

J 111 - 01
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)

COIMBATORE - 641014

VALUE ADDED COURSE

DATA ANALYTICS

(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Programme Outcome	
PO1	To apply principles of Data Science to the analysis of Business problems
PO2	To apply quantitative modeling and data analysis techniques to the solution of real world business problems
PO3	To demonstrate knowledge of statistical data analysis techniques utilized in decision making
PO4	To employ tools and technologies to analyze data.
PO5	Enhance the ability to use techniques, skills and tools necessary for computing practice
PO6	To incorporate the data analysis framework for optimal performance
PO7	Organize data sets into meaningful structures, incorporating data profiling and quality standards.
PO8	To organize information efficiently in the visual form using appropriate software
PO9	To carry out standard data visualization and formulate inference from the results
PO 10	To provide a career in software industry
Programme Specific Outcome	
PSO1	Generation of simple statistical summaries using tools or software designed for statistical analysis
PSO2	Apply data science concepts and methods to solve problems in real-world contexts
PSO3	Develop the ability to build and assess data-based models.
PSO4	Build and enhance business intelligence capabilities by adapting the appropriate technology and software solutions.
PSO5	Communicate statistical information with proper inferences for future predictions

G. R. Damodaran

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)
EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

SCHEME OF EXAMINATIONS

Level	Course Code	Course Name	MARKS			Hrs. / Week	Exam. Duration (Hrs.)	Category
			CA	TEE	TOTAL			
Certificate	20319A	Quantitative Research Methods and Analysis	25	75	100	3	3	THEORY
Certificate	20319B	Data Analysis using R	25	75	100	3	3	THEORY
Certificate	20319P	Data Analytics Lab	40	60	100	3	3	PRACTICAL

Dr.

JIII - 02
2020-21

J III - 03
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

MAPPING OF COURSES WITH PROGRAMME OUTCOME LEVELS

Course Code	Course Title	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
20319A	Quantitative Research Methods and Analysis	1	1	2	2	1	1	2	1	2	2
20319B	Data Analysis using R	2	2	1	3	2	1	2	2	3	3
20319P	Data Analytics Lab	2	3	3	1	3	1	2	2	3	3

Indicators: 1. Reasonable 2. Significant 3. Strong

J11-04
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Level	Course Code	Course Title	Theory/ Practical	Problems %	Theory %
Certificate	20319A	Quantitative Research Methods and Analysis	Theory	50	50

Objective of the Course: The course gives an introduction to quantitative research methods and statistics. The focus of the course is on the basics of descriptive and inferential statistics.

UNIT I: Data Warehousing

(Teaching hours: 8)

Introduction – Definition – Multidimensional data model – OLAP operations – Warehouse schema – Data warehousing architecture – Warehouse server – Meta data – OLAP Engine – Data warehouse backend process – Cloud data warehousing - Other features.

UNIT II: Data Analytics

(Teaching hours: 8)

Statistical data analysis – Fact based decision making using statistical data analysis – Descriptive model – Predictive model - Analyze and predict results based on historical patterns - Apply statistical methods to economic data, problems and trends.

UNIT III: Statistics for Data Engineers

(Teaching hours: 8)

Mean, Median, Variance, Standard Deviation – Regression-Correlation Test: Chi square, Spearman Rank correlation, Pearson correlation coefficient – Parametric Test: period sample t-test, Two Independent t-test, One sample t-test.(only theoretical concepts)

UNIT IV: SPSS

(Teaching hours: 8)

An overview of SPSS-Creating and editing a data file-recording variables-exploring data-selecting cases-sorting cases-merging files-Creating and editing graphs and charts-Frequencies- bar charts- histograms- percentiles.

UNIT V: Measures of Central Tendency

(Teaching hours: 8)

Measures of central tendency- variability-deviation from normality- size and stability. Cross Tabulation and chi-square analyses-The means Procedure-Bivariate Correlation, Partial-Correlations -the correlation matrix-Independent samples- paired samples-one sample tests.

JIII -05
2020 -21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Course Outcome mapping with Knowledge level

Course Outcome	CO Statement	Knowledge level
CO1	To understand the basic concepts of data model and data warehouse	K1, K2
CO2	Analyze and implementation of data analytic methods	K2, K3
CO3	To implement various rank correlation and regression methods	K3, K4
CO4	To work with Statistical Package for the Social Sciences	K5
CO5	To infer with the measures of central tendency and visualization Techniques	K4, K5

Note: K1- Remembering; K2 – Understanding; K3 – Applying; K4 – Analysing; K5 – Creating & Evaluating.

Course Outcome mapping with Programme outcome

Course outcome	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO1	3	3	2	1	2				1	
CO2		1	3	2	2	2	1	2		
CO3		2	3	2	2	3	1	1		
CO4	1		3	2	3	3	1	2	3	
CO5		2	3	2	3		1	3	3	

Indicators: 1. Reasonable 2. Significant 3.Strong

Text Books				
S.No.	Title	Author	Publishers	Publication Year & Edition
1	The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning	Ralph Kimball, Joe Casertra	Wiley Inc	2004, First Edition
2	Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner	GalitShmueli, Nitin R. Patel and Peter C. Bruce	Wiley Inc	2007, First Edition
3	Statistics for Research With a Guide to SPSS, Third Edition	George Argyrous	Sage Publications.	2011, Third

Pedagogy: Lecture, PPT presentation, e-content seminar, Assignment, Quiz, Group Discussion

J III - 06
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Level	Course Code	Course Title	Theory/ Practical	Problems %	Theory %
Certificate	20319B	Data Analysis Using R	Theory	25	75

Objective of the course: The objective of the course is to gain more knowledge about data analytic tool R. The course component includes topics such as Data retrieval, Data Cleaning, Exploring and Visualizing data.

UNIT I: Preliminaries

(Teaching hours: 8)

Introduction: About R- Downloading and Installing R – Choosing an IDE: Emacs and ESS – Eclipse/Architect – RStudio – Revolution-R – Live-R –Editors – Help files - First program. Mathematical Operations and Vectors – Assigning Variables – Special Numbers – Logical Vectors – Inspecting variables and workspace.

UNIT II: Data Structure in R

(Teaching hours: 8)

Vectors – Matrices and Arrays – Lists – Data Frames – Environments – Functions – Strings – Factors.

Unit III: Looping concepts and Packages

(Teaching hours: 8)

Flow control and Loops – Advanced Looping: Replication – Looping over Lists – Looping over arrays – Multiple-Input Apply – Split-Apply-Combine – The plyr Package.

Unit IV: Data Analysis

(Teaching hours: 8)

Data Analysis: Built-in Datasets – Reading Text files – Reading Binary files – Accessing Database.

Cleaning and Transforming: Cleaning strings – Manipulating data frames – sorting.

Unit V : Data Visualization

(Teaching hours: 8)

Exploring and Visualizing: Scatter plots – Line plots – Histograms – Box plots – Bar charts – Other Plotting packages and systems

Jm - 07
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Course Outcome mapping with Knowledge level

Course Outcome	CO Statement	Knowledge level
CO1	To understand the preliminaries of R	K1
CO2	To work with the data structures in R	K2, K3
CO3	Understanding the looping concepts and working with the Packages in R	K3
CO4	Apply the process of Data Analysis, cleaning and transformation	K4, K5
CO5	Work with Data Exploration & Visualization techniques to interpret the results	K4, K5

Note: K1- Remembering; K2 – Understanding; K3 – Applying; K4 – Analysing; K5 – Creating & Evaluating.

Course Outcome mapping with Programme outcome

Course outcome	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO1			2		2			1	1	
CO2		2	3	2	2		1	2		
CO3		2	3	2	2	3	1	1		
CO4	1		3	2		3	1	2		2
CO5		2	3	2			1	2	3	3

Indicators: 1. Reasonable 2. Significant 3.Strong

Text Books				
S.No.	Title	Author	Publishers	Publication Year & Edition
1	Learning R	Richard Cotton	O'Reilly Media	2013, First Edition
2	R Programming for Data Science	Roger D. Peng	Lean Publishing	2014, First Edition
Web References				
3	R Data camp – Online Course Contents - https://campus.datacamp.com/courses/			

Pedagogy: Lecture, PPT presentation, e-content seminar, Assignment, Quiz, Group Discussion

JIV - 08
2020-21

Dr G R DAMODARAN COLLEGE OF SCIENCE (AUTONOMOUS)
COIMBATORE - 641014
VALUE ADDED COURSE
DATA ANALYTICS
(Certificate Level)

(Under Choice Based Credit System)

EFFECTIVE FOR THE STUDENTS ADMITTED DURING THE ACADEMIC YEAR 2020-2021

Level	Course Code	Course Title	Theory/ Practical	Problems %	Theory %
Certificate	20319P	Data Analytics Lab	Practical	-	-

Objective of the course: This course provides a hands-on experience on working with MS Excel, SPSS and R to perform quantitative and data analysis.

MS Excel

1. Pivot table
2. Chart
3. Descriptive analysis
4. Correlation and Regression

SPSS

5. Input data into SPSS
6. Drawing Graphs
7. Transposing Chart
8. Estimation and Hypothesis testing

R

9. Manipulating vectors
10. Working with Datasets
11. Creating Data Frames

Course Outcome mapping with Knowledge level

Course Outcome	CO Statement	Knowledge level
CO	The relevant programming abilities is developed and the Students will demonstrate proficiency with statistical analysis of data	K5

Course Outcome mapping with Programme outcome

Course outcome	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO	2	3	3	1	3	1	2	2	3	3

Pedagogy: Demonstration, Experiments