

# Building Data Lakes on AWS

## AWS クラスルームトレーニング

### コースの説明

このコースでは、構造化データと非構造化データの両方の分析に対応するデータレイクを構築する方法を学びます。データレイクの作成に関連するサービスのコンポーネントや機能について学習します。AWS Lake Formation を使用してデータレイクを構築し、AWS Glue を使用してデータカタログを構築し、Amazon Athena を使用してデータを分析します。コースの講義とラボを通して、いくつかの一般的なデータレイクアーキテクチャについての詳細を学びます。

レベル	実施形式	所要時間
中級	クラスルームトレーニング プレゼンテーション、講義、ハンズオンラボ	1日

### コースの目標

このコースの学習内容は、以下のとおりです。

- データレイクの計画と設計手法を理解する
- AWS 上にデータレイクの構築に必要なコンポーネントとサービスを明確にする
- 適切なパーミッションを付与してデータレイクを保護する
- データレイクを中心としたデータの取り込み、保存、変換について説明する
- データレイク内のデータを分析および可視化する方法について説明する

### 対象者

このコースは以下のような方を対象としています。

- データプラットフォームエンジニア
- ソリューションアーキテクト
- IT プロフェッショナル

### 前提条件

このコースを受講するにあたって、以下の前提条件を満たしておくことをお勧めします。

# Building Data Lakes on AWS

## AWS クラスルームトレーニング

- [AWS Cloud Practitioner Essentials](#) クラスルームコースまたは [AWS Cloud Practitioner Essentials](#) でデジタルコース、[AWS Practical Startup Workshop/AWS Technical Essential](#) クラスルームコースの修了
- データ分析パイプラインの構築に関する 1 年間の経験、または [Data Analytics Fundamentals](#) デジタルコースの修了

### 登録

<https://www.aws.training/training/schedule?courseId=65039&countryName=JP&trainingProviderId=1>

### コースの概要

#### 1 日目

##### モジュール 1: データレイクの概要

- データレイクの価値について説明する
- データレイクとデータウェアハウスを比較する
- データレイクのコンポーネントについて説明する
- データレイク上に構築される一般的なアーキテクチャを理解する

##### モジュール 2: データの取り込み、カタログ化、準備

- データレイクストレージとデータ取り込みの関係について説明する
- AWS Glue クローラと、それを使用してデータカタログを作成する方法について説明する
- ストレージとクエリの効率化に向いているデータのフォーマット、パーティション化、圧縮の方法を特定する
- ラボ 1: シンプルなデータレイクをセットアップする

##### モジュール 3: データの処理と分析

- データ処理がデータレイクにどのように適用されるかを理解する
- AWS Glue を使用してデータレイク内のデータを処理する
- Amazon Athena を使用したデータレイク内のデータの分析方法について説明する

##### モジュール 4: AWS Lake Formation を使用したデータレイクの構築

- AWS Lake Formation の機能と利点について説明する
- AWS Lake Formation を使用してデータレイクを作成する

# Building Data Lakes on AWS

## AWS クラスルームトレーニング

- AWS Lake Formation のセキュリティモデルを理解する
- ラボ 2: AWS Lake Formation を使用してデータレイクを構築する

### モジュール 5: Lake Formation のその他の設定

- ブループリントとワークフローを使用して AWS Lake Formation を自動化する
- AWS Lake Formation にセキュリティとアクセスコントロールを適用する
- AWS Lake Formation FindMatches を使用したレコードの照合
- Amazon QuickSight を使用してデータを可視化する
- ラボ 3: AWS Lake Formation のブループリントを使用してデータレイクの作成を自動化する
- ラボ 4: Amazon QuickSight を使用してデータを可視化する

### モジュール 6: アーキテクチャとコースの復習

- コース修了の知識確認
- アーキテクチャの復習
- コースの復習