

# AWS Certified Data Engineer Associate



## Online Course

**ZETLAN TECHNOLOGIES**  
[www.zetlantech.com](http://www.zetlantech.com)

## Course Modules



### 1. Perform data ingestion

- Throughput and latency characteristics for AWS services
- Data ingestion patterns (for example, frequency and data history)
- Streaming data ingestion
- Batch data ingestion
- Replay ability of data ingestion pipelines
- Stateful and stateless data transactions

### 2. Transform and process data.

- Creation of ETL pipelines based on business requirements
- Volume, velocity, and variety of data
- Cloud computing and distributed computing
- How to use Apache Spark to process data
- Intermediate data staging locations

Zetlan Technologies

### 3. Orchestrate Data Pipelines

- How to integrate various AWS services to create ETL pipelines
- Event-driven architecture
- Config AWS services for data pipelines based on schedules
- Implementing and maintaining Serverless workflows
- Using notification services to send alerts







## 4. Apply Programming Concepts

- Continuous integration & continuous delivery
- SQL queries (for data source queries and data transformations)
- Infrastructure as code (IaC) for repeatable deployments
- Distributed computing
- Data structures and algorithms
- SQL query optimization

## Data Store Management

## 5. Choose a data store

- Storage platforms and their characteristics
- Storage services and configurations
- Data storage formats (for example, .csv, .txt, Parquet)
- How to align data storage with data migration requirements
- Determine the appropriate storage soln for specific access
- How to manage locks to prevent access to data

## 6. Understand data cataloguing systems

- How to create a data catalog
- Data classification based on requirements
- Components of metadata and data catalogs





## 7. Manage the lifecycle of data

- Appropriate storage solns to address hot and cold data reqmt
- How to optimize the cost of storage based on the data lifecycle
- How to delete data to meet business and legal requirements
- Data retention policies and archiving strategies
- How to protect data with appropriate resiliency and availability

## 8. Design data models and schema evolution

- Data modelling concepts
- How to ensure accuracy and trustworthiness of data
- Best practices for indexing, partition strategies, compression, etc.,
- Model structured, semi-structured, and unstructured data
- Schema evolution techniques

## Data Operations and Support

## 9. Automate data processing by using AWS services

- How to maintain and troubleshoot data processing
- API calls for data processing
- Which services accept scripting





# AWS Certified Data Engineer Associate

## 10. Analyse data by using AWS services

- Tradeoffs between provisioned services and serverless services
- SQL queries (for example, SELECT or JOIN clauses)
- How to visualize data for analysis
- When and how to apply cleansing techniques
- Data aggregation, rolling average, grouping, and pivoting

## 11. Maintain and monitor data pipelines

- How to log application data
- Best practices for performance tuning
- How to log access to AWS services
- Amazon Macie, AWS CloudTrail, and Amazon CloudWatch



## 12. Ensure data quality

- Data sampling techniques
- How to implement data skew mechanisms
- Data validation (data completeness, consistency, etc.,)

## Data Security and Governance



## 13. Apply authentication mechanisms

- VPC security networking concepts
- Differences between managed services and unmanaged services
- Methods (password-based, certificate-based, and role-based)
- Diff bet AWS managed policies and customer managed policies



# AWS Certified Data Engineer Associate

**For Enquiry: +91 8680961847**

## **14. Apply authorization mechanisms**

- Authorization methods
- Principle of least privilege as it applies to AWS security
- Role-based access control and expected access patterns
- Methods to protect data from unauthorized access across services

## **15. Ensure Data Encryption and masking**

- Data encryption options available in AWS analytics services
- Diff between client-side encryption and server-side encryption
- Protection of sensitive data
- Data anonymization, masking, and key salting

## **16. Prepare logs for audit**

- How to log application data
- How to log access to AWS services
- Centralized AWS logs



## **17. Understand data privacy and governance**

- How to protect personally identifiable information (PII)
- Data sovereignty



**Free Advice: +91 9600579474**

**[www.zetlantech.com](http://www.zetlantech.com)**





**LEARN  
REMOTELY!!**

The efficiency of online learning in terms of time management, flexibility, and the ability to access resources anytime, anywhere can be compelling.



**ZETLAN TECHNOLOGIES**  
**[www.zetlantech.com](http://www.zetlantech.com)**

**For contact: +91 8680961847  
+91 9600579474**

