

Building Data Lakes on AWS

Duration: 1 Day(s)

Course Overview

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

- Apply data lake methodologies in planning and designing a data lake
- Articulate the components and services required for building an AWS data lake
- Secure a data lake with appropriate permission
- Ingest, store, and transform data in a data lake
- Query, analyze, and visualize data within a data lake

Review this course online at <https://www.alta3.com/courses/AWS-DL>

Objectives

- Design and plan an operational data lake using AWS services
- Implement and secure data lakes with appropriate permissions
- Ingest, store, and transform data efficiently within the lake
- Visualize and analyze data using AWS analytics tools

Who Should Attend

- Data platform engineers
- Solutions Architects
- IT professionals

Prerequisites

- Architecting on AWS
- Developing on AWS
- Systems Operations on AWS

Course Outline

Module 1: Introduction to data lakes

1. Describe the value of data lakes
2. Compare data lakes and data warehouses
3. Describe the components of a data lake
4. Recognize common architectures built on data lakes

Module 2: Data ingestion, cataloging and preparation

5. Describe the relationship between data lake storage and data ingestion

6. Describe AWS Glue crawlers and how they are used to create a data catalog
7. Identify data formatting, partitioning, and compression for efficient storage and query
8. Lab 1: Set up a simple data lake

Module 3: Data processing and analytics

9. Recognize how data processing applies to a data lake
10. Use AWS Glue to process data within a data lake
11. Describe how to use Amazon Athena to analyze data in a data lake

Module 4: Building a data lake with AWS Lake Formation

12. Describe the features and benefits of AWS Lake Formation
13. Use AWS Lake Formation to create a data lake
14. Understand the AWS Lake Formation security model
15. Lab 2: Build a data lake using AWS Lake Formation

Module 5: Additional Lake Formation configurations

16. Automate AWS Lake Formation using blueprints and workflows
17. Apply security and access controls to AWS Lake Formation
18. Match records with AWS Lake Formation FindMatches
19. Visualize data with Amazon QuickSight
20. Lab 3: Automate data lake creation using AWS Lake Formation blueprints
21. Lab 4: Data visualization using Amazon QuickSight

Module 6: Architecture and course review

22. Post course knowledge check
23. Architecture review
24. Course review