

VELAGAPUDI RAMAKRISHNA
SIDDHARTHA ENGINEERING COLLEGE::VIJAYAWADA
(AUTONOMOUS)
DEPARTMENT OF CE/CSE/ECE/EEE/EIE/IT/ME
MICRO LEVEL SYLLABUS

Class	B Tech	Regulation	VR20
Subject Code	20BS4101	Year & Semester	II & IV
Title of the Subject	Statistics with R		

Unit No	Content/Topics Covered (mention Sub Topics as found in books)	Text Book	Chapter/ Section No.	Page Number
Unit I	The R Environment:	[T1]	2	13
	Command Line interface		2.1	14
	R Studio		2.2	15
	Installing R Packages		3.1	29
	Basics of R:		4	35
	Basic math		4.1	35
	Variables		4.2	36
	Data types		4.3	38
	Vectors		4.4	43
	Calling function		4.5	49
	Missing data		4.7	50
	Data. Frames		5.1	53
	Lists		5.2	61
	Matrices		5.3	68
	Arrays		5.4	71
	Reading data into R:		6	73
	Reading CSVs		6.1	73
	Excel Data		6.2	74
	Statistical Graphs:		7	83
	Base Graphs		7.1	83
	ggplot2		7.2	86
	Writing R functions		8	99-104
	Hello World		8.1	
	Function arguments		8.2	
	Return Values		8.3	
	do.call		8.4	
	control statements:–		9	105
	If and else		9.1	105
	Switch		9.2	108
	Compound tests		9.4	111
	For loops		10.1	113
	While loops		10.2	115
Unit II	Group manipulation:	[T1]	11	117
	Apply Family		11.1	117
	aggregate		11.2	120
	Plyr		11.3	124
	data.table		11.4	129
	Data Reshaping:		12	141
	rbind and cbind		12.1	141
	joins		12.2	142
	reshape2.		12.3	149
	Strings:		13	155
	Paste		13.1	155

	Sprint		13.2	156
	Extracting text		13.3	157
	Regular expressions		13.4	161
	Math Functions:	[T2]	8	
			8.1	189
	Calculating a Probability		8.1.1	190
	Cumulative sums and products		8.1.2	191
	Minima and maxima		8.1.3	191
	Calculus		8.1.4	192
	Sorting		8.3	194
	Set operations		8.5	202
Unit III	Probability Distributions:	[T1]	14	171
	Normal Distribution		14.1	171
	Binomial Distribution		14.2	176
	Poisson Distribution		14.3	182
	Basic Statistics:		15	187
	Summary statistics		15.1	187
	correlation and covariance		15.2	191
	t-tests		15.3	200
	ANOVA		15.4	207
Unit IV	Linear Models:	[T1]	16	211
	Simple Linear Regression		16.1	211
	Multiple Regression		16.2	216
	Logistic Regression		17.1	233
	Poisson Regression.		17.2	237
	Nonlinear Models:		20	297
	Nonlinear least squares		20.1	297
	Splines		20.2	300
	Generalized additive models		20.3	304
	Decision trees		20.4	310
	Random forests.		20.5	312
	Time series:		21	315
	Autoregressive Moving Average		21.1	
	VAR		21.2	322
	GARCH		21.3	327
	Clustering: K Means		22.1	337
	PAM		22.2	345
	Hierarchical Clustering		22.3	352