

CLDR-1 - CLUDERA DATA ANALYST

Categoria: **Cludera**

INFORMAZIONI SUL CORSO



Durata:
5 Giorni



Categoria:
Cludera



Qualifica Istruttore:
Docente Senior (min.
5 anni)



Dedicato a:
Analista



Produttore:
PCSNET

OBIETTIVI

Al termine del corso gli allievi avranno acquisito tutti gli elementi necessari all'implementazione di piattaforme Big Data per l'analisi ed il processamento dei dati attraverso strumenti quali Hive e Impala.

PREREQUISITI

- Buona conoscenza dei database e del linguaggio SQL
- Buona conoscenza del sistema operativo Linux

CONTENUTI

Data Provisioning

Usare comandi Data Definition Language (DDL) per creare o modificare strutture nel metastore usando Hive e Impala:

- Creare nuove tabelle usando vari tipi di dati, delimitatore e formati di file
- Creare nuove tabelle usando tabelle esistenti per definire schemi
- Aumentare le performance creando tabelle partizionate nel metastore
- Modificare tabelle per modificare schemi
- Creare viste per semplificare le query

Data Analysis

Usare istruzioni Query Language (QL) in Hive e Impala per analizzare dati sul cluster:

- Preparare report usando comandi SELECT, inclusi union e subquery
- Calcolare statistiche aggregate, come somme e medie attraverso le query
- Incrociare i dati su tabelle multiple usando i comandi di join
- Trasformare il formato di output delle query usando funzioni built-in
- Eseguire query attraverso un gruppo di righe usando le funzioni di window

INFO

Materiale didattico: Materiale didattico e relativo prezzo da concordare

Costo materiale didattico: NON incluso nel prezzo del corso

Natura del corso: Operativo (previsti lab su PC)