and Learning Methods	<p>This module has 4 weekly sessions that students work on in Weeks 9-12 of Michaelmas Term. The session will be released the Friday before the start of that week, with the synchronous session taking place on the Thursday or Friday of the week.</p> <p>Session 1 Installing and running R. The user interface. Basic operations.</p> <p>Session 2 Graphing data. Loading and saving data in R.</p> <p>Session 3 Calculating point estimates and confidence intervals.</p> <p>Session 4 Hypothesis testing.</p>

Module 3 Code	ST8003
Module 3 Title	Linear regression
ECTS weighting	10
Term	Hilary Term
Contact Hours	7
Module Personnel	Dr. John McDonagh
Learning Outcomes	<p>On successful completion of this module students should be able:</p> <p>ML03.1 To carry out an initial examination of the data to determine an appropriate regression model to use.</p> <p>ML03.2 To use a regression package (R) to apply multiple regression to simple data sets.</p> <p>ML03.3 To produce and interpret graphs for data summary and model diagnostics,</p> <p>ML03.4 To interpret the results of the model and see such modelling as the basis for more advanced statistical analysis.</p> <p>ML03.5 To construct and exploit derived variables, such as logs, products and indicator variables.</p>
Module Learning Aims	<p>Multiple linear regression – and its many variants – is the most widely used tool in applied statistics. This course will build on simple linear regression, introduced in the Base Module. The aim is to become familiar with its use, to further develop experience and confidence in the use and role of statistical modelling. As the class is diverse in terms of research area and quantitative skills, students are encouraged to conduct small analyses of data in their own research fields.</p>
Module Content	<p>Specific topics addressed in this module include:</p> <p>Review of simple linear regression model: assumptions, model fitting, estimation of coefficients and their standard errors</p> <p>The multiple linear regression model and its analysis including:</p> <p>Confidence intervals and statistical significance tests on model parameters</p> <p>Issues in the interpretation of the multiple parameters</p> <p>Analysis of variance in regression: F-tests, r-squared</p> <p>Indicator variables and interaction terms</p>

	<p>Model validation: residuals, residual plots, normal plots, diagnostics</p> <p>Introduction to logistic regression</p>
Recommended Reading List	<p>Core reading materials. The core reading material is of course the module study notes. These texts complement that material:</p> <p>Sheather, S. J. A Modern Approach to regression with R,, New York:, Springer 2009</p> <p>Neter, J., Wasserman, W. & Kutner, M.H. Applied Linear Models , 2nd edition Boston, Irwin:1989</p> <p>Kutner. M. H., Nachtsheim, C.J., Neter, J. & Li, W. Applied Linear Statistical Models, 5th, Boston: McGraw-Hill, 2005</p> <p>Davies, “The Book of R: A First Course in Programming and Statistics”, 1st edition. Published by No Starch Press.</p> <p>The R project website has an extensive online manual at https://cran.r-project.org/manuals.html including material on regression</p> <p>Supplementary reading or similar materials:</p> <p>James, Witten, Hastie and Tibshirani, “An Introduction to Statistical Learning: with Applications in R.</p>
Assessment Details	<p>The module is assessed by a series of practical tasks (50%) and a final project (50%)</p>
Weekly Session Descriptions	<p>This module has 7 weekly sessions that students work on in Weeks 1-7 of Hilary Term. The session will be released the Friday before the start of that week, with the synchronous session taking place on the Thursday or Friday of the week.</p> <p>Session 1 What is regression? Simple linear regression and the principle of least squares. Diagnostics.</p> <p>Session 2 Multiple linear regression, including confidence intervals, hypothesis tests and prediction.</p> <p>Session 3 Implementing regression in R and interpreting the output</p> <p>Session 4 Pre-processing data for better analysis: transformations and multi-collinearity.</p> <p>Session 5 Analysis of Variance: regression when the variables are discrete. Adding indicator variables and interactions to a regression analysis.</p> <p>Session 6 Introduction to logistic regression. Choosing a regression model. .</p> <p>Session 7 Implementing ANOVA and logistic regression in R and interpreting the output.</p>

Module 4 Code	STP80080
Module 4 Title	Foundations of Data Science 1
ECTS weighting	5
Term	Hilary Term
Contact Hours	5
Module Personnel	Professors Andrea Patane, Giovanni di Liberto, Emma Howard and Simon Wilson
Learning Outcomes	<p>Upon successful completion of the module the student should be able to:</p> <p>ML04.1 Identify the different skills, in computer science, statistics and elsewhere, that make up the field of data science and how they work together;</p> <p>ML04.2 Recognise which machine learning method is appropriate for a particular data analysis task, implement it and assess its performance;</p> <p>ML04.3 Undertake various statistical analyses and implement a machine learning algorithms using Python;</p>
Module Learning Aims	This module and its successor aim to give a broad overview of the aspects of data science other than statistics, broadly speaking: machine learning, distributed computing and data security and privacy. In addition, students will be introduced to another computing language, Python, to complement their knowledge of R, for data analysis. This module focusses on machine learning, and includes session that will be devoted to an industry case study.
Module Content	<p>Specific topics addressed in this module include:</p> <ul style="list-style-type: none"> • Machine learning and how it differs from statistics; non-statistical ML methods (such as case-based reasoning), regression with neural networks; classification (support vector machines, kNN); • Evaluating ML methods: cross validation, ROC; Installing and running Python. Its user interface and basic operations. • Introducing information systems and their use in managing and storing data; • Industry case study.
Recommended Reading List	

Assessment Details	The module is assessed by a series of final exam (100%)
Weekly Session Descriptions	<p>This module has 4 weekly sessions that students work on in Weeks 9-12 of Hilary Term. The session will be released the Friday before the start of that week, with the synchronous session taking place on the Thursday or Friday of the week</p> <p>Session 1 Programming in Python</p> <p>Session 2 Machine learning</p> <p>Session 3 Introduction to information systems</p> <p>Session 4 Industry case study</p>

Choose an item.

Choose an item.

Appendix B: Plagiarism Calendar Entry

PLAGIARISM GENERAL

It is clearly understood that all members of the academic community use and build on the work and ideas of others. It is commonly accepted also, however, that we build on the work and ideas of others in an open and explicit manner, and with due acknowledgement.

Plagiarism is the act of presenting the work or ideas of others as one's own, without due acknowledgement.

Plagiarism can arise from deliberate actions and also through careless thinking and/or methodology. The offence lies not in the attitude or intention of the perpetrator, but in the action and in its consequences.

It is the responsibility of the author of any work to ensure that he/she does not commit plagiarism.

Plagiarism is considered to be academically fraudulent, and an offence against academic integrity that is subject to the disciplinary procedures of the University.

Plagiarism can arise from actions such as:

(a) copying another student's work;

67 EXAMPLES OF PLAGIARISM

2. (b) enlisting another person or persons to complete an assignment on the student's behalf;
3. (c) procuring, whether with payment or otherwise, the work or ideas of another;

4. (d) quoting directly, without acknowledgement, from books, articles or other sources, either in printed, recorded or electronic format, including websites and social media;
5. (e) paraphrasing, without acknowledgement, the writings of other authors.

Examples (d) and (e) in particular can arise through careless thinking and/or methodology where students:

1. (i) fail to distinguish between their own ideas and those of others;
2. (ii) fail to take proper notes during preliminary research and therefore lose track of the sources from which the notes were drawn;
3. (iii) fail to distinguish between information which needs no acknowledgement because it is firmly in the public

domain, and information which might be widely known, but which nevertheless requires some sort of acknowledgement;

(iv) come across a distinctive methodology or idea and fail to record its source.

23

All the above serve only as examples and are not exhaustive.

68 PLAGIARISM IN THE CONTEXT OF GROUP WORK

Students should normally submit work done in co-operation with other students only when it is done with the full knowledge and permission of the lecturer concerned. Without this, submitting work which is the product of collaboration with other students may be considered to be plagiarism.

When work is submitted as the result of a group project, it is the responsibility of all students in the group to ensure, so far as is possible, that no work submitted by the group is plagiarised.

In order to avoid plagiarism in the context of collaboration and groupwork, it is particularly important to ensure that each student appropriately attributes work that is not their own.

69 SELF PLAGIARISM

No work can normally be submitted for more than one assessment for credit. Resubmitting the same work for more than one assessment for credit is normally considered self-plagiarism.

70 AVOIDING PLAGIARISM

Students should ensure the integrity of their work by seeking advice from their lecturers, tutor or supervisor on avoiding plagiarism. All schools and departments must include, in their handbooks or other literature given to students, guidelines on the appropriate methodology for the kind of work that students will be expected to undertake. In addition, a general set of guidelines for students on avoiding plagiarism is available at <http://tcd-ie.libguides.com/plagiarism>

71 If plagiarism as referred to in paragraph (1) above is suspected, the Director of Teaching and Learning (Postgraduate) or his/her designate will arrange an informal meeting with the student, the student's Supervisor and/or the academic staff member concerned, to put their suspicions to the student and give the student the opportunity to respond. Students may nominate a Graduate Students' Union representative or PG advisor to accompany them to the meeting. The student will be requested to respond in writing stating his/her agreement to attend such a meeting and confirming on which of the suggested dates and times it will be possible for them to attend. If the student does not in this manner agree to attend such a meeting, the Director of Teaching and Learning (Postgraduate), or designate, may refer the case directly to the Junior Dean, who will interview the student and may implement the procedures as referred to in Section 5 (Other General Regulations).

If the Director of Teaching and Learning (Postgraduate) forms the view that plagiarism has taken place, he/she must decide if the offence can be dealt with under the summary procedure set out below. In order for this summary procedure to be followed, all parties noted above must be in agreement and must state their agreement in writing to the Director of Teaching and Learning (Postgraduate) or designate. If one of the parties to the informal meeting withholds his/her written agreement to the application of the summary procedure, or if the facts of the case are in dispute, or if the Director of Teaching and Learning (Postgraduate) feels that the penalties provided for

under the summary procedure below are inappropriate given the circumstances of the case, he/she will refer the case directly to the Junior Dean, who will interview the student and may implement the procedures set out in Section 5 (Other General Regulations).

72 If the offence can be dealt with under the summary procedure, the Director of Teaching and Learning (Postgraduate) will recommend one of the following penalties:

(a) Level 1: Student receives an informal verbal warning. The piece of work in question is inadmissible. The student is required to rephrase and correctly reference all plagiarised elements. Other content should not be altered. The resubmitted work will be assessed and marked without penalty;

(b) Level 2: Student receives a formal written warning. The piece of work in question is inadmissible. The student is required to rephrase and correctly reference all plagiarised elements. Other content should not be altered. The resubmitted work will receive a reduced or capped mark depending on the seriousness/extent of plagiarism;

(c) Level 3: Student receives a formal written warning. The piece of work in question is inadmissible. There is no opportunity for resubmission.

Provided that the appropriate procedure has been followed and all parties in (6) above are in agreement with the proposed penalty, the Director of Teaching and Learning (Postgraduate) should in the case of a Level 1 offence, inform the Course Director and, where appropriate, the Course Office. In the case of a Level 2 or Level 3 offence, the Dean of Graduate Studies must be notified and requested to approve the recommended penalty. The Dean of Graduate Studies may approve or reject the recommended penalty, or seek further information before making a decision. If he/she considers that the penalties provided for under the summary procedure are inappropriate given the circumstances of the case, he/she may also refer the matter directly to the Junior Dean who will interview the student and may implement the procedures as referred to under conduct and college. Notwithstanding his/her decision, the Dean of Graduate Studies will inform the Junior Dean of all notified cases of Level 2 and Level 3 offences accordingly. The Junior Dean may nevertheless implement the procedures as set out in Section 5 (Other General Regulations).

73 If the case cannot normally be dealt with under summary procedures, it is deemed to be a Level 4 offence and will be referred directly to the Junior Dean. Nothing provided for under the summary procedure diminishes or prejudices the disciplinary powers of the Junior Dean under the 2010 Consolidated Statutes. to be a Level 4 offence and will be referred directly to the Junior Dean. Nothing provided for under the summary procedure diminishes or prejudices the disciplinary powers of the Junior Dean under the 2010 Consolidated Statutes.

<https://www.tcd.ie/calendar/graduate-studies-higher-degrees/complete-part-III-hl.pdf>

Appendix C: College Appeals procedures

APPEALS PROCESS FOR GRADUATE STUDENTS GENERAL

74 Every School will have a defined appeals process, with one or more appeals committees to hear student cases relating to decisions made at a School level. These include: i) appeals against the decision of a taught postgraduate course committee; and ii) appeals against the decision of a School research transfer/confirmation panel.

75 Appeals against the decision of the examiners of a research thesis will be made in the first instance to the Dean of Graduate Studies. Appeals against the decision of a School appeals committee will be made to the Dean of Graduate Studies.

76 Appeals by postgraduate students against the decision of the Dean of Graduate Studies will be made to the Academic Appeals Committee for Graduate Students (Taught) for taught graduate students or to the Academic Appeals Committee for Graduate Students (Research) for research graduate students – See Paragraph 1.24.7.

77 The Academic Appeals Committee for Graduate Students (Taught or Research) will refer its recommendation to the University Council for approval. If the Academic Appeals Committee for Graduate Students (Taught or Research) upholds a graduate student's appeal, then its recommendations must be fully implemented. Where the Academic Appeals Committee for Graduate Students (Taught or Research) cannot arrive at a decision, or where the graduate student is dissatisfied with the decision of the Committee, he/she may seek redress from the Visitor.

78 It should be noted that no appeals committee can overturn the examiners' academic verdict on the work as presented. Only changes of a procedural nature can be recommended. An appeal other than an ad misericordiam appeal cannot be made against the normal application of College academic regulations approved by the University Council.

REPRESENTATION

79 The student is entitled to have representation appropriate to the formality of the appeal hearing – i.e., a supervisor, other appropriate academic staff member (who is not a relative), a Graduate Students' Union representative, or an advisor from the Postgraduate Advisory Service. When the Academic Appeals Committee for Graduate Students (Taught or Research) hears an appeal, the graduate student must attend and is entitled to present her/his own case if this is desired, with the support of a Postgraduate Student Advisor or a representative of the Graduate Students Union.

CONFLICT OF INTEREST

80 Conflicts of interest will be avoided throughout all stages of the Postgraduate Appeals process (i.e., no person from the relevant course committee or an examiner or supervisor should serve on the appeals committee; when the Dean of Graduate Studies is also the supervisor, examiner or lecturer, a pro-Dean will be appointed).

APPEALING THE DECISION OF A TAUGHT POSTGRADUATE COURSE COMMITTEE

81 Where a graduate student has failed on the coursework component, and feels that there are mitigating circumstances, he/she may make a request, in writing, to the Course Co-ordinator/Director for permission to repeat the examination(s) and/or assessment(s). The reasons for the request must be clearly stated and supported where necessary by documentary evidence. If the graduate student is granted a supplemental examination, the results of that examination will be considered to be final, i.e., repetition of the year will not be permitted.

82 Where a graduate student has passed the coursework component but has failed on the assessment of the dissertation (including an oral examination), and is dissatisfied with the manner in which the dissertation was examined, he/she may make a request, in writing, to the Course Co-ordinator/Director. The reasons for the request must be clearly stated and supported where necessary by documentary evidence.

83 If the Course Co-ordinator refuses to grant a request, the student may make an appeal to the relevant School appeals committee. If the School appeals committee refuses to grant the appeal, or the student is not satisfied with the outcome of the appeal process, he or she may appeal the decision of the School in writing to the Dean of Graduate Studies, clearly stating under what grounds of (i) to (iii) listed in Paragraph 20 above they are appealing and what remedy they are seeking. If the Dean of Graduate Studies denies the appeal, the student may appeal this decision to the Academic Appeals Committee for Graduate Students (Taught).

Appendix D: SCSS Assessment Submission Sheet



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics

Assessment Submission Form

Student Name	
Student ID Number	
Course Title	
Module Title	
Lecturer(s)	
Assessment Title	
Date Submitted	
Word Count	

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: <http://www.tcd.ie/calendar>

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at <http://tcd-ie.libguides.com/plagiarism/ready-steady-write>

I declare that the assignment being submitted represents my own work and has not been taken from the work of others save where appropriately referenced in the body of the assignment.

Signed Date