

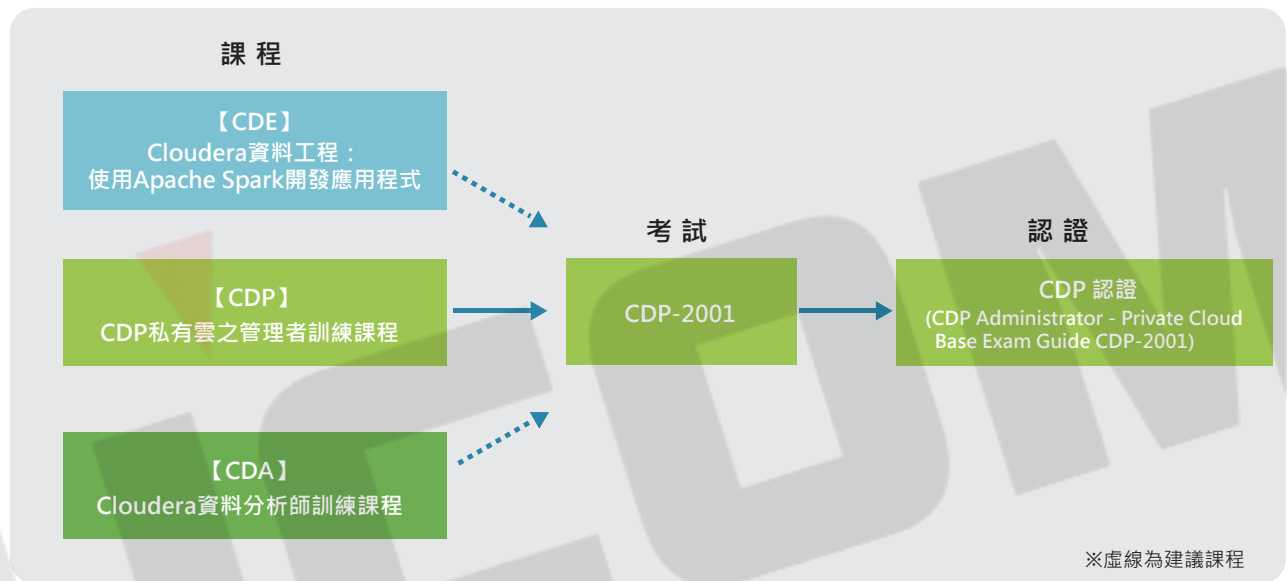
Cloudera 課程簡介

☑ 系統管理 ☑ 資料分析 ☑ 程式開發

Cloudera 認證學習流程圖

大數據時代來臨，不論您的公司是否正在尋找相關技術或是您的客戶正在尋求解決方案，Cloudera所提供的技術及架構，都能為您帶來競爭優勢，現在就開始了解Cloudera並取得相關認證，成為一個新技術的領導者，也是向您的客戶或您的雇主提出技術能力的最好證明。

CDP 認證(CDP Administrator - Private Cloud Base Exam Guide CDP-2001)



獨家課程 CDP認證 (CDP Administrator - Private Cloud Base Exam Guide CDP-2001)

CDP	<h2 style="margin: 0;">CDP私有雲之管理者訓練課程</h2> <p style="margin: 0;">Administrator Training : CDP Private Cloud Base</p> <p style="margin: 0;">時數：28小時 費用：64,000元 點數：19點 教材：Cloudera原廠教材</p>		
適合對象	1. 系統管理人員 2. IT經理 3. RDBMS資料庫管理人員 4. CDP私有雲管理者		
預備知識	基礎的Linux操作(不需要任何Apache Hadoop經驗)		
課程目標	本課程將提供學員關於使用Cloudera Manager操作與維護本地部署叢集所必要的步驟與方法。從安裝與組態到負載平衡與調校，本課程是管理CDP私有雲的管理者在面對真實世界挑戰時最佳的準備課程。 透過講師引導式討論、互動與實際操作練習，您將一覽Cloudera Data Platform的全貌並學習到下列內容：		
後續推薦課程	: Cloudera資料工程：使用Apache Spark開發應用程式		



CDA

Cloudera資料分析師訓練課程

Cloudera Data Analyst Training

時數：28小時 | 費用：64,000元 | 點數：19點 | 教材：Cloudera原廠教材



適合對象	1. 資料分析人員 2. 商業分析人員	3. 程式開發人員 4. 系統管理人員
預備知識	1. SQL 語言 2. 基本Unix/Linux操作(不需要具備任何Apache Hadoop經驗)	
課程目標	<p>本課程將教您如何使用Apache Hive、Apache Impala，讓您可以在大資料環境裡使用傳統的資料分析與商業智慧技術，同時也會展示資料專家如何使用類SQL語言存取、操作、分析複雜的大資料資料集，而不需要使用Java、Scala、Python等程式開發語言，並透過講師引導式討論、互動與實際操作練習，您將一覽Hadoop生態系的全貌。</p> <p>透過本課程您將學習到以下技能：</p> <ol style="list-style-type: none"> 1. 如何使用開放原始碼的大資料生態系工具解決傳統RDBMS無法克服的挑戰 2. 使用Apache Hive與Apache Impala提供SQL介面來存取大資料資料集 3. Hive與Impala的語法與資料格式，包含函數與子查詢 4. 建立、修改與刪除表格、視圖與資料庫，載入資料與儲存查詢結果 5. 建立、使用分區與不同的檔案格式 6. 正確地使用JOIN或UNION組合兩個或多個資料集 7. 了解分析函數與視窗函數以及如何使用它們 8. 儲存與查詢複雜或巢套資料結構 9. 處理與分析半結構化與無結構化資料 10. 優化Hive與Impala查詢的技巧 11. 使用參數、客製檔案格式、SerDes與外部指令稿來擴增Hive與Impala的功能 12. 進行某項工作前，如何決定該使用Hive、Impala或RDBMS或者採用混合模式 	
後續推薦課程	PyPD：資料分析-使用Python Pandas與Matplotlib	



2022 新課

CDE

Cloudera資料工程：使用Apache Spark開發應用程式

Cloudera Data Engineering : Developing Applications with Apache Spark

時數：28小時 | 費用：64,000元 | 點數：19點 | 教材：Cloudera原廠教材



適合對象	1. 使用Spark在Cloudera Data Platform之上撰寫應用程式的開發人員 2. 希望取得CCA Spark and Hadoop Developer與CCP Data Engineer認證的相關人員	
預備知識	1. 熟悉Python或Scala任一種程式語言 2. 基本的Linux Command Line操作	3. 基本的SQL知識
課程目標	<p>本訓練課程以實作型態教授如何在Cloudera Data Platform中使用Apache Spark開發高效能的平行運算應用程式所需要的關鍵概念與專業知識。您將學習到下列知識：如何使用Spark SQL查詢結構化資料與使用Spark Streaming對來自不同資料源的串流資料進行即時處理，此外開發者將會練習到使用Core Spark撰寫應用程式來進行ETL操作與迭代演算法。</p> <p>本課程還包含如何使用位於分散式檔案系統的大資料集，與如何在CDP叢集開發、組態與部署Spark應用程式。</p> <p>完成本課程後，能夠讓您有能力面對真實環境的挑戰與建立Spark應用程式以產生更快、更好的決策，同時能夠將互動式分析，應用於不同的使用案例、架構與產業。</p> <p>透過本課程您將學習到以下技能：</p> <ol style="list-style-type: none"> 1. 了解在CDP Cluster中資料如何分散儲存與分散處理 2. 了解如何撰寫、設定與部署Spark應用程式 3. 使用Spark Shell與Spark應用程式探索、處理與分析分散儲存的資料 4. 使用Spark SQL、DataFrame與Hive Table資料查詢 5. 使用Spark Streaming協同Kafka處理即時資料流 	
後續推薦課程	CDA：Cloudera資料分析師訓練課程	