

# Data Analytics Fundamentals

## AWS Data Analytics Fundamentals

AWS Data Analytics Fundamentals course will help you learn about the process for planning data analysis solutions and the various data analytic processes that are involved. This course takes you through five key factors that indicate the need for specific AWS services in collecting, processing, analyzing, and presenting your data. This includes learning basic architectures, value propositions, and potential use cases. The course introduces you to the AWS services and solutions to help you build and enhance data analysis solutions.

## Intended Audience

This course is intended for:

- Data architects
- Data scientists
- Data analysts

## Course Objectives

This course teaches you how to:

- Identify the characteristics of data analysis solutions and the characteristics that indicate such a solution may be required.
- Define types of data including structured, semistructured, and unstructured data.
- Define data storage types such as data lakes, data warehouses, and the Amazon Simple Storage Service (Amazon S3).
- Analyze the characteristics of and differences in batch and stream processing.
- Define how Amazon Kinesis is used to process streaming data.
- Analyze the characteristics of different storage systems for source data.
- Analyze the characteristics of online transaction processing (OLTP) and online analytical processing (OLAP) systems and their impact on the organization of data within these systems.
  
- Analyze the differences of row-based and columnar data storage methods.
- Define how Amazon EMR, AWS Glue, and Amazon Redshift each work to process, cleanse, and transform data within a data analysis solution.
- Analyze the concept of atomicity, consistency, isolation, and durability (ACID) compliance as well as basic availability, soft state, eventual consistency (BASE) compliance and how an extract, transform, load (ETL) process can help to ensure compliance.
- Explore the concept of data schemas and understand how they define data and how this information is stored in metastores.
- Analyze the concept of data versus information.

- Recognize the ways to analyze data to produce information for reports.
- Define how AWS services work together to visualize data.

## **Duration**

1 Day

## **Prerequisites**

We recommend that attendees of this course have the following prerequisites:

- Working knowledge of database concepts.
- Basic understanding of data storage, processing, and analytics.
- Experience with enterprise IT systems.

## **Course Outline**

This course covers the following concepts:

### **Lesson 1: Introduction to data analysis solutions**

Data analytics and data analysis concepts  
Introduction to the challenges of data analytics

### **Lesson 2: Volume – data storage**

Introduction to Amazon S3  
Introduction to data lakes  
Introduction to data storage methods

### **Lesson 3: Velocity – data processing**

Introduction to data processing methods  
Introduction to batch data processing  
Introduction to stream data processing

### **Lesson 4: Variety – data structure and types**

Introduction to source data storage  
Introduction to structured data stores  
Introduction to semi-structured and unstructured data stores

## Lesson 5: Veracity – cleansing and transformation

Understanding data integrity

Understanding database consistency

Introduction to the ETL process

## Lesson 6: Value – reporting and business intelligence

Introduction to analyzing data

Introduction to visualizing data

## Lesson 7: Key Takeaways

Putting the pieces together

What's next