

Course Name	IBM CE – Big Data Analytics
Course Duration	32 Hours
About Technology	<p>IBM InfoSphere BigInsights brings the power of Hadoop to the enterprise. Apache Hadoop is the open source software framework, used to reliably managing large volumes of structured and unstructured data.</p> <p>BigInsights makes it simpler for people to use Hadoop and build big data applications. It enhances this open source technology to withstand the demands of your enterprise, adding administrative, discovery, development, provisioning, and security features, along with best-in-class analytical capabilities from IBM Research. The result is that you get a more developer and user-friendly solution for complex, large scale analytics.</p> <p>InfoSphere BigInsights allows enterprises of all sizes to cost effectively manage and analyze the massive volume, variety and velocity of data that consumers and businesses create every day. InfoSphere BigInsights can help you increase operational efficiency by augmenting your data warehouse environment. It can be used as a query-able archive, allowing you to store and analyze large volumes of multi-structured data without straining the data warehouse. It can be used as a pre-processing hub, helping you to explore your data, determine what is the most valuable, and extract that data cost-effectively. It can also allow for ad hoc analysis, giving you the ability to perform analysis on all of your data.</p>
About Course	This course, developed specifically for the business analyst will teach how to capture structured, semi-structured and unstructured data from several different data source types using IBM InfoSphere BigInsights and then do manipulations and analysis on the gathered data. This course will focus on using the Graphical User Interface of InfoSphere BigInsights to collect, manipulate, analyze, view and export data.
Target Audience	<p>System administrators and developers.</p> <p>Students of:</p> <ul style="list-style-type: none"> – CS/IT/ ECE/EEE – 2nd / 3rd / 4th Year – MCA – 2nd/ 3rd Year – MTech – 1st Year – BCA/ BSC – 3rd year – MBA
Pre-requisites	Prior to attending this course, knowledge of Linux would be beneficial.
Contents	<p>After completing this course, you should be able to:</p> <p>Module 1 - Fundamentals of Big Data using Hadoop</p> <p>Unit 1 - Introduction to Big Data</p> <p>Unit 2 - Introduction to InfoSphere BigInsights</p> <p>Unit 3 - Apache Hadoop and HDFS Overview</p> <p>Exercise 1- Exploring Apache Hadoop</p> <p>Unit 4 - GPFS-FPO</p> <p>Unit 5 - BigInsights Web Console Security</p>

Unit 6 - Introduction to MapReduce
Exercise 2- MapReduce
Unit 7 - Adaptive MapReduce
Unit 8 - Setup, Configuration, and Administration of a Hadoop Cluster
Exercise 3- Hadoop Configuration
Unit 9 - Overview of Oozie
Exercise 4- Controlling Workloads with Oozie
Exercise 4- Controlling Workloads with Oozie continued
Unit 10 - Managing Job Execution
Unit 11 - Moving Data into Hadoop
Exercise 5 - Using Flume for Data Loading

Module 2 - BigInsights Analytics for Business Analysts

Unit 1 - Introduction to BigSheets
Unit 2 - Making Data Available to BigSheets
Exercise 1 - Importing Data into a Workbook
Unit 3 - Navigating a Workbook
Exercise 2 - Adding Sheets to a Workbook
Unit 4 - Expressions and Functions
Exercise 3 - Working with Functions
Unit 5 - Big SQL Integration
Exercise 4 - Big SQL Integration
Unit 6 - BigSheets Visualizations
Exercise 5 - Analyzing Social Media and Structured Data