

Subject Description Form

Subject Code	AMA3632
Subject Title	Programming for Data Science
Credit Value	3
Level	3
Pre-requisite	Nil
Objectives	This course aims at familiarizing students with the capabilities of the R language for data management and statistical analysis and computation. Students will learn the basics of R from creating variables and performing basic operations, to handling of data structures including vectors, matrices, data frames, linked lists and trees. Also, the graphical capabilities of R will be explored for data visualizations and tree creation and traversal. Students will also learn basic numerical techniques and machine learning techniques using R to tackle practical problems.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: <ul style="list-style-type: none"> a) Identify and develop problem solutions in a logical manner; b) Demonstrate programming skills logically and systematically; c) Understand the properties of various data structures; d) Develop programs for data cleaning and validating; e) Present data by summarizing and visualizing of data graphically; f) Communicate the statistical findings from data; g) Demonstrate critical data management including data manipulation, queries and merge; h) Possess the ability to adapt to other statistical software including JMP, SPSS and Minitab.
Subject Synopsis/ Indicative Syllabus	Students are required to construct and execute programs in R using elementary programming techniques. The details of the framework include: <ul style="list-style-type: none"> - Import/export of data from files or Internet; - Assign and manipulate various data structures; - Create, navigate and traverse a tree; - Use loop and condition statements; - Create user-defined functions; - Learn debugging; - Use built-in and/or external R packages for statistical calculations; - Learn basic numerical techniques with R; and - Learn basic machine learning techniques with R.
Teaching/Learning Methodology	The subject focuses on the data management with emphasis on the conceptual elements in computer programming. The lectures will be taught in a workshop mode with hands-on exercises reinforcing taught concepts. Students are required to attend the laboratory sessions, which allows them to consolidate their concepts learnt in the lectures. Other practical work helps to reinforce the programming skills learned for applications. Students will gain experience in data management by working with

	datasets given in the course.										
Assessment Methods in Alignment with Intended Learning Outcomes	Specific assessment methods/tasks	% weighting	Intended subject learning outcomes to be assessed (Please tick as appropriate)								
			a	b	c	d	e	f	g	h	
	1. Continuous Assessment (Assignments, Mini-projects, quizzes and test)	60%	✓	✓	✓	✓	✓	✓	✓	✓	
	2. Examination	40%	✓	✓	✓	✓		✓	✓		
	Total	100 %									
<p>Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes: The subject focuses on knowledge, skills and understanding of Programming for Data Science, thus, continuous assessment is the most appropriate assessment method, including 60% worth of individual assignments, mini-projects, online quizzes and a mid-term test. All these components and examination (40%) will be designed to assess the specified learning outcomes.</p>											
Student Study Effort Expected	Class contact:										
	• Lecture		26 Hrs.								
	• Tutorial		13 Hrs.								
	Other student study effort:										
	• Assignment and mini-project		40 Hrs.								
	• Self-study		41 Hrs.								
	Total student study effort:		120 Hrs.								
Reading List and References	Textbook:										
	O. Jones, R. Maillardet and A. Robinson	Introduction to Scientific Programming and Simulation using R, 2 nd Edition							CRC Press 2014		
	References:										
	Larry, Pace	Beginning R an introduction to statistical programming					Springer 2012				
Matloff, Norman S.	Art of R programming – A tour of statistical software design					No Starch Press 2011					
Black, Kelly	R object-oriented programming					Packt Publishing 2014					