

Building Data Lakes on AWS

AWS 課堂培訓

課程說明

在本課程中，您將學習如何建置支援結構化和非結構化資料分析的操作資料湖。您將學習建立資料湖所涉及之服務的元件和功能。您將使用 AWS Lake Formation 建立資料湖，使用 AWS Glue 建立資料目錄，而使用 Amazon Athena 來分析資料。本課程講座和實驗室透過探索幾種常見的資料湖架構，進一步進一步了解您的學習。

- 課程級別：中級
- 持續時間：1 天

活動

本課程包括演示文稿，講座，實作實驗室和小組練習。

課程目標

在本課程中，您將學習：

- 在規劃和設計資料湖時套用資料湖方法
- CO1. 使用已建立的資料湖方法來規劃和設計資料湖。
- CO2. 說明在 AWS 上建立資料湖所需的元件和服務。
- CO3. 說明如何使用適當的許可保護 AWS 上的資料湖。
- CO4. 比較在 AWS 資料湖中擷取、存放和轉換資料的方式。
- CO5. 分析和視覺化存放在 AWS 資料湖中的資料。
- CO6. 在 AWS 上建置和自動化資料湖的部署。
- CO7. 說明資料湖在現代資料架構中的作用。

Building Data Lakes on AWS

AWS 課堂培訓

目標受眾

本課程適用於：

- 資料平台工程師
- 方案架構師
- IT 專業人員

前提

我們建議參加本課程的參加者：

- 完成 [AWS 技術基礎](#) 課堂課程
- 擁有一年建置資料分析管道的經驗，或已完成 [資料分析基礎知識](#) 數位課程

課程大綱

單元 1：資料湖簡介

- 描述資料湖的價值
- 比較資料湖和資料倉儲
- 描述資料湖的元件
- 辨識建置在資料湖上的通用架構

單元 2：資料擷取、編目和準備

- 描述資料湖儲存與資料擷取之間的關係
- 說明 AWS Glue 爬蟲程式，以及如何使用它們建立資料型錄
- 識別資料格式化、分割和壓縮，以便有效率地儲存和查詢

Building Data Lakes on AWS

AWS 課堂培訓

單元 3：使用 AWS Lake Formation 建置資料湖

- 辨識資料處理如何套用至資料湖
- 使用 AWS Glue 處理資料湖中的資料
- 說明如何使用 Amazon Athena 分析資料湖中的資料
- 實驗室 1：使用 AWS Lake Formation 建置資料湖

單元 4：資料處理與分析

- 說明 AWS Lake Formation 的功能和好處
- 使用 AWS Lake Formation 建立資料湖
- 了解 AWS Lake Formation 安全模型
- 實驗室 2：使用 AWS Lake Formation 建置資料湖

單元 5：額外的 Lake Formation 配置

- 解釋可用的內置藍圖以創建和填充新的 Lake Formation
- 說明套用進階權限以保護資料存取和工作流程的方法。
- 描述細細整行/儲存格存取控制
- 解釋基於 Lake Formation 標籤的訪問控制機制以及具名訪問控制與基於標籤的訪問控制的不同用例
- 描述針對連線至 Lake Formation 的分析服務強制執行目錄中繼資料和基礎資料資源的精細存取原則的存取流程

單元 6：現代資料架構

- 說明現代資料架構的功能：可擴充的資料湖、目的建置分析服務、順暢的資料移動、統一控管，以及效能與成本效益
- 闡明現代資料架構中的典型資料移動：由內而外、外部進入、周邊，以及跨越共用。

Building Data Lakes on AWS

AWS 課堂培訓

- 描述建置和維護資料產品即服務的重點。
- 使用 Lake Formation 描述典型的資料網格架構，以及支援此方法的關鍵推動者
- 實驗室 3：在 Lake Formation 構建和發布數據產品

單元 7：課程總結

- 課程後知識檢查
- 建築檢討
- 課程複習