APPROVED

DATA C8Z01: Statistics using R

Module Details

Module Code:	DATA C8Z01
Full Title:	Statistics using R APPROVED
Valid From::	Semester 1 - 2020/21 (September 2020)
Language of Instruction:	English
Duration:	1 Semester
Credits::	10
Module Owner::	Fiona Lawless
Departments:	Unknown
Module Description:	The module lays solid foundations in statistics using the R programming language.

Module Learning Outcome	
On successful completion of this module the learner will be able to:	
#	**Module Learning Outcome Description**
MLO1	Write basic functions using R data structures.
MLO2	Use R to produce statistical graphics.
MLO3	Apply fundamental concepts and techniques in exploratory data analysis using R.
MLO4	Apply fundamental concepts of probability laws and recognise the appropriate probability distribution to model given problems.
MLO5	Understand, study, design the processing of data and use it to insure integrity of data.
MLO6	Construct and interpret appropriate hypothesis testing and confidence intervals for one, two and paired samples, and more than two samples. Implement hypothesis testing using R.
MLO7	Evaluate correlation and conduct analysis for simple linear regression models using R.

Pre-requisite learning
Module Recommendations *This is prior learning (or a practical skill) that is strongly recommended before enrolment in this module. You may enrol in this module if you have not acquired the recommended learning but you will have considerable difficulty in passing (i.e. achieving the learning outcomes of) the module. While the prior learning is expressed as named DkIT module(s) it also allows for learning (in another module or modules) which is equivalent to the learning specified in the named module(s).*
No recommendations listed

Module Indicative Content
R Understand and use R data structures. Understand how to use R for mathematical computation programming. Use R to produce statistical graphics and for data analysis.
Descriptive Statistics Frequency tables, measures of central tendency & variation; Graphical representation of data
Processing of Data Missing Data, Outliers
Probability Theory Basic Laws of Probability, Probabilistic Problem Solving, Bayes Theorem
Probability distributions Binomial, Poisson, Normal and other distributions, Monte Carlo method
Hypothesis Testing One Sample, Two Sample, Paired Sample and ANOVA hypothesis testing
Regression Analysis Scatterplots, Correlation & Simple Linear Regression Analysis.
Bootstrap Bootstrap methods, Density estimation.

Module Assessment

Assessment Breakdown	%
Course Work	20.00%
Project	20.00%
Practical	20.00%
Final Examination	40.00%

Module Special Regulation

Assessments

Full Time On Campus

Course Work

Assessment Type	Continuous Assessment	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40
Timing	n/a	**Learning Outcome**	1,2,3,4,5,7
Duration in minutes	30		

Assessment Description
Bi-Weekly short individual & group exercises or quizzes.

Project

Assessment Type	Project	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40
Timing	End-of-Semester	**Learning Outcome**	1,2,3,4,5,6,7
Duration in minutes	0		

Assessment Description
Data Project 2. Practical Implementation of some aspects of content of this module that involve analyzing data using R and will form part of a joint project with Applied Database Systems.

Practical

Assessment Type	Practical/Skills Evaluation	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40
Timing	n/a	**Learning Outcome**	1,2,3,4
Duration in minutes	60		

Assessment Description
An in-class lab test which will require students to use R to answers problems and do some elementary statistics

Final Examination

Assessment Type	Formal Exam	**% of Total Mark**	40
Marks Out Of	0	**Pass Mark**	0
Timing	End-of-Semester	**Learning Outcome**	3,4,5,6,7
Duration in minutes	120		

Assessment Description
n/a

Part Time On Campus

Course Work

Assessment Type	Continuous Assessment	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40
Timing	n/a	**Learning Outcome**	1,2,3,4,5,7
Duration in minutes	30		

Assessment Description
Bi-Weekly short individual & group exercises or quizzes.

Project

Assessment Type	Project	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40
Timing	End of Year	**Learning Outcome**	1,2,3,4,5,6,7
Duration in minutes	0		

Assessment Description
Data Project 2. Practical Implementation of some aspects of content of this module that involve analyzing data using R and will form part of a joint project with Applied Database Systems.

Practical

Assessment Type	Practical/Skills Evaluation	**% of Total Mark**	20
Marks Out Of	100	**Pass Mark**	40

Timing	n/a	**Learning Outcome**	1,2,3,4
Duration in minutes	60		

Assessment Description
An in-class lab test which will require students to use R to answers problems and do some elementary statistics

Final Examination

Assessment Type	Formal Exam	**% of Total Mark**	40
Marks Out Of	0	**Pass Mark**	40
Timing	End of Year	**Learning Outcome**	3,4,5,6,7
Duration in minutes	120		

Assessment Description
Final Examination

Reassessment Requirement

A repeat examination
Reassessment of this module will consist of a repeat examination. It is possible that there will also be a requirement to be reassessed in a coursework element.

Module Workload

Workload: Full Time On Campus					
Workload Type	*Contact Type*	*Workload Description*	*Frequency*	*Average Weekly Learner Workload*	*Hours*
Lecture	Contact	No Description	Every Week	2.00	2
Practical	Contact	No Description	Every Week	4.00	4
Independent Study	Non Contact	No Description	Every Week	10.00	10
				Total Weekly Learner Workload	16.00
				Total Weekly Contact Hours	6.00

Workload: Part Time On Campus					
Workload Type	*Contact Type*	*Workload Description*	*Frequency*	*Average Weekly Learner Workload*	*Hours*
Lecture	Contact		Every Week	1.00	1
Practical	Contact		Every Week	2.00	2
Independent Study	Non Contact		Every Week	5.00	5
				Total Weekly Learner Workload	8.00
				Total Weekly Contact Hours	3.00

Module Resources

Supplementary Book Resources

Andy Field, Jeremy Miles, & Zoe Field. (2013), Discovering Statistics using R, SAGE Publications, [ISBN: 9781446289136].

James, G., Witten, D., Hastie, T., Tibshirani, R.. (2017), An Introduction to Statistical Learning: with Applications in R, Springer-Verlag New York Inc., [ISBN: 9781461471370].

This module does not have any article/paper resources

Other Resources

W. N. Venables, D. M. Smith and the R Core Team. An Introduction in R,
https://cran.r-project.org/doc/manuals/r -release/R-intro.pdf

Website, Online Statistics Education: An Interactive Multimedia Course of Study,
http://onlinestatbook.com/

Website, Khan Academy,
http://www.khanacademy.org

Website, DataCamp,
http://www.datacamp.com