

Big Data Technology Fundamentals

Course code: AWSBDF

Who is the course for Individuals who are new to big data concepts, including Enterprise Solutions Architects, Big Data Solutions Architects, Data Scientists, and Data Analysts.

What we teach you This course teaches you how to:

Identify common tools and technologies that can be used to create big data solutions Understand the MapReduce programming framework, including the map, shuffle and sort, and reduce components Distinguish options available for creating a big data solution using the Hive programming framework Please register for free here.

Required skills We recommend that attendees of this course have:

Working knowledge of basic programming in a language such as Java or C# Teaching methods Professional explanation with practical samples and examples.

Course outline Module 1 – Introduction to Big Data

The Business Importance of Big Data The Hadoop Ecosystem Characteristics of Big Data Processing Big Data Tools and Techniques for Analyzing Big Data Implementing Big Data Solutions Case Study – Social Media Analytics

Module 2 – Introduction to MapReduce and Hadoop

Hadoop Architecture MapReduce Framework MapReduce Programming MapReduce and HDFS/S3 Use Case – Recommendation Engine

Module 3 – Data Analysis Using Pig Programming

Introduction to Pig Pig Data Types Representing Data in Pig Running Pig User-Defined Functions Pig vs

Traditional RDBMSs Advanced Techniques in Pig

Module 4 – Big Data Querying with Hive
Introduction to Hive Representing Data in Hive Hive Data Types Probing Data with Hive Queries Hive and AWS Use Case – Ad Hoc Analysis and Product Feedback

Who is the course for

Individuals who are new to big data concepts, including Enterprise Solutions Architects, Big Data Solutions Architects, Data Scientists, and Data Analysts.

What we teach you

This course teaches you how to:

- Identify common tools and technologies that can be used to create big data solutions
- Understand the MapReduce programming framework, including the map, shuffle and sort, and reduce components
- Distinguish options available for creating a big data solution using the Hive programming framework

Please register for free here.

Required skills

We recommend that attendees of this course have:

- Working knowledge of basic programming in a language such as Java or C#

Course outline

Module 1 – Introduction to Big Data

- The Business Importance of Big Data
- The Hadoop Ecosystem
- Characteristics of Big Data
- Processing Big Data
- Tools and Techniques for Analyzing Big Data
- Implementing Big Data Solutions
- Case Study – Social Media Analytics

Module 2 – Introduction to MapReduce and Hadoop

- Hadoop Architecture
- MapReduce Framework
- MapReduce Programming

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved

Big Data Technology Fundamentals

- MapReduce and HDFS/S3
- Use Case – Recommendation Engine

Module 3 – Data Analysis Using Pig Programming

- Introduction to Pig
- Pig Data Types
- Representing Data in Pig
- Running Pig
- User-Defined Functions
- Pig vs Traditional RDBMSs
- Advanced Techniques in Pig

Module 4 – Big Data Querying with Hive

- Introduction to Hive
- Representing Data in Hive
- Hive Data Types
- Probing Data with Hive Queries
- Hive and AWS
- Use Case – Ad Hoc Analysis and Product Feedback

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved