<https://github.com/KirkYagami/pyspark-playground.git>

docker exec da-spark-master spark-submit --master spark://spark-master:7077 --deploy-mode client --executor-memory 4G --driver-memory 2G --executor-cores 2 --conf spark.executor.memoryOverhead=512m --conf spark.serializer=org.apache.spark.serializer.KryoSerializer ./apps/data\_analysis\_book/chapter03/word\_non\_null.py