

DEA-C01: AWS Data Engineer – Associate

Build reliable data pipelines on AWS — from ingestion to the data lake to analytics — and prepare for the DEA-C01 exam.

DURATION

3 days · 24 hours

LEVEL

Intermediate

CURRICULUM

9 modules · 6 labs

DELIVERYLive online, Onsite,
Corporate

Course overview

Data engineers make data usable: they move it from source systems, shape it into trustworthy datasets, store it cost-effectively and keep pipelines running. On AWS that work spans a wide set of services, and choosing the right one for a given volume, latency and budget is the core skill this course develops.

Day one is about ingestion and transformation. You will compare batch and streaming ingestion with Kinesis, Amazon MSK, AWS DMS and AppFlow, then transform data with AWS Glue jobs and PySpark, choosing file formats and partitioning schemes that keep queries fast and cheap. Day two moves to data store management: designing an S3 data lake, cataloguing with the Glue Data Catalog, modelling in Amazon Redshift, and choosing between DynamoDB, RDS and purpose-built stores.

Day three covers operations, security and governance. You will orchestrate pipelines with Step Functions and Amazon MWAA, monitor them with CloudWatch, check data quality, and secure the lake with Lake Formation permissions, encryption and auditing. Throughout the course, SQL and Python are used in short, practical doses so that you can follow along even if you are stronger in one than the other.

Labs connect into one pipeline: raw files and a stream land in S3, are catalogued and transformed, then queried with Athena and loaded into Redshift. The final session reviews exam technique and practice questions. This course is aligned to the DEA-C01 exam and prepares you for it; confirm the current exam guide on the AWS Certification site before booking.

Highlights

- 24 hours of live, pipeline-building sessions
- Hands-on with Glue, Kinesis, Athena, Redshift and Lake Formation
- Batch and streaming patterns side by side
- Mapped to the four DEA-C01 domains
- Practice questions with detailed explanations

What's included

- One end-to-end pipeline built across the labs
- Batch and streaming ingestion compared in practice
- Glue and PySpark transformation exercises
- Data lake governance with Lake Formation
- Orchestration with Step Functions and MWAA
- Practice questions with explained answers
- Course completion certificate from Ace Skills Consulting

Who this course is for

Who should attend

- Data engineers working with or moving to AWS
- ETL and BI developers modernising pipelines
- Database administrators moving into data engineering
- Analytics engineers building data lakes

Prerequisites

- Working SQL knowledge
- Basic Python programming
- Familiarity with AWS core services (Cloud Practitioner level or above)
- Understanding of data formats such as CSV, JSON and Parquet

Detailed curriculum — 3 days, 24 hours, 9 modules, 6 hands-on labs

Day 1 Data Ingestion & Transformation

8 hours

1.1 Ingestion patterns

- Batch versus streaming ingestion
- Amazon MSK overview
- Amazon AppFlow for SaaS sources
- Kinesis Data Streams and Firehose
- AWS DMS for database replication

LAB Stream events into S3 with Kinesis Data Firehose · 45 min

1.2 Transformation with AWS Glue

- Glue jobs, DynamicFrames and PySpark
- Job bookmarks and incremental loads
- Crawlers and the Data Catalog
- Glue DataBrew for visual preparation

LAB Transform raw CSV to partitioned Parquet with a Glue job · 60 min

1.3 Designing for performance and cost

- Columnar formats and compression
- Serverless versus provisioned processing
- Partitioning and bucketing
- Amazon EMR at a glance

Day 2 Data Store Management

8 hours

2.1 Building the data lake

- S3 layout: raw, curated and consumption zones
- Open table formats such as Apache Iceberg
- Lifecycle policies and storage classes
- Schema evolution

LAB Query the curated zone with Amazon Athena · 40 min

2.2 Amazon Redshift

- Distribution and sort keys
- Redshift Serverless
- COPY, UNLOAD and Redshift Spectrum
- Materialised views and workload management

LAB Load curated data into Redshift Serverless and tune a query · 55 min

2.3 Operational data stores

- DynamoDB data modelling
- Choosing a store by access pattern
- RDS and Aurora for transactional sources

Day 3 Operations, Security, Governance & Exam Readiness

8 hours

3.1 Orchestration and operations

- Step Functions workflows
- Glue workflows and triggers
- Data quality checks
- Amazon MWAA (managed Airflow)
- Monitoring pipelines with CloudWatch

LAB Orchestrate the pipeline with Step Functions and add failure alerts · 50 min

3.2 Security and governance

- Lake Formation permissions and tags
- Masking sensitive data
- Encryption with KMS
- Auditing with CloudTrail

LAB Grant column-level access with Lake Formation · 35 min

3.3 Exam readiness

- DEA-C01 domain recap
- Service-selection question patterns
- Timed practice set with review

Skills and tools

Skills you will build

AWS Glue

PySpark

Amazon Kinesis

AWS DMS

Amazon S3 data lakes

Amazon Athena

Amazon Redshift

Lake Formation

Step Functions

Amazon MWAA

Data modelling

Data quality

Tools used in class

AWS Glue Studio

Amazon Athena

Amazon Redshift Query Editor

AWS CLI

Jupyter notebooks

Apache Airflow

Certification exam

This course is aligned to the exam below and prepares you for it. The exam itself is booked and taken with the vendor.

Exam	AWS Certified Data Engineer – Associate
Exam code	DEA-C01
Questions	65
Duration	130 minutes
Passing score	720 / 1000 (scaled)
Format	Multiple choice, Multiple response
Delivery	Online proctored, Test centre
Validity	3 years
Official page	https://aws.amazon.com/certification/certified-data-engineer-associate/

Exam skill areas covered

- Data Ingestion and Transformation
- Data Store Management
- Data Operations and Support
- Data Security and Governance

Exam details are pending verification. Confirm the current exam code, format and fee on the official vendor page before booking.

Frequently asked questions

How much coding is involved?

Moderate. You will write SQL and short PySpark scripts, with templates provided so the focus stays on design decisions.

Do I need Spark experience?

No. Spark concepts are introduced as needed, and Glue Studio offers a visual option for learners new to it.

Is this course relevant for Databricks or Snowflake users?

Yes. The ingestion, partitioning and governance principles carry over, although the labs use AWS-native services.

What should I study after the course?

We recommend revisiting the labs, then focusing on service-selection questions, which make up much of the exam.

Can our data team attend together?

Yes. Private batches can use sample data shaped like your own sources.

Book a batch or talk to an advisor

Course page: aceskillsconsulting.com/course/dea-c01-aws-data-engineer-associate · Email: info@aceskillsconsulting.com

Chennai office

Guindy, Chennai, Tamil Nadu

Delhi office

B-95, Old Post Office Street, Shakarpur
New Delhi 110092
Office: 98180 12523

Amazon Web Services, AWS and the related marks are trademarks of Amazon.com, Inc. or its affiliates. All product names, logos, brands, trademarks and registered trademarks are the property of their respective owners. Ace Skills Consulting is an independent training provider; unless explicitly stated, it is not affiliated with, endorsed by or sponsored by any vendor mentioned on this site. Exam details are provided for guidance and may change — always confirm on the vendor's official website. Curriculum generated on 23 Oct 2026; topics and order may be adjusted by the trainer to suit each batch.