

Building Data Lakes on AWS – Intensive Training («AWSB04»)

In this course, you will learn how to build an operational data lake that supports analysis of both structured and unstructured data.

Duration: 1 day

Price: 900.–

Course documents: Digital original AWS courseware

Content

You will learn the components and functionality of the services involved in creating a data lake. You will use AWS Lake Formation to build a data lake, AWS Glue to build a data catalog, and Amazon Athena to analyze data. The course lectures and labs further your learning with the exploration of several common data lake architectures.

Outline

Module 1: Introduction to data lakes

- Describe the value of data lakes
- Compare data lakes and data warehouses
- Describe the components of a data lake
- Recognize common architectures built on data lakes

Module 2: Data ingestion, cataloging, and preparation

- Describe the relationship between data lake storage and data ingestion
- Describe AWS Glue crawlers and how they are used to create a data catalog
- Identify data formatting, partitioning, and compression for efficient storage and query
- Lab 1: Set up a simple data lake

Module 3: Data processing and analytics

- Recognize how data processing applies to a data lake
- Use AWS Glue to process data within a data lake
- Describe how to use Amazon Athena to analyze data in a data lake

Module 4: Building a data lake with AWS Lake Formation

- Describe the features and benefits of AWS Lake Formation
- Use AWS Lake Formation to create a data lake
- Understand the AWS Lake Formation security model
- Lab 2: Build a data lake using AWS Lake Formation

Module 5: Additional Lake Formation configurations

- Automate AWS Lake Formation using blueprints and workflows
- Apply security and access controls to AWS Lake Formation
- Match records with AWS Lake Formation FindMatches
- Visualize data with Amazon QuickSight
- Lab 3: Automate data lake creation using AWS Lake Formation blueprints
- Lab 4: Data visualization using Amazon QuickSight

Module 6: Architecture and course review

- Post course knowledge check
- Architecture review
- Course review

Key Learnings

- Applying data lake methodologies in planning and designing a data lake
- Articulating the components and services required for building an AWS data lake
- Securing a data lake with appropriate permission
- Ingesting, storing, and transforming data in a data lake
- Querying, analyzing, and visualizing data within a data lake

Methodology & didactics

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

Target audience

This course is intended for the following job roles:

- Data Engineer
- Machine Learning & AI

Requirements

We recommend that attendees of this course:

- have taken the free digital course [Data Analytics Fundamentals](#)
- and have attended the following course (or have equivalent knowledge):
- [Architecting on AWS with Best Practice – Intensive Training \(«AWSA10»\)](#)
- [AWS Technical Essentials – Intensive Training \(«AWSE01»\)](#)

Further courses

- [Building Data Analytics Solutions Using Amazon Redshift – Intensive Training \(«AWSB06»\)](#)
- [Data Warehousing on AWS – Intensive Training \(«AWSA05»\)](#)
- [Building Batch Data Analytics Solutions on AWS – Intensive Training \(«AWSB05»\)](#)
- [Building Streaming Data Analytics Solutions on AWS – Intensive Training \(«AWSB08»\)](#)

Any questions?

We are happy to advise you on +41 44 447 21 21 or info@digicomp.ch. You can find detailed information about dates on www.digicomp.ch/courses-it-provider/amazon-web-services-aws/aws-data-engineer/course-building-data-lakes-on-aws-intensive-training