



Exploring Big Data & Data Analysis in the Big Data Ecosystem

Gain the Skills Required to Use Integrated Big Data Solutions to Acquire, Process, Integrate and Analyze Big Data

www.triveratech.com

Course Snapshot

- **Course: Exploring Big Data & Data Analysis in Big Data Ecosystem (TTDS6615)**
- **Duration:** 4 days
- **Skill-level:** Foundation-level Big Data knowledge for Intermediate skilled team members. This is not a basic class.
- **Audience:** This course is geared for Administrator, Database Administrator, Developer those who wants to use Integrate Big Data Solution to acquire, process, integrate and analyze big data
- **Hands-on Learning:** This course is approximately 50% hands-on lab to 50% lecture ratio, combining engaging lecture, demos, group activities and discussions with machine-based student labs and exercises. Student machines are required.
- **Delivery Options:** This course is available for **onsite private classroom presentation, live online virtual presentation**, or can be presented in a **flexible blended learning format** for combined onsite and remote attendees. Please also ask about our **Self-Paced / Video** or **QuickSkills / Short Course** options.
- **Customizable:** This course agenda, topics and labs can be further adjusted to target your specific training skills objectives, tools and learning goals. Please inquire for details.

Overview

In the **Exploring Big Data** course, you will learn to use Integrated Big Data Solution to acquire, process, integrate and analyze big data. In this course, you will be introduced to Big Data Cloud Service. Increase your Big Data technology portfolio by learning to use a wide range of big data acquisition, processing, integration, and analysis techniques. You will also explore engineered systems for Big Data, which provide a variety of data integration and analysis capabilities. Analysis options include Big Data SQL, Data Mining, R Enterprise, and Big Data Discovery.

Learning Objectives

This course is approximately **50% hands-on**, combining expert lecture, real-world demonstrations and group discussions with machine-based practical labs and exercises. Our engaging instructors and mentors are highly experienced practitioners who bring years of current "on-the-job" experience into every classroom.

Working in a hands-on learning environment, led by our Big Data expert instructor, students will learn to:

- Define Big Data and Identify Big Data Use Cases
- Review Big Data Management Architecture and Engineered Systems
- Describe Integrated Big Data Solution and its components.
- Examine MapReduce programs and balance MapReduce jobs
- Use NoSQL Database
- Use XQuery for Hadoop
- Install, use, and administer the Big Data Appliance
- Provide data security and enable resource management
- Examine MapReduce programs and balance MapReduce jobs
- Use the BigDataLite Virtual Machine
- Use the Hadoop Distributed File System (HDFS) to store, distribute, and replicate data across the nodes in the Hadoop cluster.
- Acquire big data using the HDFS Command Line Interface, Flume, and NoSQL Database.
- Use MapReduce and YARN for distributed processing of the data stored in the Hadoop cluster.
- Process big data using MapReduce, YARN, Hive, Pig, XQuery for Hadoop, Solr, and Spark.
- Integrate big data and warehouse data using Sqoop, Big Data Connectors, Copy to BDA, Big Data SQL, Data Integrator, and GoldenGate.
- Analyze big data using Big Data SQL, Advanced Analytics technologies, and Big Data Discovery.
- Use and manage Big Data Appliance.
- Secure your data.

- Understand Big Data Cloud Service: Key Features & Benefits

Need different skills or topics? If your team requires different topics or tools, additional skills or custom approach, this course may be further adjusted to accommodate. We offer additional Big Data / Data Science, Hadoop, programming, analytics, Python/R, and other related topics that may be blended with this course for a track that best suits your needs. Our team will collaborate with you to understand your needs and will target the course to focus on your specific learning objectives and goals.

Audience & Pre-Requisites

This course is geared for attendees who wish to use Integrated Big Data Solution to acquire, process, integrate and analyze big data. Attendees should possess the following incoming skills:

- Basic to Intermediate IT Skills, and Big Data knowledge
- Good foundational mathematics or logic skills
- Basic Linux skills, including familiarity with command-line options such as ls, cd, cp, and su

Next-Steps / Follow on Training: We offer a wide variety of Big Data, Analytics, Data Science, Hadoop, AI / Machine Learning, Programming and other related courses that can advance your skills to the next level after this course. Please see our **Big Data Training Suite** course list for details, or please inquire for next step recommendations for follow on courses or Learning Plan options based on your role and goal.

Enhanced Learning Services: Please also ask about our **Pre-Training Class OnRamp & Prep / Primer offerings, Skills Gap Assessment Services, Case Studies, Knowledge Check Quizzes, Skills Immersion Programs & Camps, Collaborative Mentoring & Coaching Services** and **Extended Learning Support & Post Training** services.

Course Topics / Agenda

Please note that this list of topics is based on our standard course offering, evolved from typical industry uses and trends. We will work with you to tune this course and level of coverage to target the skills you need most. Course topics, agenda and labs are subject to change, and may adjust during live delivery based on audience needs, participation and skill-level.

- | | | |
|---|---|---|
| • Big Data and the Information Management System | • Database | • Using Advanced Analytics |
| • Using Big Data Lite Virtual Machine | • Introduction to MapReduce | • Introducing Big Data Discovery |
| • Introduction to the Big Data Ecosystem | • Using YARN to Manage Resources | • Using the Big Data Appliance (BDA) |
| • Introduction to the Hadoop Distributed File System (HDFS) | • Overview of Apache Hive and Apache Pig | • Managing the Big Data Appliance |
| • Acquire Data using CLI, Fuse-DFS, and Flume | • Overview of Cloudera Impala, Solr, and Apache Spark | • Balancing MapReduce Jobs |
| • Using and Administering NoSQL | • Using XQuery for Hadoop | • Securing Your Data on the BDA |
| | • Options for Integrating Your Big Data | • Introduction to Big Data Cloud Service (BDCS) |
| | • Using Big Data SQL | |

Course Materials: Each student will receive a **Student Guide** with course notes, code samples, software tutorials, step-by-step written lab instructions, diagrams and related reference materials and links (as applicable). Students will also receive the project files (or code, if applicable) and solutions required for the hands-on work.

Lab Setup Made Simple. All course labs and solutions, data sets, Tableau course software (limited version, for course use only), detailed courseware, lab guides and resources (as applicable) are provided for attendees in our easy access, no installation required, remote lab environment for the duration of the course. Our tech team will help set up, test and verify lab access for each attendee prior to the course start date, ensuring a smooth start to class and successful hands-on course experience for all participants.

For More Information

Need dedicated training? All courses can be presented **onsite** or **online**, or in a **combined / flex / blended learning format**, tailored to target your specific audience, needs and learning goals. We also offer focused, flexible **short courses**, **self-paced learning** options, **recorded sessions** and more. We train beginner to advanced skills in all areas we cover, and offer **New Hire / Cohort Training**, **Boot Camps**, **Skills Immersion Programs**, **Reskilling Programs**, **Skills Migration & Transition Programs**, and more. We collaborate with you to ensure all courses are truly targeted to meet your specific needs and learning skills, maximizing your valuable training time, as well as your important budget.

Please also visit our extensive **Public Training Schedule** for training for smaller groups or individuals. Please contact us for course details, **Corporate Rates** and **Special Discount Offers**.

For more information about our dedicated training services, collaborative mentoring services, courseware licensing options, courseware development services, public course schedule, training management services, partner and reseller programs, or to see our complete list of course offerings and special offers please visit us at www.triveratech.com, email Info@triveratech.com or call us toll free at **844-475-4559**. Our pricing and services are always satisfaction guaranteed.

TRIVERA TECHNOLOGIES • Collaborative IT Training, Coaching & Courseware Solutions
www.triveratech.com • toll free +1-844-475-4559 • Info@triveratech.com • Twitter TriveraTech

ONSITE, ONLINE & BLENDED TRAINING SOLUTIONS | PUBLIC / OPEN ENROLLMENT COURSES | COURSEWARE LICENSING & DEVELOPMENT
MENTORING | ASSESSMENTS | LEARNING PLAN DEVELOPMENT | SKILLS IMMERSION PROGRAMS / RESKILLING / NEW HIRE / BOOT CAMPS
PARTNER & RESELLER PROGRAMS | CORPORATE TRAINING MANAGEMENT | VENDOR MANAGEMENT SERVICES

Trivera Technologies is a Woman-Owned Small-Business Firm

Explore Trivera's Ways to Learn...

