# BOOST
## THE LEADING TRAINING PROVIDER



Digital Transformation and Innovation

# Big Data Principles and Practices

# Course Introduction

The concept of big data has been around for years; most organizations now understand that if they capture all the data that streams into their businesses, they can apply analytics and get significant value from it. Big data analytics helps organizations harness their data and use it to identify new opportunities. That, in turn, leads to smarter business moves, more efficient operations, higher profits and happier customers.

This training course is designed to provide participants with the key concepts and knowledge of big data – the landscape, the technology behind it, business drivers, and strategic possibilities. "Big data" is a hot buzzword, but most organizations struggle to put it to practical use. Without assuming any prior knowledge of Apache Hadoop or big data management, this course teaches you how to use and manage the benefits of big data.

# Target Audience

- Business and technology leaders
- Business Unit Managers
- Business Development Consultants
- General Managers / Regional Managers
- Senior and mid-level leaders
- individual leaders of all levels in the organization
- Art Director
- Marketing Consultants
- Marketing Development Manager

# Learning Objectives

- Gain a thorough understanding of Big Data technologies, their benefits, and the value that it can deliver to industries, companies, and functions
- Analyze, process, and extract information from extremely complex and large data sets.
- Build an organization-wide Big Data program
- Develop the maturity of Big Data within their organization
- Apply a variety of use cases to drive ideation
- Act lean and agile in pursuit of Big Data objectives

# Course Outline

- **Day 01**

  **The concepts**

  - Load data how you find it
  - Process it when you can
  - Project it into various schemas on the fly
  - Push it back to where you need it

  **The basics**

  - What it's good for
  - What can't it do / disadvantages
  - Most common use cases for big data

  **Value Creation with Big Data**

  - Introduction to Big Data technologies
  - Trends in Big Data
  - Big Data applications, use cases and best practices across industries and functions
  - Data sourcing strategies and challenges
  - Ideation phase: creating first successes

    ◦ From ideation to Proof-of-Concept and minimum viable product
- **Day 02**

    **Managing Big Data Transformation**

    ◦ Big Data Maturity model
    ◦ Developing a Big Data roadmap
    ◦ What does good look like: determining your Big Data end game
    ◦ Orchestrating Big Data maturity across data, technology and people

    **Hadoop – the free platform for working with big data**

    ◦ History
    ◦ Yahoo
    ◦ Platform fragmentation
    ◦ What usage looks like in the enterprise

    **Introduction to HDFS**

    ◦ Robustness
    ◦ Data Replication
    ◦ Gotchas
- **Day 03**

    **MapReduce – the core big data function**

    ◦ Map explained
    ◦ Sort and shuffle explained
    ◦ Reduce explained

    **YARN**

    ◦ How it fits
    ◦ How it works
    ◦ Resource Manager
    ◦ Application Master

    **PIG**

    ◦ What it is
    ◦ How it works
    ◦ Compatibilities
    ◦ Advantages
    ◦ Disadvantages
- **Day 04**

**Processing Data**

- ◦ The Piggy Bank
- ◦ Loading and Illustrating the data
- ◦ Writing a Query
- ◦ Storing the Result

**HIVE**

- ◦ Data warehousing
- ◦ What it is, what it's not
- ◦ Language compatibilities
- ◦ Advantages

**OOZIE**

- ◦ What it is
- ◦ Complex workflow environments
- ◦ Reducing time-to-market
- ◦ Frequency execution
- ◦ How it works with other big data tools

• **Day 05**

**FLUME –** stream, collect, store and analyze high-volume log data

- ◦ How it works: Event, source, sink, channel, agent, and client
- ◦ How it works illustrated
- ◦ How it works demonstrated

**SPARK**

- ◦ Move over 2012 Big Data tools: Apache SPARK is the new power tool
- ◦ The new open-source cluster framework
- ◦ When SPARK performs 100 times faster
- ◦ Performance comparison of Spark and Hadoop
- ◦ What else can it do?

**Big Data Leadership**

- ◦ Lean/agile working in support of Big Data transformation
- ◦ Key success factors for adoption of Big Data at speed
- ◦ Required skills & competencies for successful digital transformation
- ◦ Understand the mindset of digital disruptors

# Confirmed Sessions

| FROM | TO | DURATION | FEES | LOCATION |
| --- | --- | --- | --- | --- |
| May 18, 2025 | May 22, 2025 | 5 days | 4250.00 $ | KSA - Riyadh |
| Aug. 11, 2025 | Aug. 15, 2025 | 5 days | 4250.00 $ | UAE - Dubai |
| Nov. 17, 2025 | Nov. 21, 2025 | 5 days | 4950.00 $ | Spain - Barcelona |