

## Apache Spark Base

CODICE	DT0021
DURATA	2 gg
PREZZO	1.400,00 €
EXAM	

### DESCRIZIONE

---

Azure Databricks è una piattaforma di analisi veloce, semplice e collaborativa che consente l'utilizzo di Apache Spark ed è integrata con i servizi e la sicurezza di Azure.

Questo corso fornirà le basi per avviare un ambiente ed elaborare in parallelo big data archiviati con una logica di distribuzione.

### TARGET

---

Analisti

### PREREQUISTI

---

Conoscenza del linguaggio SQL

### CONTENUTI

---

## Modulo 1

- Introduzione a Spark
- Architettura
- Le API disponibili
- Le strutture dati
- Muovere i primi passi con Spark in Python (pySpark)
- Piccola introduzione al linguaggio Python
- Trasformare i dati con Spark
- Esercitazioni

## Modulo 2

- Effettuare trasformazioni e calcoli avanzati
- SparkSQL

- Spark Machine Learning
- Spark Streaming (cenni)
- Ottimizzare le prestazioni di Spark
- Esercitazioni