

Big Data Fundamentals

Learn via: **Classroom**

Duration: **1 Day**

Overview

Big data is a hot topic in BI and analytics. Yet it is a complex topic that is still in the early stages of evolution. Successful big data projects that deliver real business value are challenged by multiple definitions and rapidly shifting technologies. Achieving good return on your big data investment requires strategy that focuses on purpose, people, and process before exploring data and technologies. Strategy drives planning and architecture to ensure that big data complements and does not disrupt the existing BI and analytics environment. To prepare for success with big data, start by understanding all of the pieces and how they fit together.

Prerequisites

There are no prerequisites for this course.

Who Should Attend

Business and data analysts; BI and analytics program and project managers; BI and data warehouse architects, designers, and developers; data governance and data quality professionals getting started with big data; anyone seeking to cut through the hype to understand the opportunities, challenges, and realities of the big data phenomenon.

What You Will Learn

- Common definitions of big data and the implications of each
- Key characteristics of big data and why size is not among the top five
- The structures that can be found in “unstructured” data
- Types of big data sources—streaming data, social data, sensor data, etc.
- Value opportunities and common applications for big data
- Considerations when adapting architectures, organizations, and cultures to incorporate big data
- The scope of big data processes, tools, and technologies

Outline

Module 1 – Big Data Basics

- What Is Big Data?
 - Definitions
 - Characteristics (3 V's plus 2)
 - Types of Big Data
- Why Big Data Analytics – Extending Advanced Analytics Capabilities
- Big Data Use Cases
 - Customer Understanding and Targeting
 - Business Process Optimization
 - Healthcare Advances
 - Law Enforcement and Public Safety
 - Sports Performance Improvement
 - Public Transportation and Infrastructure Advances
- Why Big Data Now? – The Driving Forces
- Kinds of Big Data – Data Variety
- Sources of Big Data
 - Web and Social Media
 - Machine to Machine
 - Other Sources (Big Transaction Data, Biometrics, Human Generated Data,

- Publicly Available Data, Legacy Documents)
- Working with Big Data – The Big Picture

Module 2 – Big Data Processes

- Business Case
 - Business Needs and Opportunities
 - Areas of Insight
 - Expected Outcomes
 - Business Value Projection
- Technical Case – Big Data Rationale
- Data Sourcing – Getting Big Data
- Data Preparation and Storage
 - Data Selection
 - Data Cleansing
 - Data Integration
 - Data Reduction
- Big Data Analytics
 - Problem Framing
 - Analytic Purpose
 - Analytic Modeling
 - Data Visualization
 - Consumption and Application

Module 3 – Big Data Architecture

- The Role of Architecture
 - What and why
- Data Architecture
- Data Storage
- Data Access
- Data Analysis
- Data Consumption
 - Process Architecture
 - Data Governance Processes
- Data Integration and Quality
 - Analytics Architecture
 - Machine Learning
 - Predictive Analytics
 - Prescriptive Analytics
 - Descriptive Analytics
 - Reporting
- Technology Architecture
 - Search and Visualization
 - Data Management and Data Access
 - Hadoop and NoSQL
- Big Data ... Big Architecture – Summary

Module 4 – Big Data Technology

- The Technology Landscape – Overview
- Infrastructure
 - Databases
 - Development and Deployment Environment
- Analytics – Data Analysis
- Data Sources – Big Data Providers
- The Core Technologies
 - MapReduce
 - Hadoop

Module 5 – Getting Started with Big Data

- Readiness Assessment – Check Your Position
- Planning and Preparation – Charting the Course
- Execution – Navigating the Course
- Post-Project Activities – At the Destination

- Best Practices – Lessons Learned
- Mistakes to Avoid – More Lessons Learned

Module 6 – Summary and Conclusion

- Summary of Key Points – A Quick Review
- References and Resources – To Learn More