

HADOOP SCHOOL OF TRAINING

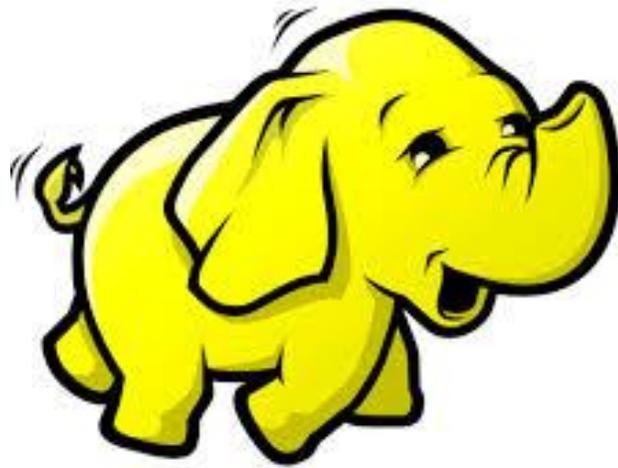


Data Analyst Program- 0 to 100



HADOOP SCHOOL OF TRAINING

Data Analyst Program- 0 to 100



Basics

Learn the basics of Big Data and hadoop

Hands On

Play with Hadoop and hadoop ecosystem

Data Analysis

Become a top notch data analyst

Data Analyst Program- 0 to 100 (40 Hours)



Overview of the course:

Data Analyst Program is a one stop course that introduces you to the domain of data analysis as well as gives you technical knowhow of the same. At the end of this course you will be able to earn a credential of data analyst and you will be capable of dealing with Terabyte scale of data and analyze it successfully.

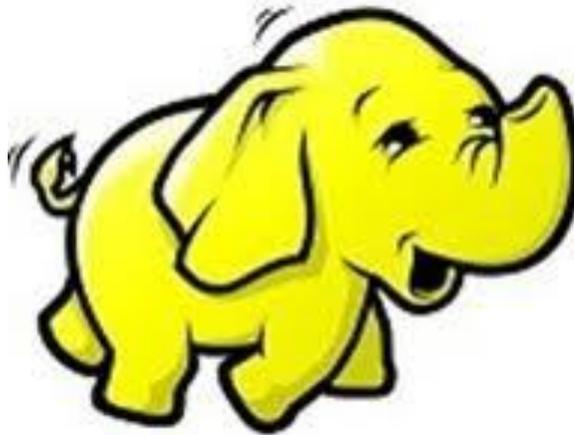
Who this course is for and not for?

For: Typically professionals with basic knowledge of software development, programming languages, and databases will find this course really helpful. Basic knowledge should be enough to succeed at this course

Not For: Students who are absolute beginners at software development as a discipline will find it difficult to follow the course

**www.learninghub.co.in; Ph: 9325793756; Email:learninghub01@gmail.com
2nd Floor, Above HDFC Bank, Next to Noble Polyclinic, MAGARPATTA CITY, PUNE**

HADOOP SCHOOL OF TRAINING



Phase 1: Hadoop
Fundamentals
(20 Hours)

Getting the Basics
Rights

Big Data

- What is Big Data
- Dimensions of Big Data
- Big Data in Advertising
- Big Data in Banking
- Big Data in Telecom
- Big Data in eCommerce
- Big Data in Healthcare
- Big Data in Defence
- Processing options of Big Data
- Hadoop as an option

Hadoop

- What is Hadoop
- How Hadoop Works
- HDFS
- Mapreduce
- How Hadoop has an edge

Hadoop Ecosystem

- Sqoop
- Oozie
- Pig
- Hive
- Flume

Hadoop Hands On

- Setting up Hadoop on a Single node cluster
- Running HDFS commands
- Running your Mapreduce program
- Running Sqoop Import and Sqoop Export
- Creating Hive tables directly from Sqoop
- Creating Hive tables
- Querying Hive tables

- Running an Oozie workflow
- Analyzing twitter data using Flume

Multinode Setup

- Setting up Multinode setup on Amazon ec2
- Setting up multimode setup on the classroom machines
- Setting up Cloudera Manager on the cloud
- Setting up Cloudera Manager on local setup

Cluster Capacity Planning

Level 1: Mini Project

Level 1: Evaluation Test
(50 marks)

HADOOP SCHOOL OF TRAINING



Phase 2: Data Analyst
(20 Hours)

Becoming a pro data analyst

Pig

- Introduction
- Basic Data Analysis
- Complex Data Analysis
- Multi Data Set Analysis
- UDFs in Pig
- Troubleshooting and Optimizing Pig
- Pig Hands On

Hive

- Introduction
- Basic Data Analysis with Hive
- Hive Data Management
- Text Processing with Hive
- Transformations in Hive
- Optimizing Hive
- Hive Hands On

Data Analysis Using Pentaho as a ETL tool

- Introduction
- Setting up Pentaho
- Loading Data to HDFS
- Loading Data to Hive
- Aggregation through Mapreduce
- Transforming Data with Hive
- Transforming Data with Pig
- Loading data from HDFS to RDBMS
- Loading Data from hive to RDBMS
- Reporting on HDFS Data
- Reporting on Hive Data

Impala

- Introduction
 - Data Analysis using Impala
- Mini Project using Pig and Hive**
Evaluation Test (50 marks)

HADOOP SCHOOL OF TRAINING

Trainer Profile

Experienced

8+ yrs of Enterprise Software Dev Exp.

Certified

Hadoop, Hbase and MapR certified

Customers

Served customers like Accenture, HP, Genpact, Mastek, and Cisco

About the trainer



Trainer's Certifications

cloudera



CCAH,CCHD, CCHSB MapR M5

Zend

SCJP

SCWCD

www.learninghub.co.in; Ph: 9325793756; Email:learninghub01@gmail.com
2nd Floor, Above HDFC Bank, Next to Noble Polyclinic, MAGARPATTA CITY, PUNE