AWS Certified Data Engineer – Associate 7-Day Study Checklist

Day 1 – Orientation & Core Concepts

- ■ Download and review the official exam guide (domains, weightings)
- ■ Understand the exam format: question types, timing, passing score
- ■ Review AWS global infrastructure (Regions, AZs, Edge Locations)
- ■ Learn about batch vs. streaming data pipelines
- ■ Make a list of AWS services relevant to data engineering

Day 2 – Data Ingestion & Transformation

- ■ Study Kinesis Data Streams and Kinesis Firehose basics
- ■ Explore Amazon MSK (managed Kafka) and AWS DMS (Database Migration Service)
- ■ Understand AWS Glue (ETL jobs, crawlers, schema registry)
- ■ Review EventBridge and AppFlow for SaaS and event-driven ingestion
- ■ Hands-on lab: Build a simple ingestion flow (e.g., Kinesis → Lambda → S3)

Day 3 – Data Storage & Modeling

- ■ Understand Amazon S3 (data lake patterns, storage classes, lifecycle)
- ■ Learn Amazon Redshift (architecture, Spectrum, concurrency scaling)
- ■ Review DynamoDB (NoSQL design, partition keys, streams)
- ■ Explore Athena for serverless querying
- ■ Lab: Build an S3 bucket with data partitioning and query via Athena

Day 4 – Processing & Orchestration

- ■ Study Amazon EMR (Spark, Hadoop basics)
- ■ Explore AWS Glue for transformations at scale
- ■ Learn orchestration tools: Step Functions, MWAA (Managed Airflow)
- ■ Review batch vs. streaming processing trade-offs
- ■ Lab: Orchestrate a pipeline (e.g., Glue → Step Functions → Redshift)

Day 5 – Monitoring, Performance & Cost Optimization

- ■ Review CloudWatch (metrics, logs, alarms)
- ■ Explore CloudTrail for governance and auditing
- ■ Learn pipeline optimization: partitioning, file formats (Parquet, ORC vs CSV/JSON)
- ■ Study cost control for data pipelines (right-sizing, storage classes, compute savings plans)
- ■ Lab: Add CloudWatch alarms to pipeline and simulate a scaling scenario

Day 6 – Security & Governance

- ■ Review IAM (roles, policies, least privilege)
- ■ Explore AWS KMS for encryption at rest and in transit
- ■ Learn Macie, Lake Formation, and governance principles
- ■ Understand compliance considerations (logging, retention, GDPR/HIPAA hints)
- ■ Lab: Secure a pipeline (e.g., S3 encryption, access control, audit logging)

Day 7 – Final Review & Practice

- ■ Take a full-length timed practice test
- ■ Review every incorrect answer and confirm understanding
- ■ Skim AWS whitepapers: Well-Architected Framework, Data Analytics Lens, Security Best Practices
- ■ Light hands-on: refresh weak areas (e.g., MSK, DMS, Redshift Spectrum)
- ■ Confirm you're consistently scoring ≥ 80% in practice tests before booking the exam