Big Data & Analytics with IBM InfoSphereBigInsights(Bigdata-Hadoop)

INTRODUCTION TO BIGDATA AND HADOOP 4 hrs

EXAMINING BIG DATA TYPES

- Defining Structured Data, Defining Unstructured Data
- Looking at Real-Time and Non-Real-Time Requirements
- Putting Big Data Together

DISTRIBUTED COMPUTING

- A Brief History of Distributed Computing.
- Understanding the Basics of Distributed Computing.
- Analytical Data Warehouses.
- Integrating Big Data with the Traditional Data Warehouse
- Big Data Analysis and the Data Warehouse
- Changing the Role of the Data Warehouse
- Examining the Future of Data Warehouse
- Big Data Analytics, Big Data Applications

Introduction to RDBMS.

- Introduction to DB2.
- DB2 Instance
- DB2 Database and Tables.
- DB2 Hands on Lab.

Hands On Lab

HADOOP DISTRIBUTED FILE SYSTEM. 4hrs

- HDFS Architecture.
- Understanding the Hadoop Distributed File System (HDFS).
- Types of Node in Hadoop Cluster.
- Name Node,Job Tracker,Secondary Name Node.
- DataNode,Task Tracker,Child Java Virtual Machine.
- Data Replication.
- HadoopFileSystemcommands .
- Working on HDFS File System.
- **HDFS Hands on Lab.**

MAPREDUCE USING INFOSPHERE BIGINSIGHTS4+4hrs

- Introduction to Map Reduce.
- Understanding Map Function
- Understanding Reduce Function
- Fault tolerance and Scheduling.
- Submitting a MapReducejob.
- Putting map and reduce Together
- Map Reduce using Java Programming Language.
- Working on JAR files.
- MapReduce Job using web console.
- Future of Map Reduce.
- Temperature data Analysis using Map Reduce.

WORKING WITH JAQL LANGUAGE 4hrs

- Introduction of JAQL approach.
- Introduction to ETL approach.
- Start the Big Insights Component.
- Setup to run JAQL.
- Some JAQL Basics.
- WHY JSON .
- Working on JSON and CSV File.
- Information stream ,Information ocean.
- Understanding Unstructured data.
- Streaming Data Analytics Through JAQL
- JAQL using Eclipse IDE.
- **JAQL Hands on Lab**

ACCESSINGHADOOP DATA USINGHIVE.4 hrs

- Understand Hive Environment.
- Introduction to Hive.
- Hive Web Interface.
- Exploring the Hive Beeline Command Line Interface (CLI).

- Perform DDL approach Through Hive.
- Working with Hive Query Language .
- Working with Databases in Hive.
- Tables in Hive.
- Managed Partitioned Tables.
- HIVE External Table.
- HIVE DML in action.
- Viewing Data from BigInsights Web Console.
- Loading Data into the Managed Non-Partitioned Tables.
- OPERATORS
- Relational Operators ,ArithmeticOperators ,Logical Operators.
- **HIVE Hands on Lab.**

INTRODUCTION TO SQOOP AND OOZIE **4 hrs**

- introduction of Sqoop
- Requirment of sqoop
- advantages of sqoop
- Working Lifecycle of oozie
- understand oozie data flow
- oozie setup and requirment
- understand oozie scheduling

PIG USING INFOSPHERE BIGINSIGHTS.4hrs

- Introduction to PIG.
- Pig Advantages and Disadvantages**.**
- Execute Pig statements from the Grunt shell.
- Execute a Pig script.
- Pass parameters to a Pig script .
- Load data for using within Pig.
- Pig Relational Operators.
- Pig Evaluation Functions.
- **PIGHands on Lab.**

BIG SQL IN DETAIL **4+4hrs**

- The Big SQL Command Line Interface.(JSqsh).
- Setting up environment of BIG SQL.
- Understanding JSqsh Connections.
- Creating your own database connection
- Working With BIG SQL Drivers.
- Setting up Connection with BigSql
- Working with BIGSql and Big Data
- Perform Various BigSql Queries

Hands on Lab and Project Development.