

Data Warehousing on AWS – Intensive Training («AWSA05»)

Learn how to design a cloud-based data warehousing solution using Amazon Redshift.

Duration: 3 days

Price: 2'700.–

Course documents: Digital original AWS courseware

Content

Note: This course is additionally enriched with content from the course [Building Data Analytics Solutions using Amazon Redshift](#).

In this course, you will learn concepts, strategies, and best practices for designing a cloud-based data warehousing solution using Amazon Redshift, the petabyte-scale data warehouse in AWS. We will demonstrate how to collect, store, and prepare data for the data warehouse by using other AWS services, such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose, and Amazon Simple Storage Service (Amazon S3). We will also explore how to use business intelligence (BI) tools to perform analysis on your data.

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
- Hands-on lab 1: Introduction to Amazon Redshift

Module 2: Introduction to Amazon Redshift

- Conceptual overview
- Real-world use cases
- Interactive Demo 1: Touring the Amazon Redshift console
- Hands-on lab 2: Launching an Amazon Redshift cluster
- RA3 Nodes and AQUA architecture
- Amazon Redshift ML

Module 3: Launching clusters

- Building the cluster
- Connecting to the cluster
- Controlling access
- Database security
- Load data
- Practice Lab 1: Load and query data in an Amazon Redshift cluster
- Optional Lab: Launching an Amazon Redshift Cluster

Module 4: Designing the database schema

- Schemas and data types
- Columnar compression
- Data distribution styles
- Data sorting methods
- Hands-on lab 3: Optimizing database schemas

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
- Redshift Data API
- SUPER Data Type
- Interactive Demo 2: Connecting your Amazon Redshift cluster using a Jupyter notebook with Data API
- Interactive Demo 3: Analyzing semi-structured data using the SUPER data type
- Hands-on lab 4: Loading real-time data into an Amazon Redshift database

Module 6: Loading data

- Preparing Data
- Data Warehousing on AWS
- Loading data using COPY
- Maintaining tables
- Concurrent write operations
- Troubleshooting load issues
- Hands-on lab 5: Loading data with the COPY command

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)
- Interactive Demo 4: Applying mixed workload management on Amazon Redshift
- Hands-on lab 6: Configuring workload management

Module 8: Amazon Redshift Spectrum

- Amazon Redshift Spectrum
- Configuring data for Amazon Redshift Spectrum
- Amazon Redshift Spectrum Queries
- Data Transformation
- Data Sharing
- Practice Lab 2: Data analytics using Amazon Redshift Spectrum
- Practice Lab 3: Data transformation and querying in Amazon Redshift
- Hands-on lab 7: Using Amazon Redshift Spectrum

Module 9: Maintaining clusters

- Audit logging
- Performance monitoring
- Events and notifications
- Hands-on lab 8: Auditing and monitoring clusters
- Resizing clusters
- Backing up and restoring clusters
- Resource tagging and limits and constraints
- Hands-on lab 9: Backing up, restoring and resizing clusters
- Optional: Analyzing and Visualizing Data

Module 10: Analyzing and visualizing data

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

Key Learnings

- Evaluating the relationship between Amazon Redshift and other Big Data systems
- Evaluating use cases for data warehousing workloads and reviewing real-world implementation of AWS data and analytic services as part of a data warehousing solution
- Choosing an appropriate Amazon Redshift node type and size for your data needs
- Understanding which security features are appropriate for Amazon Redshift, such as encryption, IAM permissions, and database permissions
- Launching an Amazon Redshift cluster and using the components, features, and functionality to implement a data warehouse in the cloud
- Using other AWS data and analytic services, such as Amazon DynamoDB, Amazon EMR, Amazon Kinesis Firehose, and Amazon S3, to contribute to the data warehousing solution
- Evaluating approaches and methodologies for designing data warehouses
- Identifying data sources and assessing requirements that affect the data warehouse design
- Designing the data warehouse to make effective use of compression, data distribution, and sort methods
- Loading and unloading data and performing data maintenance tasks
- Writing queries and evaluating query plans to optimize query performance
- Configuring the database to allocate resources such as memory to query queues and defining criteria to route certain types of queries to your configured query queues for improved processing
- Auditing, monitoring, and receiving event notifications about activities in the data warehouse by using features and services such as Amazon Redshift database audit logging, Amazon CloudTrail, Amazon CloudWatch, and Amazon Simple Notification Service (Amazon SNS)
- Preparing for operational tasks such as resizing Amazon Redshift clusters and using snapshots to back up and restore clusters
- Using a BI application to perform data analysis and visualization tasks against your data

Methodology & didactics

Instructor-led training, hands-on labs, demonstrations, and group exercises

Target audience

This course is intended for the following job roles:

- Data Analytics

Certification

IMPORTANT: This course prepares you for [AWS Data Analytics Certification](#), among other courses of the *Data Analytics* job role track

Further courses

- [Building Batch Data Analytics Solutions on AWS – Intensive Training \(«AWSB05»\)](#)

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-digital-transformation-technologies/cloud/amazon-web-services-aws/aws-data-engineer/course-data-warehousing-on-aws-intensive-training

