informatech

DATA MANAGEMENT AND BUSINESS INTELLIGENCE | DMBI-011

Certificate in Advanced Big Data and Data Analytics (CABDDA)

UK
+44 33 000 111 90
info@informatech.co.uk
https://informatech.uk
63-66 Hatton Garden Hatton Garden
EC1N 8LE , London

NL
+31 85 74 444 46
info@infomatech.nl
https://infomatech.nl
Waarderweg 50 - 2031PB
Haarlem - Netherlands

Course content

Why Attend

Big data is a change agent that challenges the ways in which organizational leaders have traditionally made decisions. This course provides participants with the confidence to store, process, analyze and present big data use cases within their organizations. This course provides a multitude of hands-on labs with Spark, a key big data technology used to solve data intensive problems. Participants will gain the knowledge and skills they need to assemble and manage a large-scale big data analytics project. Lastly, participants will work through advanced machine learning and deep learning use cases.

This is our most advanced course in our big data series following Certified Big Data and Data Analytics Practitioner (CBDDAP) and Certificate in Big Data Fundamentals (CBDF). Participants will aim to identify areas within their organization that can be improved through big data use cases, and work on an individual chosen data project during the course. By the end of the course, participants will be able to work through multiple methods and practical approaches to leverage Spark for advanced big data analytics.

This course will be highly technical with group discussions, hands-on practical exercises, and group activities being the core focus.

By the end of the course, participants will be able to:

• Understand key big data technologies, including a deep dive into Apache Spark

• Describe the main challenges and advantages of Hadoop map-reduce

• Demonstrate and discuss key technologies for big data storage and compute, such as PostgreSQL and object storage

• Discuss popular machine learning algorithms, deep learning techniques and the importance of ethics in data analytics and artificial intelligence

• Deliver a presentation demonstrating the analytics lifecycle and Spark

This is an advanced level course. It is expected that participants either have a number of years of experience utilizing big data, or have previously attended the Certified Big Data and Data Analytics Practitioner (CBDDAP) course. This course is ideal for data engineers, AI engineers and data scientists. Recommended pre-knowledge includes some python programming experience and data visualization practice.

• Big data utilization

Course content

Why Attend

- Big data analytics structures and technologies
- Ethics and integrity for big data and AI development
- Big data storage
- Apache Spark best practices

Course outline

Big Data Analytics Use Cases

- How can big data projects meet organizational needs
- Big data examples:
- Netflix
- LinkedIn
- Facebook
- Google
- Orbitz
- Dell
- Others
- Best practices in project design
- Assessing the current state of your organization
- Choosing datasets for course projects

Storing Big Data

- Big data architectures and paradigms
- The Hadoop Ecosystem

Course content

Course outline

- Overview of Hadoop
- Hadoop Distributed File System (HDFS)
- Massively parallel processing (MPP) versus distributed in-memory applications
- RDBMSs vs NoSQL DBs
- PostgreSQL
- MongoDB
- Cassandra
- Streaming data
- Data-warehousing versus Data Mart
- Intro to Apache Spark
- Big data SQL hands-on-labs

Computing Big Data

- How to access big data
- Role of cloud computing
- Data movement risk
- Networking and co-location
- Apache Spark lab
- Big data extract, transform, load (ETL) big data compute technologies
- Distributed compute
- High performance clusters vs Apache Spark
- Streaming: Storm, Spark structured streaming
- Apache Spark ETL labs

Course content

Course outline

• Apache Spark data engineering

Big Data Advanced Analytics and AI

• Analytics Lifecycle

• Apache Spark vs Pandas

• Big data machine learning & deep learning in Spark

• Importance of ethics in AI

• Automl & Hyperparameter tuning

Course Big Data Projects

• Identify analytical opportunities in an organization

• Define and assess the problem

• Describe the impact and use of data to address the problem

• Identify potential data sources

• Design a data analytics project

• Access, explore, analyze and visualize chosen dataset for project

• Present project insights in course

Seminar dates

Available seminar dates

Live dates and pricing for Certificate in Advanced Big Data and Data Analytics (CABDDA) generated from the course details page.

Date	Location	Format	Fee
22 - 26 June 2026	Frankfurt - Germany	Classroom	**€3,250.-**
13 - 17 July 2026	Rome - Italy	Classroom	**€4,250.-**
17 - 21 August 2026	Kuala lumpur - Malaysia	Classroom	**€2,250.-**
21 - 25 September 2026	Barcelona - Spain	Classroom	**€3,850.-**
19 - 23 October 2026	London - U.K	Classroom	**€4,200.-**
2 - 6 November 2026	Munich - Germany	Classroom	**€3,450.-**
21 - 25 December 2026	Amsterdam - Netherlands	Classroom	**€4,250.-**

Live online option	Online delivery is available at €1,850.-.