



# Hindi Vidya Prachar Samiti's RAMNIRANJAN JHUNJHUNWALA COLLEGE (AUTONOMOUS)

(Also known as R. J. College of Arts, Science & Commerce as per UGC Notification)

Affiliated to UNIVERSITY OF MUMBAI II Recognized by UGC under 2f & 12B  
NAAC Accredited 'A GRADE' with CGPA 3.50

Knowledge is all Ambrosia

CERTIFICATE  
COURSE IN

BIG DATA  
AND  
HADOOP

A

T

S

M

C



[www.rjcollege.edu.in](http://www.rjcollege.edu.in)



[rjcollege@rjcollege.edu.in](mailto:rjcollege@rjcollege.edu.in)



+91 22 25151763



Opposite Railway Station, Ghatkopar (W),  
Mumbai 400 086, Maharashtra, INDIA.



# ABOUT US

Hindi Vidya Prachar Samiti was incepted on the auspicious day of Shri Krishna Janmashtami, 15th August 1938. A brain child of a visionary Late Shri Nandkishore Singh Jairamji, samiti was established with the objectives of catering to the educational needs of the Hindi speaking community. Ramniranjan Jhunjhunwala College came into existence in 1963, enabling a larger section of the society to take advantage of the facilities provided for higher education.

From 1999-2000 the College has added a number of self-financing courses like B.M.S., B.B.I., B.Sc. in Computer Science, Information Technology, Biotechnology, M.Sc. in Computer Science, Biotechnology and Information Technology as well as add on courses, which further hone the special skills of the students.

The college has been reaccredited with 'A' Grade by NAAC in 2014 with a CGPA 3.50 and received the Best College Award (2007-2008) of the University of Mumbai. The College has been bestowed with IMC "Ramkrishna Bajaj Performance Excellence Trophy", 2010.

The Principal of the college was awarded "Best Teacher" by Government of Maharashtra in 2011.

Government of Maharashtra conferred the college with "JAAGAR JAANIVANCHA" (First in Mumbai Suburban- in 2013 and Second in Mumbai Suburban- in 2014) for safety of girls.

**C**ourse Code: **RJCSBDH**

**D**uration: **30** hours

**C**redits : **02**

# COURSE OUTCOME

The students will be able to:

- Identify Big Data and its Business Implications.
- List the components of Hadoop and Hadoop Eco-System
- Access and Process Data on Distributed File System
- Manage Job Execution in Hadoop Environment
- Develop Big Data Solutions using Hadoop Eco System
- Analyse Infosphere Big Insights Big Data Recommendations





# COURSE OBJECTIVE

- Understand the Big Data Platform and its Use cases
- Provide an overview of Apache Hadoop
- Provide HDFS Concepts and Interfacing with HDFS
- Understand Map Reduce Jobs
- Provide hands on Hadoop Eco System
- Apply analytics on Structured, Unstructured Data.
- Exposure to Data Analytics with R.

## Unit I

- Understand different technology trends, Big Data market and different job roles in Big Data.
- Understand what Hadoop is for, and how it works.
- Master the concepts of Hadoop framework.
- Understand complex architectures of Hadoop and its component

## Unit II

- Hadoop installation on your machine
- Setup and installation of Hive, pig, sqoop, hbase
- Hbase architecture and data saving paradigm
- Understand how MapReduce, Hive and Pig can be used to analyze big data sets
- Demos: Running HDFS commands, Hive queries, Pig queries MapReduce program execution with Bank Dataset
- Sample data sets and scripts (HDFS commands, Hive sample queries, Pig sample queries, Data Pipeline sample queries)

## Unit III

- Start writing your own codes in Hive and Pig to process huge volumes of data
- Design your own data pipeline using Pig and Hive
- Understand modern data architecture: Data Lake

## Unit Iv

- Apache spark and its benefits over Hadoop
- Apache spark and different language of approach for data analyzing
- Apache Sqoop share data from severe data to Hdfs hadoop
- Apache kafka introduction and project use case
- Hands on practices of code in lab

# COURSE CONTENT

## MEDIUM OF INSTRUCTION

English



## EVALUATION

Continuous evaluation theory and practical's (Virtual)

**40% PASSING**



## WHO SHOULD DO IT?

For SY/TY BSC Computer Science students