

Reading materials

Books

- [Spark: The Definitive Guide \(http://shop.oreilly.com/product/0636920034957.do\)](http://shop.oreilly.com/product/0636920034957.do)
- [Intro to Statistical Learning \(http://faculty.marshall.usc.edu/gareth-james/isl/\)](http://faculty.marshall.usc.edu/gareth-james/isl/)
- [Elements of Statistical Learning \(https://web.stanford.edu/~hastie/ElemStatLearn/\)](https://web.stanford.edu/~hastie/ElemStatLearn/)
- [Deep Learning by Ian Goodfellow \(http://www.deeplearningbook.org/\)](http://www.deeplearningbook.org/)

Blogs

- [Netflix uses Spark \(http://bit.ly/2Fkx4Mm\)](http://bit.ly/2Fkx4Mm)
- [Capital One uses Spark to detect Fraud \(https://youtu.be/q5HFMVoN_rc\)](https://youtu.be/q5HFMVoN_rc)

References

- [Programming Guide-Spark SQL, DataFrames, Datasets \(http://spark.apache.org/docs/latest/sql-programming-guide.html\)](http://spark.apache.org/docs/latest/sql-programming-guide.html)
- [PySpark 2.4.4 documentations \(http://spark.apache.org/docs/latest/api/python/index.html\)](http://spark.apache.org/docs/latest/api/python/index.html)

Instruction materials

Videos

- [Spark Summit YouTube Channel \(https://www.youtube.com/channel/UCRzsq7k4-kT-h3TDUBQ82-w\)](https://www.youtube.com/channel/UCRzsq7k4-kT-h3TDUBQ82-w)
- [Using Apache Spark 2.0 to Analyze the City of San Francisco's Open Data \(https://youtu.be/K14plpZgy_c\)](https://youtu.be/K14plpZgy_c)

Hand-on labs

Tutorials

- [Spark-Summit-June-2016 \(https://github.com/erdcpatel/Spark-Summit-June-2016\)](https://github.com/erdcpatel/Spark-Summit-June-2016)
- <https://www.digitalocean.com/community/tutorials/how-to-install-anaconda-on-ubuntu-18-04-quickstart>
(<https://www.digitalocean.com/community/tutorials/how-to-install-anaconda-on-ubuntu-18-04-quickstart>)
- <http://spark.apache.org/docs/latest/quick-start.html> (<http://spark.apache.org/docs/latest/quick-start.html>)

-
- <https://www.datacamp.com/community/tutorials/apache-spark-python#gs.fMIlqxM>
(<https://www.datacamp.com/community/tutorials/apache-spark-python#gs.fMIlqxM>).

Packages

- TensorFrames - <https://spark-packages.org/package/databricks/tensorframes> (<https://spark-packages.org/package/databricks/tensorframes>)
- Spark Deep Learning - <https://spark-packages.org/package/databricks/spark-deep-learning> (<https://spark-packages.org/package/databricks/spark-deep-learning>)

In []: