Fundamentals of Big Data Training

Duration: **3 Day**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Overview

The **Fundamentals of Big Data** course is a three-day introductory programme for professionals who want to understand Big Data technologies, Data Science concepts, and the data models that underpin modern analytical platforms.

The course is particularly suitable for organisations and technical professionals evaluating which **NoSQL technologies and Big Data solutions** may be appropriate for their environment. It also provides a strong first step for learners beginning a broader professional journey into Data Science.

Throughout the programme, participants explore different ways of modelling data, the core differences between SQL and NoSQL systems, and several of the major technologies used within the modern Big Data ecosystem.

Platforms including Neo4j, Hadoop, Spark, and MongoDB are introduced, with practical discussion of how they store, query, process, and analyse data.

The course also introduces fundamental Data Science concepts together with the mathematical and statistical thinking that supports analytical work. Participants learn how Python can be used to work with different data sources, interact with NoSQL databases through APIs, and support Big Data analysis.

By the end of the course, learners will understand the main technologies within the Big Data landscape, recognise different data modelling approaches, and have a clearer understanding of how Hadoop and Spark can be applied to large-scale analytical problems.

Prerequisites

Participants should have:

- Previous experience using Python
- Experience working with data

Familiarity with the following areas will also be beneficial:

- Basic Python programming
- Tabular data
- Data analysis
- Fundamental database concepts
- Basic SQL querying concepts

Fundamentals of Big Data Training

Duration: **3 Day**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Who Should Attend

This course is particularly suitable for:

- Existing Python developers
- Data Analysts
- Technical professionals moving into Big Data
- Learners beginning a Data Science development path
- Technical teams evaluating NoSQL technologies
- Developers working with data platforms
- Analysts who need to work with larger and more varied datasets

What You Will Learn

By the end of the course, participants will understand:

- The fundamentals of Data Science
- The role of a Data Scientist
- Different data models
- The differences between SQL and NoSQL
- What a NoSQL database is
- Major categories of NoSQL databases
- The role of Hadoop within Big Data
- The fundamentals of Spark
- Graph database concepts
- Document-oriented databases

Participants will also be able to:

- Use common Data Science terminology with greater confidence
- Work with Python's native data structures and storage approaches
- Query NoSQL databases through Python APIs
- Perform foundational Big Data processing using Hadoop
- Analyse large datasets using Spark
- Evaluate technology options for different data problems
- Approach analytics problems from a Data Science perspective

Outline

Introduction to Data Science

This module introduces the fundamental concepts behind Data Science and its role within modern organisations.

Topics include:

- What is Data Science?
- The role of a Data Scientist

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

- Roles within Data Science teams
- Technical and analytical skills
- Data Science methods
- Asking useful analytical questions
- Answering insight questions
- Applying Data Science to business problems
- Science in Business
- Data-driven decision-making

Participants learn that effective Data Science is not simply about tools; it is also about framing the right questions, understanding available data, and creating useful insights.

Python for Data Science

This section focuses on Python as a tool for data processing and analytics.

Topics include:

- Using Python for Data Science
- Python data structures
- Data manipulation
- Reading and writing data
- Dataset processing with Python
- Python-based analytics
- Connecting to NoSQL systems
- Working with APIs

Participants develop their ability to use Python as a common interface across different data technologies.

Data Analytics and Statistical Inference

This section introduces the statistical thinking used within data analysis.

Topics include:

- Data Analytics fundamentals
- Descriptive analysis
- Statistical inference
- Understanding data distributions
- Evaluating relationships within data
- Producing insights
- Interpreting analytical results
- Translating business questions into analytical problems

The aim is to help learners see Big Data technologies as part of a broader analytical process rather than as isolated infrastructure tools.

Data Models

This module introduces different approaches to storing and representing data.

Topics include:

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- Structured data
- Semi-structured data
- Unstructured data
- Relational data models
- Document models
- Graph models
- Key-value approaches
- Advantages of different data models
- Selecting a model for a specific use case

SQL and NoSQL

Participants compare relational and non-relational database approaches.

Topics include:

- What is a SQL database?
- Relational database structures
- What is a NoSQL database?
- SQL versus NoSQL
- Schema approaches
- Scalability
- Distributed data
- NoSQL use cases
- Categories of NoSQL databases

Big Data with Neo4j

This section introduces graph databases using Neo4j.

Topics include:

- Graph databases
- Nodes
- Relationships
- Properties
- Graph data modelling
- Connected data
- Neo4j fundamentals
- Graph querying
- Graph database use cases

Participants examine why graph databases can be useful when relationships between data points are especially important.

Big Data with Hadoop

This module introduces the fundamentals of the Hadoop ecosystem.

Topics include:

- What is Hadoop?
- Distributed computing

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- Hadoop architecture
- Distributed storage
- Storing large datasets
- Parallel data processing
- Hadoop use cases
- Python and Hadoop scenarios

Participants learn how large datasets can be distributed and processed across multiple systems.

Big Data with Spark

This section introduces Apache Spark and its role in scalable analytics.

Topics include:

- What is Spark?
- Spark architecture
- Distributed processing
- In-memory processing
- Analysing large datasets
- Data transformation with Spark
- Python and Spark
- Analytical workloads
- Key differences between Hadoop and Spark

Participants examine how Spark can support fast and scalable analytical processing.

Big Data with Mongo

This section introduces document-oriented NoSQL databases through MongoDB.

Topics include:

- Document databases
- MongoDB fundamentals
- Documents
- Collections
- Flexible schemas
- CRUD operations
- Using MongoDB through Python APIs
- MongoDB queries
- Document database use cases

Participants learn how structured and semi-structured data can be represented using document-based models.

Analytics Case Study

The final section brings together the concepts and technologies covered during the course.

In the case study, participants:

- Define a business problem

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- Evaluate suitable data models
- Examine available data sources
- Select a SQL or NoSQL approach
- Choose an appropriate Big Data technology
- Process data with Python
- Evaluate analytical outcomes

This exercise helps learners understand how different Big Data technologies can be combined within a realistic analytical scenario.