

Spark and Hadoop Services @ CERN

- SWAN service: <https://swan.cern.ch>
- Hadoop and Spark service @CERN
 - <https://hadoop-user-guide.web.cern.ch/getstart/access/>

Spark Training Material @ CERN

- CERN IT Spark course and notebooks:
 - <https://github.com/cerndb/SparkTraining>
 - <https://sparktraining.web.cern.ch>
- Example notebooks, SWAN Gallery:
 - https://swan-gallery.web.cern.ch/apache_spark
- Notes and tips on using Spark:
 - https://github.com/LucaCanali/Miscellaneous/tree/master/Spark_Notes
 - https://github.com/LucaCanali/Miscellaneous/tree/master/Spark_Physics



Recommended Reading Material

- **Learning Spark**, 2nd Edition, Jules Damji, Denny Lee, Brooke Wenig, Tathagata Das, O'Reilly 2020
- **Apache Spark: The definitive guide**, Bill Chambers and Matei Zaharia, O'Reilly 2018
- **Spark in Action**, Second Edition, Jean-Georges Perrin, Manning 2020
- **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**, 3rd Edition, Aurélien Géron, O' Reilly 2022

- Read the books online using O'Reilly safari's subscription of the CERN Library
 - <https://learning.oreilly.com/home/>
 - Use your CERN email to log in, you will be redirected to CERN SSO

- Spark project documentation: <https://spark.apache.org/docs/latest/index.html>

