ReadMe:

1. Open a terminal, locate and 'cd' into kafka, then run:

start zookeper server: bin/zookeeper-server-start.sh config/zookeeper.properties
start Kafka server: bin/kafka-server-start.sh config/server.properties

2. Open a tab in terminal window and create a topic:
bin/kafka-topics.sh --create --bootstrap-server localhost:9092 --replication-factor 1 --partitions 1 --topic <topic_name>

to see list of topics:
bin/kafka-topics.sh --list --bootstrap-server localhost:9092

3. Open a tab in terminal window and start a producer:
bin/kafka-console-producer.sh --broker-list localhost:9092 --topic <topic_name>

4. Open a tab in terminal window and start a consumer
bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic <topic_name> --from-beginning

5. Given spark is installed and path is saved in the environment variable, run:
spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.11:2.4.0 --class <className> <Path to jar file> <Kafka topic name> <consumerKey> <consumerSecret> <accessToken> <accessTokenSecret>
Eg:- spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.11:2.4.0 --class Part1 kafka-assembly-0.1.jar topic2 <consumerKey> <consumerSecret> <accessToken> <accessTokenSecret>

Get the jar file from here: https://kafka-twitter-bucket.s3.amazonaws.com/kafka-assembly-0.1.jar

6. Download and install Elasticsearch, Kibana and Logstash (configure it too).
./bin/elasticsearch (after 'cd' into Elasticsearch directory)
./bin/kibana (after 'cd' into Kibana directory)
create a file logstash-simple.conf with following
content:

```
input {
        kafka {
                bootstrap_servers => "localhost:9092"
                topics => ["YourTopic"]
                }
        }
```

```
output {
        elasticsearch {
                hosts => ["localhost:9200"]
                index => "YourTopic-index"
                }
        }
```

Then run logstash: bin/logstash -f logstash-simple.conf (after 'cd' into Logstash directory)

7. Get visualization on the http://localhost:5601