



# INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal, Hyderabad – 500043

## COMPUTER SCIENCE AND ENGINEERING

### List of Laboratory Experiments

DATA SCIENCE LABORATORY								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P		C	CIA	SEE
BCSC12	Core	0	0	4	2	30	70	100
Contact Classes: Nil	Tutorial Classes: Nil	Practical Classes: 36			Total Classes:36			
Branch: CSE	Semester: III	Academic Year: 2021-22			Regulation: PG21			
<p><b>Course overview:</b> This course will expose the students to R programming environment, introduces to sampling and exploring data. It also provides a foundation in both probability theory and mathematical statistics and provides an indication of the relevance and importance of the theory in solving practical problems in the real world.</p>								
<p><b>Course objectives:</b> The students will try to learn:</p> <ol style="list-style-type: none"> <li>1. The concept of R objects.</li> <li>2. Make use of different types of datasets for analysis using R.</li> <li>3. Relations among variables using statistical analysis.</li> <li>4. The different data models for predictions using R</li> </ol>								
<p><b>Course outcomes:</b> After successful completion of the course, students will be able to:</p> <p>CO1 Make use of the container Data types for display the functional values. CO2 Demonstrate the reading and writing operations from web, disk. CO3 Illustrate data with effective visualizations for exploring the data. CO4 Outline the correlation and causation for relationship of different variables. CO5 Analyze the linear data models in data exploration. CO6 Develop the classification and clustering models to label the data.</p>								
WEEK NO	EXPERIMENT NAME							CO
WEEK – I	<b>R AS CALCULATOR APPLICATION</b>							CO1
	<ol style="list-style-type: none"> <li>a. Using with and without R objects on console</li> <li>b. Using mathematical functions on console</li> <li>c. Write an R script, to create R objects for calculator application and save in a specified location in disk</li> </ol>							
WEEK – II	<b>DESCRIPTIVE STATISTICS IN R</b>							CO1
	<ol style="list-style-type: none"> <li>a. Write an R script to find basic descriptive statistics using summary, str, quartile function on mtcars&amp; cars data sets.</li> <li>b. Write an R script to find subset of dataset by using subset (), aggregate () functions on iris dataset.</li> </ol>							
WEEK – III	<b>READING AND WRITING DIFFERENT TYPES OF DATASETS</b>							CO2
	<ol style="list-style-type: none"> <li>a. Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location.</li> </ol>							

	<ul style="list-style-type: none"> <li>b. Reading Excel data sheet in R.</li> <li>c. Reading XML dataset in R.</li> </ul>	
<b>WEEK – IV</b>	<b>VISUALIZATIONS</b>	<b>CO3</b>
	<ul style="list-style-type: none"> <li>a. Find the data distributions using box and scatter plot.</li> <li>b. Find the outliers using plot.</li> <li>c. Plot the histogram, bar chart and pie chart on sample data.</li> </ul>	
<b>WEEK – V</b>	<b>CORRELATION AND COVARIANCE</b>	<b>CO4</b>
	<ul style="list-style-type: none"> <li>a. Find the correlation matrix.</li> <li>b. Plot the correlation plot on dataset and visualize giving an overview of relationships among data on iris data.</li> <li>c. Analysis of covariance: variance (ANOVA), if data have categorical variables on iris data.</li> </ul>	
<b>WEEK – VI</b>	<b>REGRESSION MODEL</b>	<b>CO5</b>
	Import a data from web storage. Name the dataset and now do Logistic Regression to find out relation between variables that are affecting the admission of a student in an institute based on his or her GRE score, GPA obtained and rank of the student. Also check the model is fit or not. require (foreign), require (MASS).	
<b>WEEK – VII</b>	<b>MULTIPLE REGRESSION MODEL</b>	<b>CO5</b>
	Apply multiple regressions, if data have a continuous independent variable. Apply on above dataset.	
<b>WEEK –VIII</b>	<b>REGRESSION MODEL FOR PREDICTION</b>	<b>CO5</b>
	Apply regression Model techniques to predict the data on above dataset.	
<b>WEEK - IX</b>	<b>CLASSIFICATION MODEL</b>	<b>CO6</b>
	<ul style="list-style-type: none"> <li>a. Install relevant package for classification.</li> <li>b. Choose classifier for classification problem.</li> <li>c. Evaluate the performance of classifier.</li> </ul>	
<b>WEEK - X</b>	<b>CLUSTERING MODEL</b>	<b>CO6</b>
	<ul style="list-style-type: none"> <li>a. Clustering algorithms for unsupervised classification.</li> <li>b. Plot the cluster data using R visualizations.</li> </ul>	