

AWS Certified Big Data - Specialty (BDS-C00) Exam Prep 2026

AWS BDS-C00 Big Data Specialty. Hadoop, EMR, Kinesis, and large-scale data processing.

IT Certifications Instructor: Brian Taylor • 5 sections • 40 lessons • 14.00h total

[View this course online](#) [Browse all courses](#)

About this course

This course prepares you for the AWS Certified Big Data Specialty (BDS-C00) exam. We cover Hadoop, EMR, Kinesis, Redshift, and data processing at scale.

You will work through hands-on labs, real big data scenarios, and practice exams. The course includes video lessons and downloadable resources.

Course curriculum

Foundations and Core Concepts

- **Understanding the BDS-C00 Exam Blueprint (12m)**
We'll break down the exam domains, question formats, and scoring so you know exactly what to expect on test day.
- **AWS Identity and Access Management (IAM) for Big Data (18m)**
Learn how to configure secure IAM roles and policies for data pipelines and analytics services.
- **Amazon S3: Storage Classes and Lifecycle Policies (22m)**
Master S3 best practices for cost-effective data storage, including Intelligent-Tiering and lifecycle management.
- **VPC Fundamentals for Data Services (15m)**
A step-by-step guide to setting up VPCs, subnets, and security groups to isolate your data workloads.
- **Data Transfer and Network Optimization (9m)**
We'll show you how to minimize costs and latency when moving large datasets into and out of AWS.
- **AWS CloudTrail and CloudWatch for Data Auditing (14m)**
Learn to monitor data service activity and set up essential alerts for your analytics environment.
- **Cost Management for Big Data Workloads (11m)**
A practical look at budgeting, cost allocation tags, and using the Pricing Calculator for data projects.

Data Collection and Ingestion

- **Amazon Kinesis Data Streams vs. Kinesis Firehose (25m)**
A deep dive into when to use each Kinesis service for real-time data ingestion with a streaming case study.
- **Ingesting Data with AWS Snowball and Snowmobile (16m)**
Learn the physical data transfer options and how to plan a large-scale data migration project.
- **Using AWS DataSync for Hybrid Cloud Transfers (13m)**
We'll demonstrate how to automate and accelerate data movement between on-premises storage and AWS.
- **Streaming Ingestion with Amazon Kinesis Data Analytics (21m)**
A problem-solving session on processing streaming data in real-time with SQL and Java applications.
- **Batch Ingestion Patterns with AWS Glue (19m)**
Learn to use Glue crawlers and jobs to catalog and move data from various sources into your data lake.
- **Amazon MSK for Apache Kafka Workloads (28m)**
Understand the architecture of Managed Streaming for Kafka and how it fits into enterprise data pipelines.
- **Ingesting IoT Data with AWS IoT Core (17m)**
A practical guide to connecting IoT devices and routing their data to services like S3 and Kinesis.
- **Database Migration with AWS DMS (24m)**
Learn to migrate on-premises databases to AWS cloud data stores with minimal downtime.

- **Designing a Resilient Data Ingestion Strategy (20m)**

A capstone lesson combining concepts to design a fault-tolerant ingestion pipeline for a retail company.

Data Storage and Management

- **Amazon Redshift: Cluster Architecture and Distribution Styles (32m)**

A deep dive into Redshift internals, helping you choose the right distribution and sort keys for performance.

- **Optimizing Redshift with Spectrum and Concurrency Scaling (26m)**

We'll show you how to use Redshift Spectrum for S3 queries and Concurrency Scaling for peak loads.

- **Amazon DynamoDB: Partitioning, Indexes, and Capacity Modes (30m)**

Master NoSQL design patterns for high-performance applications on DynamoDB.

- **Building a Data Lake with AWS Lake Formation (23m)**

A step-by-step guide to creating a secure, centralized data lake with fine-grained access controls.

- **Amazon DocumentDB and MongoDB Compatibility (15m)**

Learn when to use DocumentDB for your document database workloads on AWS.

- **Comparing HDFS vs. S3 for Big Data Storage (11m)**

A practical comparison to help you decide when to use EMR's HDFS vs. S3 as your primary storage.

- **Data Encryption at Rest and in Transit (18m)**

A real case study analysis of using KMS to encrypt data across S3, Redshift, and RDS.

Data Processing and Transformation

- **Amazon EMR: Cluster Sizing and Instance Selection (29m)**

Learn to optimize EMR clusters for cost and performance for Spark, Hive, and Presto workloads.

- **Running Spark Jobs on EMR with EMRFS (27m)**

A hands-on lesson covering EMRFS, consistent views, and running Spark applications on your cluster.

- **AWS Glue ETL: PySpark and DynamicFrames (31m)**

We'll write a practical Glue ETL job using PySpark to transform and clean complex data formats.

- **Amazon Athena: Querying S3 Data Lakes with SQL (24m)**

A problem-solving session on optimizing Athena queries, partitioning, and using SerDe libraries.

- **AWS Step Functions for Orchestration (20m)**

Learn to build resilient, visual workflows to coordinate multiple AWS services into a data pipeline.

- **AWS Batch for Event-Driven Processing (18m)**

Understand how to run batch computing workloads at any scale without managing servers.

- **Amazon Kinesis Data Analytics for Streaming ETL (22m)**

We'll transform streaming data in real-time using SQL applications in Kinesis Data Analytics.

- **Choosing the Right Processing Service (12m)**

A decision-making framework for when to use EMR, Glue, Athena, or Kinesis for your processing needs.

Data Analysis and Visualization

- **Amazon Redshift Data Warehouse Design (30m)**
Learn to design a star schema, optimize queries, and manage workloads in a Redshift data warehouse.
- **Amazon QuickSight: SPICE, Analysis, and Dashboards (25m)**
A practical guide to building interactive dashboards and using QuickSight's in-memory SPICE engine.
- **Integrating Machine Learning with Amazon SageMaker (35m)**
We'll show you how to build, train, and deploy ML models to gain insights from your big data.
- **Using Amazon Comprehend for Natural Language Processing (19m)**
A real case study on analyzing customer feedback text to extract sentiment and key phrases.
- **Amazon Elasticsearch Service (OpenSearch) for Log Analytics (28m)**
Learn to ingest, index, and visualize log data for operational intelligence and monitoring.
- **Amazon Neptune for Graph Data Analysis (16m)**
A problem-solving session on using graph databases to model complex relationships like social networks.
- **Amazon Forecast for Time-Series Predictions (21m)**
Learn to use machine learning to generate accurate forecasts from your historical time-series data.
- **Amazon Textract for Document Text Extraction (14m)**
A practical lesson on extracting text, forms, and tables from scanned documents for analysis.
- **Building a Complete Analytics Solution (26m)**
A capstone project where you'll design an end-to-end solution from ingestion to visualization for a use case.

Discover our blog

Learn more with free articles written by our professors and guest experts.

Latest article: **Flight Attendant Preparation: What to Study Before You Apply**

[Visit the blog](#) [All courses](#) [This course](#)