# CSCI 5800 – Big Data Systems

## Assignment #6 – Exercise 3 - Kafka

How to set up Apache Kafka

<https://tecadmin.net/how-to-install-apache-kafka-on-ubuntu-20-04/>

Before executing the following commands, you must change the active directory to the Kafka root folder:

*cd kafka*

# Start Zookeeper

Open a **new terminal** and start zookeeper using the following command

Keep this terminal open (Note: when you are done, you need to stop zookeeper by pressing ctrl+c).

Text

Description automatically generated

# Start Kafka Server

Open a **new terminal** and start the Kafka server using the following command:

*bin/kafka-server-start.sh config/server.properties*

Keep this terminal open (Note, when you are done, you need to stop the Kafka server by pressing ctrl+c).

Text

Description automatically generated

# Create a New Topic

Open a **new terminal** and create a topic named *topictest* with replication factor 1 and 1 partition using the following command:

*bin/kafka-topics.sh --create –bootstrap-server localhost:9092 --replication-factor 1 --partitions 1 --topic topictest*

To make sure that the topic is created, you can list the existing topics using the following command (yo do NOT need to open a new terminal):

*bin/kafka-topics.sh --list –bootstrap-server localhost:9092*

Text

Description automatically generated

# Create Producers

You can create a new producer using the following command:

*bin/kafka-console-producer.sh --broker-list localhost:9092 --topic topictest*

Now, the producer is ready to send messages to the broker. By default, each line will be sent as a separate message. Do not send any messages yet.

Text

Description automatically generated

# Create Consumers

Open a **new terminal**. Create a consumer and subscribe it to the *topictest*, using the following command:

*bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic topictest --from-beginning*

Text

Description automatically generated

Now, the consumer is ready to read messages from the topic. Go back to your producer and send a message. You should see the message in the consumer terminal right away.

You can create multiple consumers subscribing to the same topic. It will not help if the topic has only 1 partition. With multiple partitions, consumers will share the work by reading from different partitions, if they belong to the same work group.

# Kafka Java API

Note that before using the Kafka API you need to start the Zookeeper and the Kafka server if they are already running, you do not need to do anything. Also, you need to create a topic. Here, we assume *topictest* has been already created with one partition and replication factor has been set to 1.

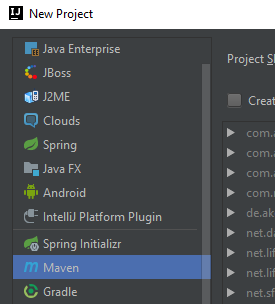
Open the IntelliJ IDE:

*cd intellij*

*cd bin*

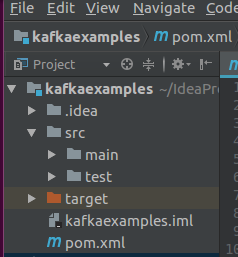
*./idea.sh*

Create a new *Maven* project.





From the project navigator, open the pom.xml. You need to add the Kafka dependencies to the project by adding the following lines before the *</project>* tag:



<dependencies>

<dependency>

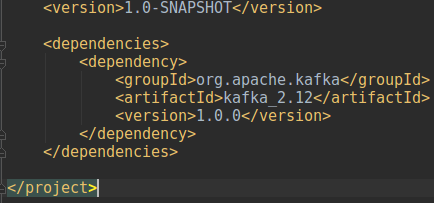
<groupId>org.apache.kafka</groupId>

<artifactId>kafka\_2.12</artifactId>

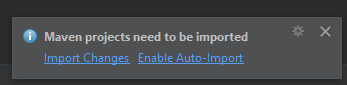
<version>1.0.0</version>

</dependency>

</dependencies>

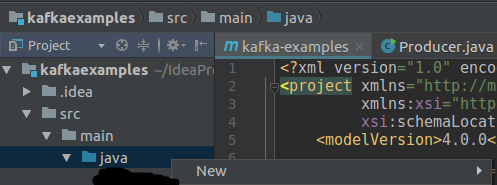


IntelliJ will ask for your permission to download and import the dependencies. Click on the Enable-Auto-Import. Wait for IntelliJ to download and import the required dependencies.



## 6.1. Create a Producer

Create a new Java class under srcmainjava and name it Producer



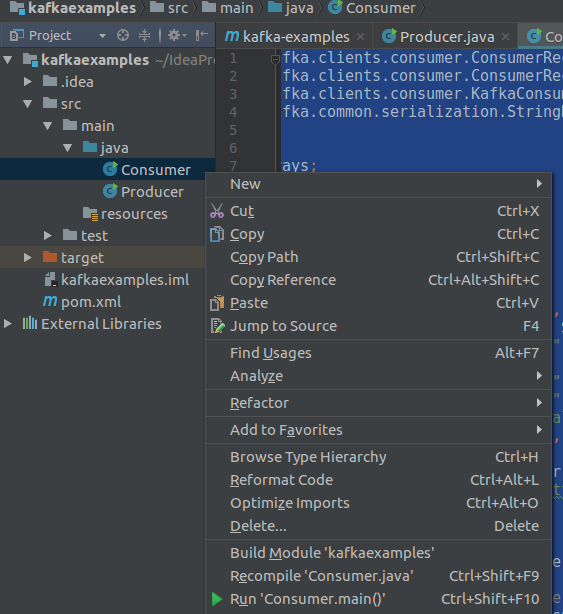
Copy the provided code from the *Producer.java* and paste it here. Go through the code and try to understand it.

Note that you can find the list of all properties and their default values here (you do not need to add anything): <https://kafka.apache.org/documentation/#producerconfigs>

Now, create a new Java class under srcmainjava and name it *Consumer*.

Copy the provided code from the Consumer.java and paste it here. Go through the code and try to understand it.

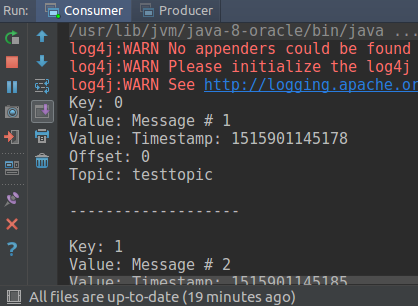
Note that you can find the list of all properties and their default values here: <https://kafka.apache.org/documentation/#consumerconfigs>



Now run the Consumer by right clicking on the Consumer class form the project navigation pane and clicking on Run ‘Consumer.main()’.

Now, the consumer is ready to read messages from the broker.

Run the producer to produce and send messages to the broker. You should be able to see messages in the consumer console shortly.



# To-Dos:

**Continue in the assignment jupyter-notebook, “Exercise 3 – Kafka” section.**