**Kafka Spark Integration**

1. Start producer for wc\_input topic we created in another exercise

bin/kafka-console-producer.sh --broker-list localhost:9092 --topic wc\_input

1. Download assembly JAR from

<https://mvnrepository.com/artifact/org.apache.spark/spark-streaming-kafka-assembly_2.10/1.6.0>

1. Run python spark code as

*spark-submit --jars spark-streaming-kafka-assembly\_2.10-1.6.0.jar kafka\_wordcount.py localhost:2181 wc\_input*

**Note:** In case you are using scala/java API, a sample command would look like following:

*spark-submit --class kafkaspark.KafkaSparkStream --master local[2] --packages org.apache.spark:spark-streaming-kafka\_2.10:1.6.0 --deploy-mode client kafkasparkstreaming\_2.12-0.1.jar*

Why don’t you try using --packages option for python too?

Sample command:

If you want to test it out on spark-shell then you may launch it as:

spark-shell *--packages org.apache.spark:spark-streaming-kafka\_2.10:1.6.0*

**Reference:**

* <https://mvnrepository.com/artifact/org.apache.spark/spark-streaming-kafka-assembly_2.10/1.6.0>
* <https://spark.apache.org/docs/1.6.0/streaming-kafka-integration.html>