

Kafka Word Count Scala

1. Create topic wc_input for input stream

```
bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic wc_input
```

2. Create topic wc_output for output stream

```
bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic wc_output --config cleanup.policy=compact
```

3. Start a producer for input stream wc_input

```
bin/kafka-console-producer.sh --broker-list localhost:9092 --topic wc_input
```

4. Start a consumer for output stream wc_output

```
bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic wc_output --from-beginning --formatter kafka.tools.DefaultMessageFormatter --property print.key=true --property print.value=true --property key.deserializer=org.apache.kafka.common.serialization.StringDeserializer --property value.deserializer=org.apache.kafka.common.serialization.LongDeserializer
```

5. Start scala program which consumes wc_input, does some transformations and publish data on wc_output

```
java -jar <Your jar file name>
```

Reference:

- <https://kafka.apache.org/documentation/streams/>
- <https://kafka.apache.org/11/documentation/streams/quickstart>