Dataskills S.r.l.
Via Dabusti 57 Casteggio – 27045 PV – Italy
CF e P.IVA 02575080185 REA 284697
Tel: +39 0383 19561559 Fax: +39 1782733569

DATA SKILLS
UNDERSTANDING THE WORLD

BIG DATA ANALYTICS

WITH SPARK

DESCRIPTION

Email: info@dataskills.it

Apache Spark is becoming central in the World of Big Data. Spark is a system for distributed calculus, which is fault tolerant and easily scalable. It is able to work efficiently and quickly with huge amounts of data. Spark has many integrated features which add to the platform a few interesting functionalities: SparkSQL, Spark Straming, Spark ML and MLib e GraphX.

The course offers a good understanding of Spark, its main components, and its various features from both a technical standpoint, and through practical exercises.

Such exercises might be carried out in Scala, Python, or R. (Please specify your preference before

enrolling).

MAIN TOPICS

- Introduction to Spark
- The Big Data architecture
- Working with RDD (only Python or Scala)
- Spark SQL and DataFrames
- Machine Learning with Spark ML and Mlib
- Spark Streaming (only Python or Scala)

PREREQUISITES

It is suggested to attend the introductory course: "Big Data Analytics: all you need to know about Big Data".

TAKEAWAYS

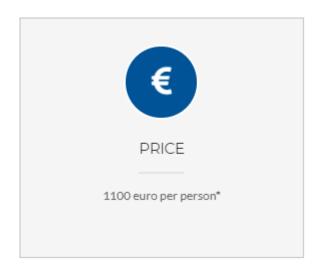
Participants will learn the latest Big Data technologies applied to Business Intelligence. The course offers the possibility to work with real-life cases, similar to those that might be encountered within the firm

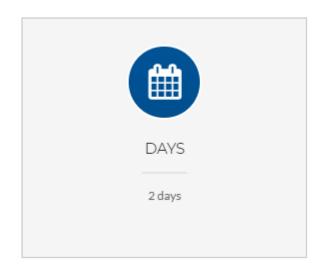
Dataskills S.r.l. Via Dabusti 57 Casteggio – 27045 PV – Italy CF e P.IVA 02575080185 REA 284697 Tel: +39 0383 19561559 Fax: +39 1782733569 DATA SKILLS
UNDERSTANDING THE WORLD

Email: info@dataskills.it

BIG DATA ANALYTICS

WITH SPARK





*Discounts applicable to groups of attenders *Prices do not comprise travels