

CLDR-1 - CLUDERA DATA ANALYST

Categoria: **Cloudera**

INFORMAZIONI SUL CORSO



Durata:
5 Giorni



Categoria:
Cloudera



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.

PREREQUISITI

- Buona conoscenza della programmazione strutturata derivata dall'uso di un qualsiasi linguaggio di programmazione
- Buona conoscenza dei database e del SQL
- Buona conoscenza del sistema operativo Linux

CONTENUTI

Trasformare, memorizzare e strutturare i dati per le interrogazioni

- Importare dati da Mysql dentro HDFS usando Sqoop
- Esporta i dati su database Mysql usando Sqoop
- Muovere i dati nelle tabelle nei metastore
- Trasformare valori, colonne o file delimitati prima dell'analisi

Provvedere strutture ai dati

- 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 multiple tabelle 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)