Data Science Using R

Course Objectives

Data science & Big Data Analytics have been touted as the most promising domains f or IT professionals . Getting into this domain early on can have a far - reaching impact on your career . This course will provide you a comprehensive and hands - on training on R programming .

Objective of the course are as follows:

- To help you understand Big data analytics concept
- To enable you to get familiar with the life cyc le and the methodology of Analytics
- To help you get a hands - on Machine Learning experience on R

W ho s h o u l d attend this c o u r s e ?

- This course is meant f or IT professionals, who are into various development roles and would like to get into data science/business analytics?

- This course can be taken by data base professionals, who wish to enter into data science domain .

Course Content

Main Modules	

Introduction	What is Data Science
	Data Science life cycle
	Data Science phases :
	    ▫ Data Acquisition
	    ▫ Data Cleaning
	    ▫ Data Manipulation
	    ▫ Data Analysis (Statistical and machine learning algorithms) to make sense of data
	    ▫ Data Visualization

Introduction to R Language	▫ What is R ? ▫ Data science & R ▫ Components of R ▫ Installing R ▫ Using command line in R ▫ Introduction to R Studio (IDE) ▫ Finding Help & solving issues in R
R Language Constructs	▫ Data types in R ▫ Program Structure in R ▫ Flow Control: For loop ▫ If condition ▫ While conditions ns and repeat loop ▫ Debugging tools ▫ Concatenation n of Data ▫ Combining Vars , c bind, r bind ▫ Sapply, apply, tapply functions ▫ Buillt - in functions in R

Using R for Analytics – Data Acquisition & Cleaning	□ R d at a impor t p ac k age introduction □ Importing g da ta f rom various sources □ Tools f or D at a Acquisition □ U s in g R f or data cleaning
Using R for Analytics – Data Manipulation	□ Appendi ng da ta to a v ec tor □ C om bin in g m ult i pl e ve c tor s □ Lis t m anag em ent □ Mer gi ng d ataf r am es □ D at a trans formati on □ Strings and dates □ Outlier detection □ H and ling N As a nd M is s in g Va lu es □ Matr ic es a nd Ar r a ys □ Lo gic al o perat io ns □ R e lat i on al o per ator s □ Ac c es s i ng V ar ia bl es □ Matr ix Mu lt i pl ic a ti on a nd In v er s i on □ Managi ng S ubs et of d ata □ C har ac t er m ani pu la ti o n □ D at a ag gr e gat i on □ Su bsc r i pt in g
Statistics with R	□ C om put in g b as ic s ta ti s tic s □ C om par i ng m ea ns of t w o s am ples □ T es ting a c or r e l at io n f or s i gn if ic a nc e □ T es ting a pr op or t io n □ C l as s ic a l tes ts (t, z,F) □ An al ys is of v ar i anc e (AN O V A) □ Sum m ar i zi ng D a ta □ D at a Mu ng in g Bas ic s
Machine Learning & Predictive Modeling	□ W hat is m achine lear n in g □ Su per v is e d a n d unsupervise d learning m od els □ Decision Trees □ R andom Forests □ K - m eans c lus t er i ng □ N aïve Bayes C classification Associ at i on R u les M in in g □ Examples
Using R for Analytics – Data Visualization	□ W hat is D a ta V is u al i z at io n? □ D at a Vis ua l i zat i on to o ls in th e m ar k et □ U s ing graphic al f unctions in R f or data visualizati on □ Li ne Plots □ Bar P lo ts

	□ Bar P lots f or Populati on
	□ Histogram & B ox P lot
	□ Pi e chart
	□ Plotting with bas e graphic s
	□ Plotting w it h Lattice graphic s
	□ Plotting and coloring in R
Project Case Study	□ Two c as e s t u dies will be us e d t o showcas e t he learnings and also f or participants to practice c e