
AWS – Data Warehousing on Amazon Web Services

Description

This 3-day course introduces you to concepts, strategies, and best practices for designing a cloudbased data warehousing solution using Amazon Redshift, the petabyte-scale data warehouse in AWS. This course demonstrates how to collect, store, and prepare data for the data warehouse by using AWS services such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis, and Amazon S3. Additionally, this course demonstrates how to use Amazon QuickSight to perform analysis on your data.

Classroom Registration Price (CHF)

2500

Virtual Classroom Registration Price (CHF)

2500

Course Content

Module 1: Introduction to Data Warehousing

- Relational databases
- Data warehousing concepts
- The intersection of data warehousing and big data
- Overview of data management in AWS

Module 2: Introduction to Amazon Redshift

- Conceptual overview
- Real-world use cases

Module 3: Launching clusters

- Building the cluster
- Connecting to the cluster
- Controlling access
- Database security
- Load data

Module 4: Designing the database schema

- Schemas and data types
- Columnar compression
- Data distribution styles
- Data sorting methods

Module 5: Identifying data sources

- Data sources overview
- Amazon S3
- Amazon DynamoDB
- Amazon EMR
- Amazon Kinesis Data Firehose

-
- AWS Lambda Database Loader for Amazon Redshift

Module 6: Loading data

- Preparing Data
- Loading data using COPY
- Maintaining tables
- Concurrent write operations
- Troubleshooting load issues

Module 7: Writing queries and tuning for performance

- Amazon Redshift SQL
- User-Defined Functions (UDFs)
- Factors that affect query performance
- The EXPLAIN command and query plans
- Workload Management (WLM)

Module 8: Amazon Redshift Spectrum

- Amazon Redshift Spectrum
- Configuring data for Amazon Redshift Spectrum
- Amazon Redshift Spectrum Queries

Module 9: Maintaining clusters

- Audit logging
- Performance monitoring
- Events and notifications
- Resizing clusters
- Backing up and restoring clusters
- Resource tagging and limits and constraints

Module 10: Analyzing and visualizing data

- Power of visualizations
- Building dashboards
- Amazon QuickSight editions and features

Lab / Exercises

- Official AWS Labs

Documentation

- Digital courseware included

Participant profiles

- Database Architects
- Database Administrators
- Database Developers
- Data Analysts
- Data Scientists

Prerequisites

- Familiarity with relational databases and database design concepts
- Completed the [AWS Technical Essentials](#) classroom training or have equivalent experience

Objectives

- Discuss the core concepts of data warehousing, and the intersection between data warehousing and big data solutions
- Launch an Amazon Redshift cluster and use the components, features, and functionality to implement a data warehouse in the cloud
- Use other AWS data and analytic services, such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis, and Amazon S3, to contribute to the data warehousing solution
- Architect the data warehouse
- Identify performance issues, optimize queries, and tune the database for better performance
- Use Amazon Redshift Spectrum to analyze data directly from an Amazon S3 bucket
- Use Amazon QuickSight to perform data analysis and visualization tasks against the data warehouse

Niveau

Intermédiaire

Duration (in Days)

3

Reference

AWS-12