



**Engineering Staff College of India**  
Autonomous Organ of The Institution of Engineers (India)  
Old Bombay Road, Gachi Bowli, Hyderabad – 500 032. TS, India

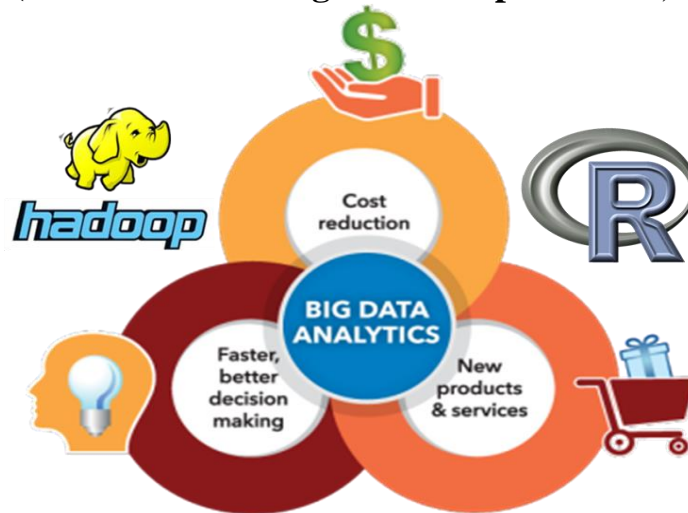


**INFORMATION TECHNOLOGY DIVISION**

**FACULTY DEVELOPMENT PROGRAMME**

# **Big Data Analytics Using Hadoop & R tool**

(Hands-on Training on Hadoop & R tool)



**13 – 16 March 2018**



(An ISO 9001:2008 Certified, AICTE & CEA Recognized Institution)

**Centre for Promotion of Professional Excellence**

## Introduction

In today's cut throat competition, business firms are focusing on predicting responses to proposed actions before actually spending valuable resources like people, time and money all of which are crucial for building an effective decision making. Data is the most crucial element for any organization. Irrespective of the size of the business, transactional data is growing exponentially with time in terms of volume challenging not only the computational capacity but also timely decision making resulting leaving great difficulties for business managers. As such organizations are heavily relying on the process of examining the ever increasing data to uncover hidden patterns, unknown correlations and other useful information.

**Data Analytics** on big data enables business firms to continuously drive innovation and make the best possible decisions. Business Managers, Data Analysts, Data Managers etc. can conduct data driven experiments to interpret data and create data models by using the most versatile tool like **Hadoop / R Tool** for identifying hidden data patterns from historical data so as to design new strategies for product and services marketing.

**Hadoop** is a leading open source Big Data framework. 73% of organizations have already invested or plan to invest in big data by 2016. The Big Data industry is expected to grow to \$48.4 billion by 2018, as per Forbes estimates. This course trains you in Hadoop components like HDFS and MapReduce and augments that with other ecosystem components like Hive, Pig and HBase.

Advanced **R Tool** allows participants to realize the power of advanced statistical analysis to understand data characteristics, identify trends and relationships among data.

This Training is organized as an interactive session to provide an understanding and effective application of advanced tools for Data Analytics like Hadoop & R tool.

## Objectives

- The objective of this Training is to provide Hands-on practices Training in the space of data management and use of **Hadoop / R tool/ Spark** for rapid analytics.

## Course Coverage

- **Overview about Data Analytics & Bigdata**
- Popular Tools
- Role of Data Scientist
- Analytics Methodology
- Problem Definition
- Technical Presentation on Data Mining.
- **Data Analytics using R tool**
- Introduction to R
- Data Exploration with R
- Understanding R data structure
- Data Exploration with R, Importing data
- Manipulating Data, Using functions in R
- R Programming
- Data Visualization with R Linear
- Case Study
- **Data Analytics using Hadoop**
- Introduction Hadoop
- Hadoop Environment Setup
- HDFS Architecture
- HDFS Commands
- MapReduce
- MapReduce Programs
- MapReduce and Job Execution
- Hadoop Serialization and Compression
- Advanced MapReduce Programming
- Streaming
- Introduction to **Spark**
- Machine learning with Spark

## Methodology

Methodology of the programme includes class room sessions with Lecture and Hands-on Practical's / case studies, debates, sharing of experiences etc. All the sessions will be interactive demanding active participation from all the members.

## Target Participants

- Faculty / Professors / Research scholars/ Technical staff members of Engineering Colleges and Universities.
- Business Managers, Data Managers, Data Analysts, Business executives Technical Managers, Scientist and People working in Government, Public sector, Private sector and Defence organizations
- ETL developers/ Project Managers / Testing Professionals.

## Benefits to the participants

- Gain insights about identifying target customers for business.
- Understand factors techniques for refining business processes that influence business outcome
- Understand the need for data analytics for maximizing organizational profitability.
- Hands-on practical experience on Hadoop & R tool.

## Programme Venue, Dates & Timings

**Venue:** Engineering Staff College of India (ESCI) Campus, Old Bombay Road, Gachi Bowli, Hyderabad. 500032. TS, India.

### Dates

**13 – 16 March, 2018**

### Timings

On the first day Registration will commence at 09:00 hrs. On all other days the programme timings will be from 09:45 – 17:15 hrs with breaks in between for tea and lunch.

## Course Director



**Mr. Syed Azgar , MBA(IT), RHCE, MCSA**  
Faculty & Manager-IT,  
Information Technology Division,  
Engineering Staff College of India, Hyderabad.

## Course Fee

**₹ 20,000/- (Residential Fee)** per participant. Fee includes course material, course kit, twin-sharing / single AC accommodation as per availability, Breakfast, Lunch, Dinner, Tea / Coffee and Snacks during the actual days of training programme.

## DISCOUNTS

- ❖ **Non – Residential Fee** - 10% discount on course fee is allowed for non- residential participants.
- ❖ **Group Discount** - Additional 10% discount for three or more participants if sponsored by the same organization.

**(All discounts are applicable only if fee is received at ESCI before commencement of the programme)**

**GST @18%** is to be paid extra and above the training fee as training is also brought under the purview of Service Tax in Finance Bill 2010. **PAN Card No.** AAATT3439Q. **GST No:** 36AAATT3439Q1ZV, **HS No.:** 999293 (under commercial training or coaching services – clause 65(105) (ZZC) of Finance act – 1994).

Programme fee is to be paid in in favour of **“THE INSTITUTION OF ENGINEERS (INDIA) – ENGINEERING STAFF COLLEGE OF INDIA”** in the form of demand draft payable at Hyderabad. Alternatively the payment may be made by **Electronic Fund Transfer (EFT)** to ESCI – **Axis Bank A/c No. 912010049234564** with The Axis Bank Ltd, Old Mumbai Hwy, Cyberhills Colony, P Janardhan Reddy Nagar, GachiBowli Hyderabad-500032 by NEFT/ RTGS/ IFSC Code No. UTIB 0000733 – MICR No.500211020. **While using EFT method of payment, please ensure to communicate us your company name, our invoice reference and programme title.**

**Register Online:** <http://www.escihyd.org/trainings/information-technology/>

To register, manually please send your nominations giving details of name, designation, contact address, email address, mobiles no, telephone and fax number of the participant along with the details of mode of payment of fee, addressed to:

### The Head

Information Technology Division  
Engineering Staff College of India  
Gachi Bowli, Hyderabad – 500 032  
Phone: 66304100 (EPABX) / 040 – 66304123/24/25 (Direct), Fax: 040 - 23000336  
Email: [it@escihyd.org](mailto:it@escihyd.org), Portal: [www.escihyd.org](http://www.escihyd.org)

## A Certificate of participation will be awarded to each participant on conclusion of the programme.

- ESCI encourages participants to present case studies from their respective organizations.
- For the convenience of outstation participants, ESCI will facilitate pick-up and drop from Airport / Railway Stations / Bus Stations, if travel plans are received at least 3 days in advance along with mobile number by fax or email. The charges shall be paid by the participant directly to the Cab.
- ESCI provides complimentary accommodation and boarding to the participants one day before commencement (Check-in 1200 h) and one day after conclusion (Check-out 1200 h) of the programme duration. Overstay charges will be applicable as per ESCI rules (subject to availability of accommodation)
- Well developed Information Centre and Internet facilities are available to the participant