**Resources about PySpark**

Official Document

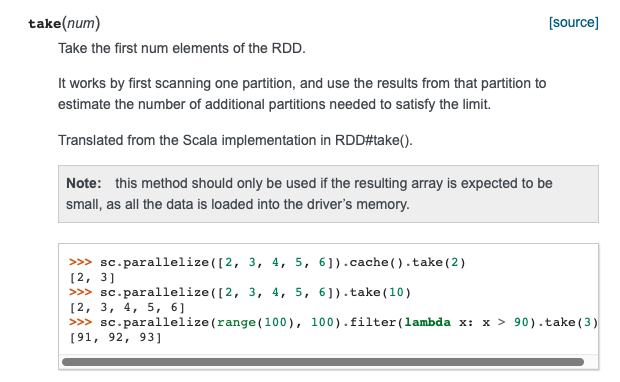
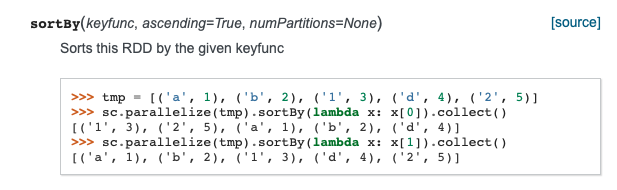
<https://spark.apache.org/docs/latest/api/python/pyspark.html#pyspark.SparkContext.wholeTextFiles>

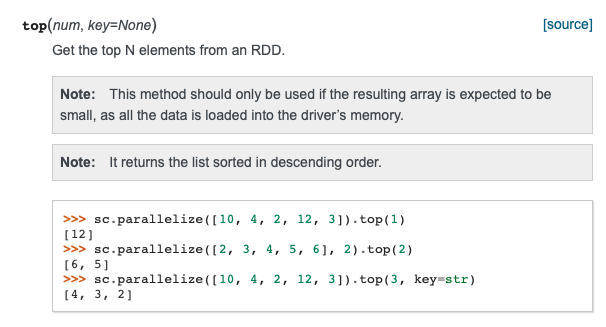
First Steps With PySpark and Big Data Processing

<https://realpython.com/pyspark-intro/#hello-world-in-pyspark>

How to sort by value efficiently in PySpark?

<https://stackoverflow.com/questions/33706408/how-to-sort-by-value-efficiently-in-pyspark>





**Errors & Solutions**

Error:

py4j.protocol.Py4JJavaError: An error occurred while calling z:org.apache.spark.api.python.PythonRDD.collectAndServe.

: java.lang.IllegalArgumentException: Unsupported class file major version 57

Solution:

Spark use Java 8 while my Java is Java 13

<https://stackoverflow.com/questions/53583199/pyspark-error-unsupported-class-file-major-version-55>

JDK 13 Uninstall

<https://explainjava.com/uninstall-java-macos/>

JDK 8 Installation for OS X

<https://docs.oracle.com/javase/8/docs/technotes/guides/install/mac_jdk.html#CHDBADCG>