

Big Data - Hadoop

Product Code: ST-SD-30009

Duration: 100 hrs

Prerequisite:

- Basics of programming (Java knowledge will be added advantage)
- Basics of SQL

Take Aways:

- Understand what is data
- Understand difference between data and Information
- What businesses do with Information
- Understanding solution for Big Data: Hadoop
- Understand how HDFS and other distributed file systems are different from traditional file systems
- Process big data with help of frameworks like MapReduce
- Process data using PIG
- Process data using HIVE
- Schedule processing data using Oozie
- Use of NoSQL database: HBase
- Load data in HDFS using SQOOP, Flume
- Use of basic reporting tool

Course Contents:

- BIG DATA
- HADOOP
- HDFS
- PIG
- Hive
- MapReduce
- Oozie
- Sqoop
- Flume
- Zookeeper & HBase (NoSQL)
- Introduction to BI with BIRT

- Spark
- Project

Recommended Next Course:

- Data Analyst
- ETL DeveloperData Structures

SEED Infotech Ltd.