

Building Data Analytics Solutions Using Amazon Redshift (BDASAR)

ID AW-BDASAR Duración 1 día

Quién debería asistir

This course is intended for:

- Data warehouse engineers
- Data platform engineers
- Architects and operators who build and manage data analytics pipelines

Prerrequisitos

Students familiar with combining AWS technologies to support data lakes or other data-driven workloads will benefit from this course. We recommend that attendees of this course have:

- Completed either [AWS Technical Essentials \(AWSE\)](#) or Architecting on [Architecting on AWS \(AWSA\)](#)
- Completed [Building Data Lakes on AWS \(BDLA\)](#)

Objetivos del curso

In this course, you will learn to:

- Compare the features and benefits of data warehouses, data lakes, and modern data architectures
- Design and implement a data warehouse analytics solution
- Identify and apply appropriate techniques, including compression, to optimize data storage
- Select and deploy appropriate options to ingest, transform, and store data
- Choose the appropriate instance and node types, clusters, auto scaling, and network topology for a particular business use case
- Understand how data storage and processing affect the analysis and visualization mechanisms needed to gain actionable business insights
- Secure data at rest and in transit
- Monitor analytics workloads to identify and remediate problems
- Apply cost management best practices

Contenido del curso

Module A: Overview of Data Analytics and the Data Pipeline

- Data analytics use cases
- Using the data pipeline for analytics

Module 1: Using Amazon Redshift in the Data Analytics Pipeline

- Why Amazon Redshift for data warehousing?
- Overview of Amazon Redshift

Module 2: Introduction to Amazon Redshift

- Amazon Redshift architecture
- Interactive Demo 1: Touring the Amazon Redshift console
- Amazon Redshift features
- Practice Lab 1: Setting up your data warehouse using Amazon Redshift

Module 3: Ingestion and Storage

- Ingestion
- Interactive Demo 2: Connecting your Amazon Redshift cluster using a Jupyter notebook with Data API
- Data distribution and storage
- Interactive Demo 3: Analyzing semi-structured data using the SUPER data type
- Querying data in Amazon Redshift
- Practice Lab 2: Data analytics using Amazon Redshift Spectrum

Module 4: Processing and Optimizing Data

- Data transformation
- Advanced querying
- Practice Lab 3: Data transformation and querying in Amazon Redshift
- Resource management
- Interactive Demo 4: Applying mixed workload management on Amazon Redshift
- Automation and optimization

Module 5: Security and Monitoring of Amazon Redshift Clusters



- Securing the Amazon Redshift cluster
- Monitoring and troubleshooting Amazon Redshift clusters

Module 6: Designing Data Warehouse Analytics Solutions

- Data warehouse use case review
- Activity: Designing a data warehouse analytics workflow

Module B: Developing Modern Data Architectures on AWS

- Modern data architectures

Centros de Entrenamiento Mundial

