

# ICT707 Data Science Practice

**School:** School of Science, Technology and Engineering

2021 | Semester 2

UniSC Southbank

**BLENDED  
LEARNING**

Most of your course is on campus but you may be able to do some components of this course online.

Online

**ONLINE**

You can do this course without coming onto campus, unless your program has specified a mandatory onsite requirement.

*Please go to [unisc.edu.au](http://unisc.edu.au) for up to date information on the teaching sessions and campuses where this course is usually offered.*

## 1. What is this course about?

### 1.1. Description

This course examines big data processing and analysis, using a modern framework such as Hadoop or Apache Spark. You will learn how to build data processing tools that can run on cloud computing systems and can scale up to process massive data sets. You will apply these skills to build tools that can generate business insights.

### 1.2. How will this course be delivered?

ACTIVITY	HOURS	BEGINNING WEEK	FREQUENCY
<b>BLENDED LEARNING</b>			
<b>Online</b> – Pre-recorded concept videos and associated activity	1hr	Not applicable	12 times
<b>Tutorial/Workshop 1</b> – On campus tutorial	2hrs	Not applicable	11 times
<b>ONLINE</b>			
<b>Online</b> – Pre-recorded concept videos and associated activity	1hr	Not applicable	12 times
<b>Tutorial/Workshop 1</b> – Interactive zoom tutorial	2hrs	Not applicable	11 times

### 1.3. Course Topics

- Spark Runtime and RDD
- Pair RDD and Files
- DataFrame and SparkSQL
- Hadoop
- MapReduce
- Parallel Computing
- Machine Learning with Spark
- Advanced Spark Programming

## 2. What level is this course?

700 Level (Specialised)

Demonstrating a specialised body of knowledge and set of skills for professional practice or further learning. Advanced application of knowledge and skills in unfamiliar contexts.

### 3. What is the unit value of this course?

12 units

### 4. How does this course contribute to my learning?

COURSE LEARNING OUTCOMES		GRADUATE QUALITIES
On successful completion of this course, you should be able to...		Completing these tasks successfully will contribute to you becoming...
1	Design and build programs that can load, transform, analyse and store big data using cloud computing techniques.	Knowledgeable Creative and critical thinker
2	Apply data mining, analysis and visualisation techniques to big data to gain business insights.	Creative and critical thinker Empowered
3	Research and apply theory and practice of scalable distributed data analysis within the discipline.	Knowledgeable Empowered
4	Demonstrate and justify the use of big data analysis skills to develop innovative solutions to business problems.	Creative and critical thinker Engaged
5	Demonstrate critical and creative thinking to identify and solve complex business problems and arrive at innovative solutions.	Creative and critical thinker

### 5. Am I eligible to enrol in this course?

Refer to the [UniSC Glossary of terms](#) for definitions of “pre-requisites, co-requisites and anti-requisites”.

#### 5.1. Pre-requisites

ICT705 and ICT706 and enrolled in a Postgraduate Program

#### 5.2. Co-requisites

Not applicable

#### 5.3. Anti-requisites

Not applicable

#### 5.4. Specific assumed prior knowledge and skills (where applicable)

Not applicable

#### 5.5. Microcredential Information

Not applicable

### 6. How am I going to be assessed?

#### 6.1. Grading Scale

Standard Grading (GRD)

High Distinction (HD), Distinction (DN), Credit (CR), Pass (PS), Fail (FL).

#### 6.2. Details of early feedback on progress

Task 1 is a test involving basic concepts, principles, and skills of data science practice, which will be the basis for the understanding of Spark programming.

### 6.3. Assessment tasks

DELIVERY MODE	TASK NO.	ASSESSMENT PRODUCT	INDIVIDUAL OR GROUP	WEIGHTING %	WHAT IS THE DURATION / LENGTH?	WHEN SHOULD I SUBMIT?	WHERE SHOULD I SUBMIT IT?
All	1	Examination - not Centrally Scheduled	Individual	20%	60min	Week 5	Online Test (Quiz)
All	2	Examination - not Centrally Scheduled	Individual	50%	90min	Week 9	Online Assignment Submission with plagiarism check
All	3	Artefact - Technical and Scientific, and Written Piece	Individual	30%	Big data analysis + 1,000 word report	Week 12	Online Assignment Submission with plagiarism check

#### All - Assessment Task 1: Big data test

<b>GOAL:</b>	To build your knowledge of big-data processing skills and problem-solving techniques.		
<b>PRODUCT:</b>	Examination - not Centrally Scheduled		
<b>AUTHORSHIP STATEMENT:</b>			
<b>FORMAT:</b>	Coding test based on the content of Week 1 – 4. This task will help to build your knowledge of basic Spark programming. Further details of this assessment will be given on Blackboard.		
<b>CRITERIA:</b>	<b>No.</b>		<b>Learning Outcome assessed</b>
	1	Analysis of the given problem	4
	2	Application of relevant programming concepts	1
	3	Accuracy of the program output	2
<b>GENERIC SKILLS:</b>			

#### All - Assessment Task 2: Mid-semester test

<b>GOAL:</b>	To demonstrate understanding of the theory and practice of scalable distributed data analysis.		
<b>PRODUCT:</b>	Examination - not Centrally Scheduled		
<b>AUTHORSHIP STATEMENT:</b>			
<b>FORMAT:</b>	This is an individual assessment. Answer a set of questions about big data analysis theory and practice		
<b>CRITERIA:</b>	<b>No.</b>		<b>Learning Outcome assessed</b>
	1	Comprehension, application and communication of definitions and concepts used in big data processing	3
	2	Comparison and selection of alternative data analysis techniques	4
	3	Demonstration of your understanding of data analysis theory	3
<b>GENERIC SKILLS:</b>			

### All - Assessment Task 3: Big data assignment

<b>GOAL:</b>	To demonstrate a comprehensive view of big data analysis in terms of definitions and concepts, techniques, and producing big-data solutions to business problems.		
<b>PRODUCT:</b>	Artefact - Technical and Scientific, and Written Piece		
<b>AUTHORSHIP STATEMENT:</b>			
<b>FORMAT:</b>	A program that uses big-data analysis techniques to solve a business problem, plus a report (1000 words) describing and justifying the design of that program.		
<b>CRITERIA:</b>	<b>No.</b>		<b>Learning Outcome assessed</b>
	1	Knowledge of complex problem-solving and/or analytical processes appropriate to their business discipline	4
	2	Demonstrate reflective thinking for complex problem solving and decision making in a business context	5
	3	Application of relevant programming concepts	1
	4	Adherence to program output and recommended programming styles	2
<b>GENERIC SKILLS:</b>			

## 7. Directed study hours

A 12-unit course will have total of 150 learning hours which will include directed study hours (including online if required), self-directed learning and completion of assessable tasks. Student workload is calculated at 12.5 learning hours per one unit.

## 8. What resources do I need to undertake this course?

Please note: Course information, including specific information of recommended readings, learning activities, resources, weekly readings, etc. are available on the course Canvas site– Please log in as soon as possible.

### 8.1. Prescribed text(s) or course reader

You need regular access to the resource(s) below. Many texts are available as ebooks through the [Library](#) at no additional cost.

REQUIRED?	AUTHOR	YEAR	TITLE	EDITION	PUBLISHER
Required	Holden Karau, Andy Konwinski, Patrick Wendell and Matei Zaharia	2015	Learning Spark: Lightning-fast data analysis		O'Reilly Media, Inc

### 8.2. Specific requirements

You must have a computer (Desktop or Laptop) that you can install Python and Spark software on, in order to be able to practice the programming skills outside lecture and workshop times.

## 9. How are risks managed in this course?

Health and safety risks for this course have been assessed as low. It is your responsibility to review course material, search online, discuss with lecturers and peers and understand the health and safety risks associated with your specific course of study and to familiarise yourself with the University's general health and safety principles by reviewing the [online induction training for students](#), and following the instructions of the University staff.

## 10. What administrative information is relevant to this course?

### 10.1. Assessment: Academic Integrity

Academic integrity is the ethical standard of university participation. It ensures that students graduate as a result of proving they are competent in their discipline. This is integral in maintaining the value of academic qualifications. Each industry has expectations and standards of the skills and knowledge within that discipline and these are reflected in assessment.

Academic integrity means that you do not engage in any activity that is considered to be academic fraud; including plagiarism, collusion or outsourcing any part of any assessment item to any other person. You are expected to be honest and ethical by completing all work yourself and indicating in your work which ideas and information were developed by you and which were taken from others. You cannot provide your assessment work to others. You are also expected to provide evidence of wide and critical reading, usually by using appropriate academic references.

In order to minimise incidents of academic fraud, this course may require that some of its assessment tasks, when submitted to Canvas, are electronically checked through Turnitin. This software allows for text comparisons to be made between your submitted assessment item and all other work to which Turnitin has access.

### 10.2. Assessment: Additional Requirements

Eligibility for Supplementary Assessment Your eligibility for supplementary assessment in a course is dependent of the following conditions applying: The final mark is in the percentage range 47% to 49.4% The course is graded using the Standard Grading scale You have not failed an assessment task in the course due to academic misconduct

### 10.3. Assessment: Submission penalties

Late submission of assessment tasks may be penalised at the following maximum rate: - 5% (of the assessment task's identified value) per day for the first two days from the date identified as the due date for the assessment task. - 10% (of the assessment task's identified value) for the third day - 20% (of the assessment task's identified value) for the fourth day and subsequent days up to and including seven days from the date identified as the due date for the assessment task. - A result of zero is awarded for an assessment task submitted after seven days from the date identified as the due date for the assessment task. Weekdays and weekends are included in the calculation of days late. To request an extension you must contact your course coordinator to negotiate an outcome.

### 10.4. Links to relevant University policy and procedures

For more information on Academic Learning & Teaching categories including:

- Assessment: Courses and Coursework Programs
- Review of Assessment and Final Grades
- Supplementary Assessment
- Central Examinations
- Deferred Examinations
- Student Conduct
- Students with a Disability

For more information, visit <https://www.usc.edu.au/explore/policies-and-procedures#academic-learning-and-teaching>

### 10.5. Student Charter

UniSC is committed to excellence in teaching, research and engagement in an environment that is inclusive, inspiring, safe and respectful. The [Student Charter](#) sets out what students can expect from the University, and what in turn is expected of students, to achieve these outcomes.

### 10.6. General Enquiries

#### In person:

- **UniSC Sunshine Coast** - Student Central, Ground Floor, Building C, 90 Sippy Downs Drive, Sippy Downs
- **UniSC Moreton Bay** - Service Centre, Ground Floor, Foundation Building, Gympie Road, Petrie
- **UniSC SouthBank** - Student Central, Building A4 (SW1), 52 Merivale Street, South Brisbane
- **UniSC Gympie** - Student Central, 71 Cartwright Road, Gympie
- **UniSC Fraser Coast** - Student Central, Student Central, Building A, 161 Old Maryborough Rd, Hervey Bay
- **UniSC Caboolture** - Student Central, Level 1 Building J, Cnr Manley and Tallon Street, Caboolture

**Tel:** +61 7 5430 2890

**Email:** [studentcentral@usc.edu.au](mailto:studentcentral@usc.edu.au)