

Laboratorio 2. (1° giorno - pomeriggio)

Obiettivi

- Prendere confidenza con Spark : DataFrame, Dataset e RDD
- Widget di visualizzazione nel Notebook
- Lettura di file da HDFS

Spark - Documentazione

- <https://spark.apache.org/docs/latest/>
- <https://spark.apache.org/docs/latest/api/scala/index.html#org.apache.spark.sql.Dataset>

Esercizi

Da Zapellin caricare le *note* in *Spark/Lab2/Es1*