KNOWLEDGE
For Training

1. Course Details:

Course Code : DSBA-003
Course Title : AWS Big Data
Course Duration : 5 Days

2. Course Overview

- Big Data on AWS introduces you to cloud-based big data solutions such as Amazon EMR, Amazon Redshift, Amazon Kinesis and the rest of the AWS big data platform. In this course, we show you how to use Amazon to process data using the broad ecosystem of Hadoop tools like Hive and Hue.

3. Learning Outcomes:

- Fit AWS solutions inside of a big data ecosystem
- Identify the components of an Amazon EMR cluster
- Launch and configure an Amazon EMR cluster
- Leverage common programming frameworks available for Amazon EMR including Hive, Pig, and Streaming
- Use in-memory analytics with Spark on Amazon EMR
- Choose appropriate AWS data storage options
- Leverage Amazon Redshift to efficiently store and analyze data
- Comprehend and manage costs and security for a big data solution
- Secure a Big Data solution
- Identify options for ingesting, transferring, and compressing data
- Orchestrate big data workflows using AWS Data Pipeline

4. Course Contents:

- Overview of Big Data
- Big Data Ingestion and Transfer
- Lab 1: Using Amazon Kinesis to Stream and Analyze Apache Server Log Data
- Big Data Storage Solutions
- Big Data Processing and Analytics
- Lab 2: Using Amazon Athena to Query Log Data from Amazon S3
- Apache Hadoop and Amazon EMR
- Lab 3: Storing and Querying Data on Amazon DynamoDB
- Using Amazon EMR
- Hadoop Programming Frameworks

- Lab 4: Processing Server Logs with Hive on Amazon EMR
- Web Interfaces on Amazon EMR
- Lab 5: Running Pig Scripts in Hue on Amazon EMR
- Apache Spark on Amazon EMR
- Lab 6: Processing NY Taxi data using Spark on Amazon EMR
- Amazon Redshift and Big Data
- Visualizing and Orchestrating Big Data
- Lab 7: Using TIBCO Spotfire to Visualize Data
- Managing Big Data Costs
- Big Data Design Patterns

5. Our Methodology:

- Make coaching and monitoring innovative and using modern
- Media training also using on the go training by using interactive means and focusing on
- The exercises, practical applications and real situations study.
- Live delivery method, instructor-led training.
- Experienced consultant, trainers, and professional.
- Qualified trainer with high-level experience

6. Attendance Reports:

- Send daily attendance reports to trainingdep
- Send full attendance report to training dep. by the end of thecourse
- Attend 100 % from the course days also providedaily
- Issue attendance certificate for participant who attend minimum 80% from the course duration

7. Pre/Post Reports:

- Pre- assessment before starting training.
- Post assessment after finish training
- Full report for the deferent between Pre-& Post assessment

8. Who Should Attend?

- Business Analysts
- Financial Analysts
- Managers, Heads of Departments, and Directors

- Data Center
- Individuals responsible for designing and implementing big data solutions,
- namely Solutions Architects
- Data Scientists and Data Analysts interested in learning about the services and architecture patterns behind big data solutions on AWS