Building Batch Data Analytics Solutions on AWS

In this course, you will learn to build batch data analytics solutions using Amazon EMR, an enterprise-grade Apache Spark and Apache Hadoop managed service. You will learn how Amazon EMR integrates with open-source projects such as Apache Hive, Hue, and HBase, and with AWS services such as AWS Glue and AWS Lake Formation. The course addresses data collection, ingestion, cataloging, storage, and processing components in the context of Spark and Hadoop. You will learn to use EMR Notebooks to support both analytics and machine learning workloads. You will also learn to apply security, performance, and cost management best practices to the operation of Amazon EMR.

This course includes presentations, interactive demos, practice labs, discussions, and class exercises.

Course Contents
Module A: Overview of Data Analytics and the Data Pipeline
Module 1: Introduction to Amazon EMR
Module 2: Data Analytics Pipeline Using Amazon EMR: Ingestion and Storage
Module 3: High-Performance Batch Data Analytics Using Apache Spark on Amazon EMR
Module 4: Processing and Analyzing Batch Data with Amazon EMR and Hive
Module 5: Serverless Data Processing
Module 6: Security and Monitoring of Amazon EMR Clusters
Module 7: Designing Batch Data Analytics Solutions
Module B: Developing Modern Data Architectures on AWS

You have access to the labs for another 14 days after the course. This way you can repeat exercises or deepen them individually.

E-Book You will receive the original course documentation by Amazon Web Services as an e-book.

Target Group
This course is intended for:
• Data platform engineers
• Architects and operators who build and manage data analytics pipelines

Individuals with at least one year of experience managing open source data frameworks such as Apache Spark or Apache Hadoop will benefit from this course.

Prerequisites
Students with a minimum one-year experience managing open-source data frameworks such as Apache
Spark or Apache Hadoop will benefit from this course.

We suggest the AWS Hadoop Fundamentals course for those that need a refresher on Apache Hadoop.
We recommend that attendees of this course have:
• Completed either AWS Technical Essentials or Architecting on AWS
• Completed either Building Data Lakes on AWS or Getting Started with AWS Glue

Status 03/26/2024

This Course in the Web
You can find the up-to-date information and options for ordering under the following link:
www.experteach-training.com/go/**AWBB**

Reservation
On our Website, you can reserve a course seat for 7 days free of charge and in an non-committal manner. This can also be done by phone under +49 6074/4868-0.

Guaranteed Course Dates
To ensure reliable planning, we are continuously offering a wide range of guaranteed course dates.

Your Tailor-Made Course!
We can precisely customize this course to your project and the corresponding requirements.

Training		Prices, excl. of V.A.T.
Classes in Germany	**1 Day**	**€ 795**
Online Training	**1 Day**	**€ 795**
Dates upon request		

Amazon Web Services (AWS)

ExperTeach

Table of Contents
Building Batch Data Analytics Solutions on AWS