

Preparing with Cloudera Data Engineering

powered by Apache Spark, Hive, and Airflow

This hands-on course teaches the key concepts and skills developers need to develop high-performance, parallel applications on the Cloudera Data Platform (CDP) with Apache Spark.

Practical exercises allow you to practice writing Spark applications that integrate with CDP core components. You will learn how to use Spark SQL to query structured data, how to use Hive functions to ingest and denormalize data, and how to work with "big data" stored in a distributed file system.

Course Contents

- HDFS Introduction
- YARN Introduction
- Working with RDDs
- Working with DataFrames
- Introduction to Apache Hive
- Working with Apache Hive
- Hive and Spark Integration
- Distributed Processing Challenges
- Spark Distributed Processing
- Spark Distributed Persistence
- Data Engineering Service
- Workload XM
- Appendix: Working with Datasets in Scala

E-Book You will receive the original documentation as a E-Book (pdf).

Target Group

This course is intended for developers and data engineers.

Prerequisites

You are expected to have basic knowledge of Linux and basic knowledge of the programming languages Python or Scala. Basic knowledge of SQL is helpful. Previous knowledge of Spark and Hadoop is not required.

Course Target

- Distribute, store and process data in a CDP cluster
- Write, configure and deploy Apache Spark applications
- Use Spark interpreters and Spark applications to explore, process and analyze distributed data
- Query data with Spark SQL, DataFrames and Hive tables
- Deploying a Spark application on the Data Engineering Service

This Course in the Web



You can find the up-to-date information and options for ordering under the following link:

www.expertech-training.com/go/CLSP

Reservation

On our Website, you can reserve a course seat for 7 days free of charge and in a 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	4 Days	€ 3,495
Online Training	4 Days	€ 3,495
Dates upon request		

Status 03/09/2025



Our Training Offer for You:



Classroom Training

The live training experience in our training centers or on your premises on-site.



Online Training

Attend your course in online mode – without travel and accommodation expenses and effort.



Hybrid Training

Classroom & online in one course – choose how you wish to participate in your training event.



In-house Training

We create your customized training concepts for your project.



Guaranteed Course Dates

ExperTeach guaranteed course dates allow for secure and reliable planning.

Awards Won by ExperTeach



ExperTeach Benelux B.V.

Ceresstraat 1 · 4811 CA Breda · Phone: +49 6074 4868-0 · Fax: +49 6074 4868-109 · info@experteach.de · www.experteach-training.com