Training: AWS
Data Engineering on AWS

TRAINING GOALS:

This comprehensive 3-day instructor-led training provides a deep dive into data engineering practices and solutions on Amazon Web Services (AWS). Participants will learn how to design, build, optimize, and secure data engineering solutions by using AWS services. Topics range from foundational concepts to hands-on implementation of data lakes, data warehouses, and both batch and streaming data pipelines. This course equips data professionals with the skills needed to architect and manage modern data solutions at scale.

Course objectives

In this course, you will learn to do the following:

- Design and implement scalable data lakes and data warehouses on AWS.
- Build, optimize, and secure batch data processing pipelines.
- Develop and manage streaming data solutions.
- Apply best practices for data governance and security.
- Automate data engineering workflows by using AWS services.
- Implement access control and security measures for data solutions.

Intended audience

- Data engineers
- Solutions architects
- DevOps engineers
- IT professionals
- Data analysts looking to expand into data engineering

CONSPECT:

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- Module 1: Data Engineering Roles and Key Concepts
  - The role of a data engineer
  - Data discovery for a data analytics system
  - AWS services for data workflows
  - Continuous integration and continuous delivery
  - Networking considerations
- Module 2: Designing and Implementing Data Lakes
  - Data lake introduction
  - Data lake storage
  - Ingest data
  - Catalog data
  - Transform data
  - Serve data for consumption
  - Lab: Setting up a Data Lake on AWS
- Module 3: Optimizing and Securing Data Lake Solutions
  - Optimizing performance
  - Security using Lake Formation
  - Setting permissions with Lake Formation
  - Security and governance
  - Troubleshooting
  - Lab: Automating Data Lake Creation using AWS Lake Formation Blueprints
- Module 4: Data Warehouse Architecture and Design Principles
  - Introduction to data warehouses
  - Amazon Redshift overview
  - Ingesting data into Amazon Redshift
  - Processing data
  - Serving data for consumption
  - Lab: Setting up a Data Warehouse using Amazon Redshift Serverless
- Module 5: Performance Optimization Techniques for Data Warehouses
  - Monitoring and optimization options
  - Data optimization in Amazon Redshift
  - Query optimization in Amazon Redshift
  - Data orchestration
- Module 6: Security and Access Control for Data Warehouses
  - Authentication and access control in Amazon Redshift
  - Data security in Amazon Redshift

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- Lab: Working with Amazon Redshift
- Module 7: Designing Batch Data Pipelines
  - Introduction to batch data pipelines
  - Designing a batch data pipeline
  - Ingesting batch data
- Module 8: Implementing Strategies for Batch Data Pipelines
  - Processing and transforming data
  - Transforming data formats
  - Integrating your data
  - Cataloging data
  - Serving data for consumption
  - Lab: A Day in the Life of a Data Engineer
- Module 9: Optimizing, Orchestrating, and Securing Batch Data Pipelines
  - Optimizing the batch data pipeline
  - Orchestrating the batch data pipeline
  - Securing the batch data pipeline
  - Lab: Orchestrating Data Processing in Spark using AWS Step Functions
- Module 10: Streaming Data Architecture Patterns
  - Introduction to streaming data pipelines
  - Ingesting data from stream sources
  - Storing streaming data
  - Processing streaming data
  - Analyzing streaming data
  - Lab: Streaming Analytics with Amazon Managed Service for Apache Flink
- Module 11: Optimizing and Securing Streaming Solutions
  - Optimizing a streaming data solution
  - Securing a streaming data pipeline
  - Lab: Access Control with Amazon Managed Streaming for Apache Kafka
- Module 12: Compliance and Cost Optimization
  - Compliance considerations
  - Cost optimization tools
- Module 13: Course Wrap-Up

REQUIREMENTS:

We recommend that attendees of this course have the following:

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

- Basic understanding of AWS services
- Familiarity with database concepts
- Basic programming or scripting knowledge
- Understanding of data processing fundamentals

Difficulty level



CERTIFICATE:

The participants will obtain certificates signed by AWS (course completion).

This course also helps you prepare for the AWS Certified Data Engineer – Associate DEA-C01 exam and this way gain the AWS Certified Data Engineer – Associate title – associate level. AWS certification exams are offered at Pearson Vue test centers worldwide
https://home.pearsonvue.com/Clients/AWS.aspx

TRAINER:

AWS Authorized Instructor (AAI)

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70

BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007