

# ANSH INFOTECH

Development | Training | Consultancy  
An ISO Certified company 9001:2015

Course Name – **BigData**

Category - **Advance**

Duration - **2 Months**

**Pre-Requisites:** Python, basic knowledge of DMBS like My-SQL, Familiarization with any Object oriented programming language

## COURSE CONTENT

- Python
- Welcome to Big Data
  - Why is Big Data
  - 7V's of Big Data
  - Installing Hadoop
  - Hadoop Overview
  - Hadoop Ecosystem
- Basic Linux Commands
- Hadoop's Core
  - HDFS
  - MapReduce
  - Programming with MapReduce
- Pig
  - Introduction to Ambrai
  - Introduction to Pig
- Spark
  - Why Spark
  - Resilient Distributed Dataset
- Relational Data in Hadoop
  - Introduction to Hive
  - Working of Hive
  - Sql with Hadoop(Sqoop)
- Non-relational data in Hadoop
  - Introduction to NoSql
  - Why NoSql?
  - HBase
  - Cassandra
  - MongoDB
- Interactive Query Engines
  - Apache Drill
  - Apache phoenix
  - Presto
- Cluster Management
  - Yarn
  - Tez

# ANSH INFOTECH

Development | Training | Consultancy

An ISO Certified company 9001:2015

- Mesos
- Zookeeper
- Oozie
- Zepplin
- Hue
- Analysis of Data
  - Kafka
  - Flume
- Real Life Application(Project)

