



INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal, Hyderabad -500 043

COMPUTER SCIENCE AND ENGINEERING

TUTORIAL QUESTION BANK

Course Title	R PROGRAMING				
Course Code	ACS808				
Programme	B.Tech				
Semester	VII	CSE			
Course Type	SKILL				
Regulation	IARE - R16				
Course Structure	Theory			Practical	
	Lectures	Tutorials	Credits	Laboratory	Credits
	-	-	-	-	-
Chief Coordinator	Mr. S Laxman Kumar, Assistant Professor				
Course Faculty	Mr. S Laxman Kumar, Assistant Professor Ms. B Anupama, Assistant Professor				

COURSE OBJECTIVES:

The course should enable the students to:

I.	Summarize the fundamental knowledge on basics of data science and R programming.
II.	Develop programs in R language for understanding and visualization of data using statistical functions and plots.
III.	Understand a range of machine learning algorithms along with their strengths and weaknesses.
IV.	Learn to apply hypotheses and data into actionable predictions.
V.	Prepare documentation and present data in the form of graphs for multivariate data

COURSE OUTCOMES (COs):

I	Understand the process and different stages of data science and relevant data descriptions in R
II	Illustrate various SQL, NOSQL databases connecting with R and perform correlation and regression analysis
III	Evaluate different data models and perform clustering analysis.
IV	Solve various real time problems using artificial neural networks techniques and comparing different learning algorithms.
V	Explore on various ways to deliver results through documentation and plots of multivariate data and matrix data

COURSE LEARNING OUTCOMES (CLOs):

Sl. No.	Description
ACS808.01	Understand and develop relevant programming abilities
ACS808.02	Understand and intuition of the whole process line of extracting knowledge from data
ACS808.03	Equip with the fundamental knowledge on basics of data science and R programming
ACS808.04	Critically analyze and evaluate variety of NoSQL databases.
ACS808.05	Develop the ability to build and assess data-based models.
ACS808.06	Analyze data analysis and make models using regression analysis
ACS808.07	Familiarize with variety of machine learning tasks: clustering, dimensionality reduction, regression and classification
ACS808.08	Understand how to formalize practical problems using methods of machine learning
ACS808.09	Understand neural networks techniques solve real time problems
ACS808.10	Understand the different learning algorithms
ACS808.11	Chose a appropriate learning Algorithms to solve particular problems
ACS808.12	Based on delivering results make a documentation for various results sets
ACS808.13	Understand how to plot graphs for multivariate and matrix data

TUTORIAL QUESTION BANK

S. No	QUESTION	Blooms Taxonomy Level	Course Outcomes	Course Learning Outcomes
UNIT - I				
INTRODUCTION TO DATA SCIENCE AND R PROGRAMMING				
Part - A (Short Answer Questions)				
1	List the major steps involved in data science project.	Remember	CO 1	ACS808.01
2	Identify the steps in creating and running an R script?	Remember	CO 1	ACS808.02
3	Write the steps to create a calculator application with R objects?	Understand	CO 1	ACS808.02
4	Write the basic set of steps followed in data management?	Remember	CO 1	ACS808.02
5	State the any few data sources in data collection step.	Understand	CO 1	ACS808.02
6	Draw the block diagram of data science project process.	Remember	CO 1	ACS808.03
7	State the features of R Language.	Understand	CO 1	ACS808.03
8	Write the various classes of data types ANSI C supports?	Remember	CO 1	ACS808.03
9	State which of the following are valid identifiers. If not invalid, state the reason. a. Q b. C c. c d. t e. pi	Remember	CO 1	ACS808.05
10	Write expression in R a) $e^4 + \log 2$ b) $24X53$ c) $\log 210$ c) $\log 102$	Remember	CO 1	ACS808.03
11	What are the different forms of data types and how to test the data type in R? Give one example for each	Understand	CO 1	ACS808.02
12	Differentiate R while comparing with other programming languages?	Remember	CO 1	ACS808.03
13	How to set up the R environment?	Understand	CO 1	ACS808.02
14	Explain R as calculator using basic operations and inbuilt functions with suitable example?	Remember	CO 1	ACS808.02
15	What are different basic components in R?	Remember	CO 1	ACS808.02
16	How can you load a .csv file in R?	Understand	CO 1	ACS808.02
17	What are the different components of grammar of graphics?	Understand	CO 1	ACS808.02
18	What is R markdown? What is the use of it003F	Remember	CO 1	ACS808.02
19	How do you install a package in R?	Understand	CO 1	ACS808.05
20	Explain about confusion matrix in R?	Understand	CO 1	ACS808.05
Part - B (Long Answer Questions)				
1	Elaborate the following R objects. a)vector b)data frame c)matrixd)list	Remember	CO 1	ACS808.01
2	Compute the given <u>mathematical</u> formula and display on console in R. $X =$	Remember	CO 1	ACS808.01
3	Describe the multiple ways to read and write data from disc and web. Elaborate R functions to choose character, numerical input dynamically from user.	Remember	CO 1	ACS808.01
4	List the inbuilt summary functions to apply on vectors. Create a vector and apply all functions on it.	Understand	CO 1	ACS808.01
5	Identify the different ways to access the R objects. List the different data types in R with suitable example.	Understand	CO 1	ACS808.02

6	The price of one kg of rice is Rs. 40.75 and one kg of sugar is Rs. 30. Write R program to get the total amount of 2kg rice and 5kg sugar purchase.	Understand	CO 1	ACS808.02
7	Elaborate the process of working with different data files. Write about filechoose () function.	Remember	CO 1	ACS808.02
8	List the inbuilt summary functions to apply on vectors. Create a vector and apply all functions on it.	Understand	CO 1	ACS808.02
9	How to get system date in R? Generate sequence of previous and coming 10 dates from today in R.	Understand	CO 1	ACS808.02
10	What are the different ways to read the dataset? How to create and rename a variable in R? What are the read write methods available in R and explain?	Understand	CO 1	ACS808.02
11	Create an Array with name "MySales" with 30 observations using following methods: a) By using the array with dimensions 3, 5 and 2. b) By using Vector method.	Remember	CO 1	ACS808.02
12	That is data frame and how to create a data frame using the following data: Height GPA 66 3.80 62 3.78 63 3.88 70 3.72 74 3.69 Write an R program to prepare the inventory.	Understand	CO 1	ACS808.05
13	How will you identify and treat the missing value and outlier data in R?	Understand	CO 1	ACS808.05
14	Compare the different forms of data types and create a list object with restaurant menu items. Give inbuilt functions on list object with example Write R script to create and display list object of stores items having: {Fruits: {orange, mango, apple, watermelon, banana} Juices: {app, fruit, slice} Milkshakes: {Mango, papaya, sapota, pineapple}}.	Understand	CO 1	ACS808.05
15	Let two vectors x <- c(1,3, 5) and y <- c(3, 2, 10), What is size of the expression rbind(x, y) output? Differences between rbind and cbind. Create a matrix using rbind and cbind.	Remember	CO 1	ACS808.04
16	Create a data frame with a = c(1, 2, 3), b = c(4, 5, 6), c(7, 8, 9) and find the value of the following 1a.) How do I select the c (4, 5, 6)? 1b.) How do I select the 1? 1c.) How do I select the 5? 1d.) What is df [, 3]? 1e.) What is df[1,]? 1f.) What is df[2, 2]?	Understand	CO 1	ACS808.04
17	Find the output in R: If x is defined as x <- list(2, "a", "b", TRUE). What does x[[1]]? If x is defined as x <- list(2, "a", "b", TRUE). What does x[[2]] give me? iii) If x, y are two vectors x <- 1:4 and y <- 2:3. What is produced by If x, y are two vectors, x <- 1:4 and a vector y <- 2. What is produced by the expression x + y? Find x ³	Remember	CO 1	ACS808.03
18	Analyze the control structures with conditional statements in R with suitable example. Create a user defined function fact(j) to return the factorial of j using functions in R.	Understand	CO 1	ACS808.03
19	Describe the incremental process model advantages and disadvantages with the help of the diagram.	Remember	CO 1	ACS808.02
20	Name some functions available in "dplyr" package?	Understand	CO 1	ACS808.04

UNIT - II
SQL, NoSQL, DATA ANALYSIS

Part – A (Short Answer Questions)

1	State the packages and function required to run SQL queries in R?	Understand	CO 2	ACS808.06
2	List the packages and function required to read Excel sheet data to R object?	Understand	CO 2	ACS808.06
3	Explain the drawbacks in different ways to access Excel sheet in R.	Understand	CO 2	ACS808.06
4	What are other names of NoSQL?	Remember	CO 2	ACS808.06
5	List the features of NoSQL.	Remember	CO 2	ACS808.06
6	Give any five examples of NoSQL databases.	Understand	CO 2	ACS808.07
7	Differentiate SQL and No SQL databases.	Understand	CO 2	ACS808.07
8	Compare the SQL and NoSQL in terms of Data storage model and Schema.	Remember	CO 2	ACS808.07
9	What are the assumptions of regression Modelling	Understand	CO 2	ACS808.08
10	What is ANOVA?	Remember	CO 2	ACS808.08
11	Compare NoSQL & RDBMS	Remember	CO 2	ACS808.08
12	What is NoSQL?	Understand	CO 2	ACS808.08
13	What are the features of NoSQL?	Remember	CO 2	ACS808.08
14	Explain the difference between NoSQL v/s Relational database?	Understand	CO 2	ACS808.08
15	Explain “Polyglot Persistence” in NoSQL?	Understand	CO 2	ACS808.08
16	How does NoSQL DB budget memory?	Remember	CO 2	ACS808.08
17	Does NoSQL Database Interact With Oracle Database?	Understand	CO 2	ACS808.08
18	What is the difference between NoSQL & Mysql DBs’?	Remember	CO 2	ACS808.08
19	Explain Oracle NoSQL database?	Remember	CO 2	ACS808.08
20	When should use a NoSQL database instead of a relational database?	Understand	CO 2	ACS808.08

Part - B (Long Answer Questions)

1	Write the R script which include relevant packages and procedure to access .csv and .xsl files. Elaborate with suitable example.	Understand	CO 2	ACS808.08
2	What is the difference between SQL database and NoSQL Database?	Remember	CO 2	ACS808.06
3	Explain different approaches in R to connect with Excel	Remember	CO 2	ACS808.06
4	Define basic classification based on data model, with examples.	Understand	CO 2	ACS808.06
5	Explain Covariance and Correlation with example. Infer the relation among attributes with respect to correlation coefficient.	Understand	CO 2	ACS808.06
6	Calculate the correlation coefficient of iris dataset. What preprocess steps isneeded for iris data set.	Remember	CO 2	ACS808.06
7	Write an R program to find the correlation coefficient of iris data set and find the exact relation by using linear regression model.	Understand	CO 2	ACS808.07
8	Distinguish simple and multiple regression analysis and its applications working with numerical and categorical data?	Understand	CO 2	ACS808.07
9	Differentiate SQL and No SQL databases in detail?	Understand	CO 2	ACS808.08
10	What is forecasting give examples?	Understand	CO 2	ACS808.08
11	Write R program to extract sampleXML data from web and steps to convert as dataframe. Specify the needed packages and functions.	Remember	CO 2	ACS808.08
12	Distinguish simple and multiple regression analysis and its applications working with numerical and categorical data?	Remember	CO 2	ACS808.08
13	What are residuals? Define in Regression analysis.	Understand	CO 2	ACS808.08
14	Generate prediction model using linear regression for finding	Understand	CO 2	ACS808.08

	relative relation among variables. Write a R script to get a linear equation $y=mx+c$ form for the heart weight and body weight in cats dataset.			
15	Compute the covariance matrix and correlation matrix for the four numerical attributes. Interpret the statistical findings to know more about hidden nature in data.	Remember	CO 2	ACS808.08
16	Write an R script to connect with Excel, read the contents of sheet and load into R object.	Understand	CO 2	ACS808.08
17	Write a R program to perform the following: a. Import a data from webstorage. b. Name the dataset with suitable identifier c. Perform 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. d. Check the model is fit or not.	Remember	CO 2	ACS808.08
18	Write R program to perform the following: a. Find the correlation matrix of iris dataset b. Plot the correlation plot on dataset and visualize giving an overview of relationships among data on iris data. c. Perform analysis of covariance	Understand	CO 2	ACS808.08
19	How to script NoSQL DB configuration?	Remember	CO 2	ACS808.08
20	What is the CAP theorem? How is it applicable to NoSQL systems?	Understand	CO 2	ACS808.08

UNIT-III DATA MODELS

Part – A (Short Answer Questions)

1.	State the applications of regression models?	Remember	CO 3	ACS808.11
2.	List various types of regression models?	Understand	CO 3	ACS808.11
3.	State the difference between numerical and categorical parameters?	Understand	CO 3	ACS808.11
4.	Justify a single model on data is suggestible.	Remember	CO 3	ACS808.11
5.	State the various types of functions in R to support linear regression?	Understand	CO 3	ACS808.11
6.	State the various attributes to evaluate the multiple regression?	Understand	CO 3	ACS808.11
7.	State the residuals impact in linear model.	Remember	CO 3	ACS808.11
8.	What are the evaluating measures in regression models.	Understand	CO 3	ACS808.11
9.	What is the role of machine learning algorithms in data model.	Remember	CO 3	ACS808.11
10.	State the steps to evaluate the data model.	Understand	CO 3	ACS808.12
11.	Differentiate between univariate, bivariate and multivariate analysis?	Remember	CO 3	ACS808.11
12.	How will you define supervised and unsupervised learning?	Understand	CO 3	ACS808.11
13.	What you mean by Type I error and Type II error in Hypothesis testing?	Understand	CO 3	ACS808.12
14.	How will you evaluate your regression model based on R^2 , Adjusted R^2 and tolerance?	Remember	CO 3	ACS808.12
15.	What is Logistic regression? How will you evaluate your Logistic regression model?	Understand	CO 3	ACS808.13
16.	What is the difference between ANOVA and t-test?	Understand	CO 3	ACS808.13
17.	What is the difference between Overfitting and Underfitting?	Remember	CO 3	ACS808.15
18.	What are the steps involved in an analytics project?	Understand	CO 3	ACS808.14
19.	What all are the main packages used in Python for Data science and Machine Learning	Remember	CO 3	ACS808.13
20.	How will you define your number of clusters in K-Means clustering algorithm?	Remember	CO 3	ACS808.13

Part – B (Long Answer Questions)

1	Outline about the learning of a model? Write any four learning techniques and in each case give the expression for weight-updating.	Understand	CO 3	ACS808.11
2	Describe the limitations of the perception model. How to create and evaluate a data model. Describe with one case study.	Remember	CO 3	ACS808.11
3	List out the applications of Machine learning with example. Discuss about a consistent learner and what it means for a set of training examples to be linearly separable.	Remember	CO 3	ACS808.11
4	Write the steps to describe the process to create and evaluate the data model for the given data.	Understand	CO 3	ACS808.11
5	Predict whether an email is a spam and should be delivered to the Junk folder. Suggest suitable data model.	Understand	CO 3	ACS808.11
6	Discuss about clustering in detail? Elaborate k means algorithm.	Understand	CO 3	ACS808.12
7	Write a R script to implement	Understand	CO 3	ACS808.12
8	State about conditional probability in the bayes theorem for different type of data classification problem with suitable example.	Remember	CO 3	ACS808.12
9	Calculate the Jaccard coefficient for the given data $p = 1\ 0\ 0\ 0\ 0\ 0$ $q = 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1$;	Remember	CO 3	ACS808.13
10	List the different types clustering. Write about k-nn algorithm. Write a R script to cluster the mtcars dataset using k-nn algorithm.	Remember	CO 3	ACS808.13
11	Find the suitable Machine learning algorithm for the following real time problems: a. Predicting HousingPrices b. AnalyzingSentiment c. Finding Similar Documents d. RecommendingProducts	Understand	CO 3	ACS808.14
12	How can you identify, the best fit data model from the given dataset.	Remember	CO 3	ACS808.14
13	Describe the prediction model in terms of the following measures for best fit: Residual standard error, Multiple R-squared, F-statistic, p-value	Remember	CO 3	ACS808.14
14	How kNN is different from K-Means clustering?	Understand	CO 3	ACS808.14
15	What are the assumptions required for linear regression?	Remember	CO 3	ACS808.15
16	What are neural networks? What are the types of neural networks?	Understand	CO 3	ACS808.15
17	Why use artificial neural networks? What are its advantages?	Remember	CO 3	ACS808.15
18	How are artificial neural networks different from normal computers?	Understand	CO 3	ACS808.12
19	What is simple artificial neuron?	Remember	CO 3	ACS808.15
20	How Artificial Neurons Learn?	Understand	CO 3	ACS808.15

**UNIT-IV
ARTIFICIAL NEURAL NETWORKS AND EVALUATION**

1	Draw the basic structure of neural network.	Understand	CO 4	ACS808.16
2	List various types of problems for neural network learning?	Remember	CO 4	ACS808.16
3	State the difference between neural network and multi-layer neural network?	Understand	CO 4	ACS808.16
4	What is meant by perceptions in neural network?	Understand	CO 4	ACS808.16
5	State the various types of layers in neural network design?	Remember	CO 4	ACS808.16
6	State the remarks in multi-layer perceptions?	Understand	CO 4	ACS808.16
7	State the role of error function in neural network.	Understand	CO 4	ACS808.16
8	What are the evaluating measures of neural network model?	Remember	CO 4	ACS808.16
9	Define Multi-layer networks	Understand	CO 4	ACS808.18
10	Give a short note on feed forward networks	Remember	CO 4	ACS808.19

11	How artificial neural networks can be applied in future?	Remember	CO 4	ACS808.19
12	What can you do with an nn and what not?	Understand	CO 4	ACS808.19
13	What are cases and variables?	Understand	CO 4	ACS808.17
14	What are the population, sample, training set, design set, validation set, and test set?	Remember	CO 4	ACS808.17
15	How to avoid overflow in the logistic function?	Understand	CO 4	ACS808.18
16	What are common uses and/or applications for AI?	Understand	CO 4	ACS808.18
17	Why is image recognition a key function of AI?	Remember	CO 4	ACS808.18
18	Explain Neural Network architecture.	Understand	CO 4	ACS808.20
19	Explain Unsupervised Learning Neural Networks.	Understand	CO 4	ACS808.16
20	Explain Supervised Learning Neural Networks.	Understand	CO 4	ACS808.16
Part – B (Long Answer Questions)				
1	Outline about the learning of a model? Write any four learning techniques and in each case give the expression for weight-updating.	Understand	CO 4	ACS808.17
2	Describe the limitations of the perception model. How to create and evaluate a data model. Describe with one case study.	Understand	CO 4	ACS808.17
3	List out the applications of Machine learning with example. Discuss about a consistent learner and what it means for a set of training examples to be linearly separable.	Understand	CO 4	ACS808.19
4	Write the steps to describe the process to create and evaluate the data model for the given data.	Remember	CO 4	ACS808.16
5	Predict whether an email is a spam and should be delivered to the Junk folder. Suggest suitable data model.	Understand	CO 4	ACS808.16
6	Give the basic structure of neural network and different types of ANN with real time examples.	Understand	CO 4	ACS808.19
7	State the suitable classification algorithm for pima Indian diabetes dataset classification. Improve the accuracy by performing multiple preprocess steps.	Remember	CO 4	ACS808.19
8	How to evaluate hypothesis of different types with examples? Explain the basics of sampling theory.	Understand	CO 4	ACS808.18
9	Discuss the difference of error in two hypotheses. Differentiate the MAP (maximum a posteriori) and ML (maximum likelihood) hypothesis. Give an example of a scenario in which a MAP hypothesis is preferable to an ML hypothesis.	Remember	CO 4	ACS808.19
10	Compare the learning algorithms with example in terms of problem nature, accuracy and error rate.	Understand	CO 4	ACS808.20
11	Find the suitable Machine learning algorithm for the following real time problems: a. Predicting Housing Prices b. Analyzing Sentiment c. Finding Similar Documents d. Recommending Products	Remember	CO 4	ACS808.19
12	How can you identify, the best fit data model from the given dataset.	Remember	CO 4	ACS808.19
13	Describe the prediction model in terms of the following measures for best fit: Residual standard error, Multiple R-squared, F-statistic, p-value	Remember	CO 4	ACS808.18
14	How to evaluate hypothesis of the given problem. Describe the basic principle of sampling theory.	Understand	CO 4	ACS808.16
15	List the steps to improve the accuracy of neural network data model.	Understand	CO 4	ACS808.20
UNIT-V				
GRAPHS AND PLOT MODELS				
Part - A (Short Answer Questions)				

1	State the applications of regression models?	Understand	CO 5	ACS808.21
2	List various types of regression models?	Remember	CO 5	ACS808.21
3	State the difference between numerical and categorical parameters?	Remember	CO 5	ACS808.21
4	Justify a single model on data is suggestible.	Remember	CO 5	ACS808.21
5	State the various types of functions in R to support linear regression?	Understand	CO 5	ACS808.21
6	State the various attributes to evaluate the multiple regression?	Remember	CO 5	ACS808.21
7	State the residuals impact in linear model.	Understand	CO 5	ACS808.22
8	What are the evaluating measures in regression models.	Understand	CO 5	ACS808.22
9	What is the role of machine learning algorithms in data model.	Remember	CO 5	ACS808.22
10	State the steps to evaluate the data model.	Understand	CO 5	ACS808.22
11	What is Graphical Data Analysis with R?	Remember	CO 5	ACS808.22
12	What is Histograms?	Remember	CO 5	ACS808.22
13	How to set the axis labels and title of the r plots?	Understand	CO 5	ACS808.24
14	How to create a Stack Bar Chart?	Remember	CO 5	ACS808.24
15	How to create a Bar Chart ?	Remember	CO 5	ACS808.23
16	How to create a Histogram ?	Understand	CO 5	ACS808.23
17	How to create a Scatter Plot?	Understand	CO 5	ACS808.24
18	How to create a Box Plot ?	Understand	CO 5	ACS808.24
19	How to create an Area Chart ?	Remember	CO 5	ACS808.25
20	How to create a heat map ?	Understand	CO 5	ACS808.24
Part - B (Long Answer Questions)				
1	Generalize the graphical analysis in data analysis? List the various plots in R and explain in detail.	Understand	CO 5	ACS808.22
2	How to plot the word (text) data based on frequency of words. Write R script to plot a data frame having: {df1: {sea,river,pond,lake,pool} df2: {3,5,8,10,34}} using relevant plot.	Remember	CO 5	ACS808.22
3	List out the applications of Machine learning with example. Discuss about a consistent learner and what it means for a set of training examples to be linearly separable.	Remember	CO 5	ACS808.24
4	Write the steps to plot the numerical data at different granularity levels.	Understand	CO 5	ACS808.24
5	Plot the k-means clustering algorithm outcome. State the relevant packages and functions in R	Remember	CO 5	ACS808.24
6	How would you get the multiple plots in single window?	Remember	CO 5	ACS808.22
7	Elaborate how to export a graph using graphics parameters. How to export the text data to plot with example.	Understand	CO 5	ACS808.23
8	Describe boxplot() of iris\$petal length attribute. Specify the observations of plot.	Remember	CO 5	ACS808.21
9	Plot the regression model along with residuals.	Remember	CO 5	ACS808.21
10	Write a R script for creating a boxplot of iris sepal length attribute.	Understand	CO 5	ACS808.21
11	Describe the various plots in R to visualize the data and explain the purpose of each plot in detail.	Remember	CO 5	ACS808.21
12	Write R script to plot a data frame having: {df1: {red,green,blue,pink,black} df2: {3,5,8,10,34}} using relevant plot.	Remember	CO 5	ACS808.22
13	List out the steps to plot the data models with relevant packages.	Remember	CO 5	ACS808.22
14	How to partition the window to get more number of plots. Discuss on single and multi object plots in R.	Understand	CO 5	ACS808.23
15	Discuss about the residuals with respect to observed values? State a case study to show the fitted line and residuals in logistic regression.	Remember	CO 5	ACS808.21

16	How to create a Correlogram ?	Remember	CO 5	ACS808.22
17	How to plot a geographical map ?	Understand	CO 5	ACS808.23
18	How to plot a data set in single command ?	Remember	CO 5	ACS808.21
19	What is Multiple Linear regression?	Remember	CO 5	ACS808.22
20	What is Machine Learning?	Understand	CO 5	ACS808.23

Prepared By:

Mr. S Laxman Kumar, Assistant Professor

HOD, CSE