Building Data Analytics Solutions Using Amazon Redshift

Cursusduur: 1 Dag   **Cursuscode: GK7379**

Beschrijving:

In this course, you will build a data analytics solution using Amazon Redshift, a cloud data warehouse service. The course focuses on the data collection, ingestion, cataloging, storage, and processing components of the analytics pipeline. You will learn to integrate Amazon Redshift with a data lake to support both analytics and machine learning workloads. You will also learn to apply security, performance, and cost management best practices to the operation of Amazon Redshift.

Doelgroep:

This course is intended for data warehouse engineers, data platform engineers, and architects and operators who build and manage data analytics pipelines.

Doelstelling:

- 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

Vereiste kennis en vaardigheden:

Students with a minimum one-year experience managing data warehouses will benefit from this course.

We recommend that attendees of this course have:

- Completed either AWS Technical Essentials or Architecting on AWS
- Completed Building Data Lakes on AWS

Cursusinhoud:

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: Load and query data in an Amazon Redshift cluster

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
- Interactive demo 5: Amazon Redshift cluster resizing from the dc2.large to ra3.xlplus cluster

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

Nadere informatie:

Neem voor nadere informatie of boekingen contact op met onze Customer Service Desk 030 - 60 89 444

info@globalknowledge.nl

www.globalknowledge.com/nl-nl/

Iepenhoeve 5, 3438 MR Nieuwegein