

MGIT - Hyderabad
B.Tech in INFORMATION TECHNOLOGY
DATA SCIENCE using R Programming
Date : 06/08/2024 to 09/08/2024

Course Objectives:

- Understand the R Programming Language.
- Exposure on Solving data science problems.
- Understand The classification and Regression Model

Course Outcomes:

- Illustrate the use of various data structures.
- Analyze and manipulate Data using Pandas
- Creating static, animated, and interactive visualizations using Matplotlib.
- Understand the implementation procedures for the machine learning algorithms.
- Apply appropriate data sets to the Machine Learning algorithms
- Identify and apply Machine Learning algorithms to solve real-world problems.

Date	Day	LIST OF EXPERIMENTS
06/08/2024	Day 1	1- R AS CALCULATOR APPLICATION
		a) Using with and without R objects on console
		b) Using mathematical functions on console
		c) Write an R script, to create R objects for calculator application and save in a specified location in disk
		2- DESCRIPTIVE STATISTICS IN R
		a) Write an R script to find basic descriptive statistics using summary
		b) Write an R script to find subset of dataset by using subset ()
		3- READING AND WRITING DIFFERENT TYPES OF DATASETS
		a. Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location.
		b. Reading Excel data sheet in R.
		c. Reading XML dataset in R.
07/08/2024	Day 2	4- VISUALIZATIONS
		a) Find the data distributions using a box and scatter plot.
		b) Find the outliers using a plot.
		c) Plot the histogram, bar chart and pie chart on sample data
		5- CORRELATION AND COVARIANCE
		a) Find the correlation matrix.
		b) Plot the correlation plot on dataset and visualize giving an overview of relationships among data on iris data.
		c) Analysis of covariance: variance (ANOVA), if data have categorical variables on iris data
		6- REGRESSION MODEL
		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 a institute based on his or her GRE score, GPA

08/08/2024	Day 3	obtained and rank of the student. Also check the model is fit or not.
		require (foreign), require (MASS).
		7- MULTIPLE REGRESSION MODEL
		Apply multiple regressions, if data have a continuous independent variable. Apply on above dataset.
		8- REGRESSION MODEL FOR PREDICTION
		Apply regression Model techniques to predict the data on above dataset
09/08/2024	Day 4	9- CLASSIFICATION MODEL
		a) Install relevant packages for classification.
		b) Choose a classifier for classification problems.
		c) Evaluate the performance of the classifier.
		10) CLUSTERING MODEL
		a) Install relevant packages for classification.
		b) Choose a classifier for classification problems.
		c) Evaluate the performance of the classifier.
Datasets Detils: Visualizations: Iris data, Regression: Admission_Predict, Classification: Exploring Survival on the Titanic		