



**DEPARTMENT OF COMPUTER ENGINEERING
ONE WEEK ONLINE FACULTY DEVELOPMENT PROGRAM ON
"R-Language for Analytics and Data Science"**

**Organized by
Department of Computer Engineering, Jaihind College of Engineering, Kuran
(Pune)**

In association with IIT Bombay Spoken Tutorial

on

Sunday, 10th May, 2020 to Friday, 15th May, 2020

FDP SCHEDULE

Day #	Session #	Topics to be covered	Activities
Day 1 10/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Installing R and RStudio on Linux • Installing R and RStudio on Windows • Introduction to basics of R 	Find out how to calculate median-using Help button and web search
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Introduction to Data Frames in R • Introduction to RStudio • Introduction to R script 	<ul style="list-style-type: none"> • Calculate mean and median of the data frame CO2 • Take the help of R for all the commands shown in this tutorial • Import dataset from the internet directly or through file • Check whether plyr package is installed on your machine. If yes load it. •
Day 2 11/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Working directories in RStudio • Indexing and Slicing Data Frames 	<ul style="list-style-type: none"> • Create a new folder on your computer and make it your working directory. • Create a script and save it on desktop as testscript.R • Load myfirstscript.R (created in this tutorial) in R and run it • Create a subset from

			captaincy data frame with the captains who have played > 20 matches and lost < 14 matches.
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Creating Matrices using Data Frames • Operations on Matrices and Data Frames 	<ul style="list-style-type: none"> • Consider 2 vectors $c(9,10,11,12)$ and $c(13,14,15,16)$. Create a 4 by 2 matrix from these two vectors. • Add another vector $c(17,18,19,20)$ as a column to the previous matrix. •
Day 3 12/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Merging and Importing Data • Data types and Factors 	<ul style="list-style-type: none"> • Using built-in dataset iris, implement all the functions we have learnt in this tutorial • . Using built-in dataset iris, find out the categorical variables. • Can you find a variable which is categorical, but R reads as numeric? If yes, change it to categorical.
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Lists and its Operations 	<ul style="list-style-type: none"> • Create a numeric vector $c(1:5)$ and a 5 by 3 matrix with elements from 1 to 15. • Create a named list with vector, matrix and iris data set. • Retrieve the iris data set from the list using dollar operator and indexing method. • State the differences between the results obtained by using dollar operator and indexing method of accessing iris
Day 4 13/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Plotting Histograms and Pie Chart • Plotting Bar Charts and Scatter Plot 	<ul style="list-style-type: none"> • Read the file moviesData.csv. Create a histogram of the object named imdb_num votes in this file. • Create a pie chart of the object mpaa_rating. • Save both the plots. • Read the file moviesData.csv. Create a bar chart of critics score for the first 10 movies.

			<ul style="list-style-type: none"> • Create a scatter plot of imdb_rating and imdb_num votes to see their relation. • Save both the plots.
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Introduction to ggplot2 • Aesthetic Mapping in ggplot2 	<ul style="list-style-type: none"> • Consider the built-in data set mtcars. Find the numerical variables in this data set. • Make a scatter plot from the objects named mpg and wt in this data set. • Save the plot in .jpeg format.
Day 5 14/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Data Manipulation using dplyr Package • More Functions in dplyr Package 	<ul style="list-style-type: none"> • Consider the built-in data set mtcars. Find the cars with hp greater than 100 and cyl equal to 3. • Arrange the mtcars data set based on mpg variable. • Use the built-in data set airquality. Using select function, select the variables Ozone, Wind, and Temp in this data set. • Use the built-in data set mtcars. Rename the variables mpg and cyl with MilesPerGallon and Cylinder, respectively.
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Pipe Operator 	<ul style="list-style-type: none"> • 1. Use the built-in data set iris. Using the pipe operator, group the flowers by their Species. • Summarise the grouped data by the mean of Sepal.Length and Sepal.Width.
Day 6 15/05/2020	Session – 1 10.00 am to 12.30 pm	<ul style="list-style-type: none"> • Conditional Statements 	<ul style="list-style-type: none"> • Use the built-in data set iris. Find the Species, in which Sepal.Length is greater than Petal.Length. • Count all such Species.
	Session – 2 2.00 pm to 4.30 pm	<ul style="list-style-type: none"> • Functions in R 	<ul style="list-style-type: none"> • Create a function which computes combination of two numbers. • Create a function which takes a natural number as an argument, and prints Fibonacci series. For example, consider fibonacci(5). It should print the first 5 elements of Fibonacci series, i.e. 1, 1, 2, 3, 5.

Note:

1. You have to watch all 22 videos given in R Test course at Spoken Tutorial web site according to day wise and session wise.
2. Do practice simultaneously to clear your doubts.
3. Read Instruction sheets carefully before goin to R Test.

Date of Exam: Saturday, 16th May, 2020

Ms S.B. Jadhav

Computer Faculty

Prof. C. S. Arya

FDP Coordinator

Prof. Wavhal D.N.

HOD