ASSIGNMENT KAFKA-2

Task 1:

Create a java program MyKafkaProducer.java that takes a file name and delimiter as input arguments.

It should read the content of file line by line.

Fields in the file are in following order

1. Kafka Topic Name

1. Key

1. value

For every line, insert the key and value to the respective Kafka broker in a fire and forget mode.

After record is sent, it should print appropriate message on screen.

Pass dataset_producer.txt as the input file and - as delimiter.

Task 2:

Modify the previous program MyKafkaProducer.java and create a new Java program KafkaProducerWithAck.java.

This should perform the same task as of KafkaProducer.java with some modification.

When passing any data to a topic, it should wait for acknowledgement. After acknowledgement is received from the broker, it should print the key and value, which has been written to a specified topic.

The application should attempt for three retries before giving any exception.

Pass dataset_producer.txt as the input file and - as delimiter.

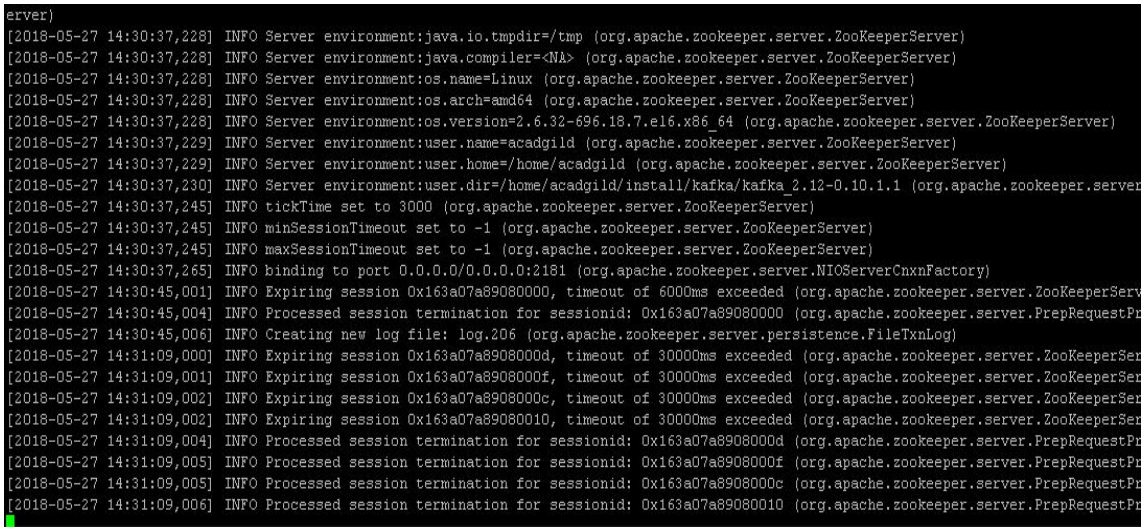
EXPLANATION: TO START WITH THE KAFKA WE NEED TO START THE FOLLOWING:

1)ZOOKEEPER SERVER: TO START ZOOKEEPER WE NEED TO GO INTO ZOOKEEPER BIN FOLDER WHERE THE NECESSARY FILES ARE PRESENT FOR STARTING THE ZOOKEEPER. SO WE NEED TO CHANGE THE DIRECTORY TO "cd $ZOOKEEPER_HOME". THEN GO TO BIN FOLDER "cd bin". THEN ENTER BELOW:-

CODE: zkServer.sh start

THIS WILL START THE ZOOKEEPER SERVICES. THEN WE NEED TO START THE ZOOKEEPER

CLIENT FOR THAT WE USE BELOW CODE.



CODE: cd $KAFKA_HOME
./bin/zookeeper-server-start.sh ./config/zookeeper.properties

2) KAFKA SERVER "BROKER": NOW WE NEED TO CHANGE DIRECTORY TO KAFKA BY "cd $KAFKA_HOME". THEN ENTER BELOW CODE:

CODE: ./bin/kafka-server-start.sh ./config/server.properties

```
[acadgild@localhost ~]$ cd $KAFKA_HOME
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost kafka_2.12-0.10.1.1]$ ./bin/kafka-server-start.sh ./config/server.properties
[2018-05-27 14:32:21,604] INFO KafkaConfig values:
        advertised.host.name = null
        advertised.listeners = null
        advertised.port = null
        authorizer.class.name =
        auto.create.topics.enable = true
        auto.leader.rebalance.enable = true
        background.threads = 10
        broker.id = 0
        broker.id.generation.enable = true
        broker.rack = null
        compression.type = producer
        connections.max.idle.ms = 600000
        controlled.shutdown.enable = true
        controlled.shutdown.max.retries = 3
        controlled.shutdown.retry.backoff.ms = 5000
        controller.socket.timeout.ms = 30000
        default.replication.factor = 1
        delete.topic.enable = false
        fetch.purgatory.purge.interval.requests = 1000
        group.max.session.timeout.ms = 300000
        group.min.session.timeout.ms = 6000
        host.name =
        inter.broker.protocol.version = 0.10.1-IV2
        leader.imbalance.check.interval.seconds = 300
        leader.imbalance.per.broker.percentage = 10
        listeners = null
        log.cleaner.backoff.ms = 15000
        log.cleaner.dedupe.buffer.size = 134217728
        log.cleaner.delete.retention.ms = 86400000
        log.cleaner.enable = true
        log.cleaner.io.buffer.load.factor = 0.9
```

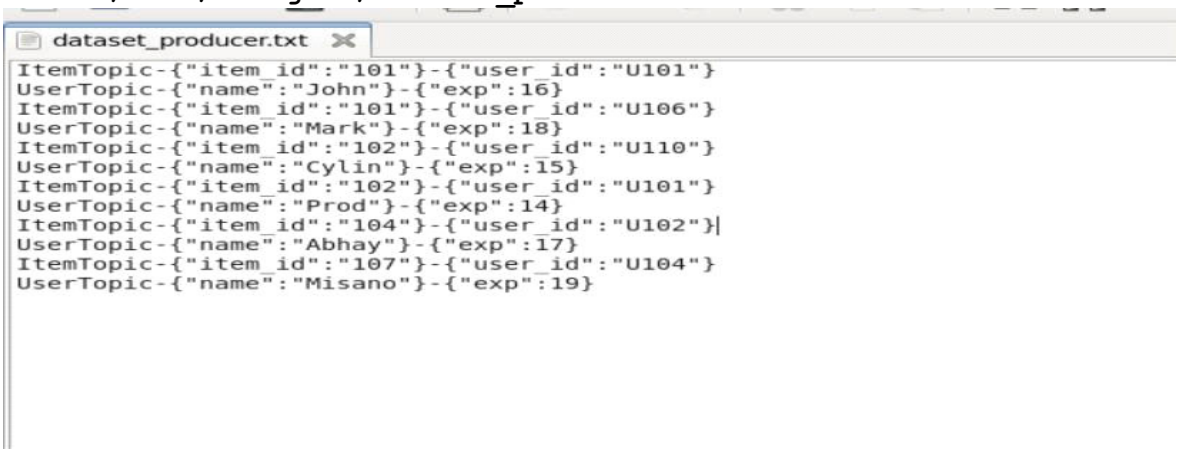
```
        socket.receive.buffer.bytes = 102400
        socket.request.max.bytes = 104857600
        socket.send.buffer.bytes = 102400
        ssl.cipher.suites = null
        ssl.client.auth = none
        ssl.enabled.protocols = [TLSv1.2, TLSv1.1, TLSv1]
        ssl.endpoint.identification.algorithm = null
        ssl.key.password = null
        ssl.keymanager.algorithm = SunX509
        ssl.keystore.location = null
        ssl.keystore.password = null
        ssl.keystore.type = JKS
        ssl.protocol = TLS
        ssl.provider = null
        ssl.secure.random.implementation = null
        ssl.trustmanager.algorithm = PKIX
        ssl.truststore.location = null
        ssl.truststore.password = null
        ssl.truststore.type = JKS
        unclean.leader.election.enable = true
        zookeeper.connect = localhost:2181
        zookeeper.connection.timeout.ms = 6000
        zookeeper.session.timeout.ms = 6000
        zookeeper.set.acl = false
        zookeeper.sync.time.ms = 2000
 (kafka.server.KafkaConfig)
[2018-05-27 14:32:21,730] INFO starting (kafka.server.KafkaServer)
[2018-05-27 14:32:21,752] INFO [ThrottledRequestReaper-Fetch], Starting  (kafka.server.ClientQuotaManager$Throttled
[2018-05-27 14:32:21,754] INFO [ThrottledRequestReaper-Produce], Starting  (kafka.server.ClientQuotaManager$Throttl
[2018-05-27 14:32:21,773] INFO Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2018-05-27 14:32:21,796] INFO Starting ZkClient event thread. (org.IOItec.zkclient.ZkEventThread)
[2018-05-27 14:32:21,803] INFO Client environment:zookeeper.version=3.4.8--1, built on 02/06/2016 03:18 GMT (org.ap
[2018-05-27 14:32:21,803] INFO Client environment:host.name=localhost (org.apache.zookeeper.ZooKeeper)
[2018-05-27 14:32:21,803] INFO Client environment:java.version=1.8.0_151 (org.apache.zookeeper.ZooKeeper)
[2018-05-27 14:32:21,803] INFO Client environment:java.vendor=Oracle Corporation (org.apache.zookeeper.ZooKeeper)
[2018-05-27 14:32:21,803] INFO Client environment:java.home=/usr/java/jdk1.8.0_151/jre (org.apache.zookeeper.ZooKee
[2018-05-27 14:32:21,803] INFO Client environment:java.class.path=:/home/acadgild/install/kafka/kafka_2.12-0.10.1.1
me/acadgild/install/kafka/kafka_2.12-0.10.1.1/bin/../libs/argparse4j-0.5.0.jar:/home/acadgild/install/kafka/kafka_2
me/acadgild/install/kafka/kafka_2.12-0.10.1.1/bin/../libs/connect-file-0.10.1.1.jar:/home/acadgild/install/kafka/ka
```

```
[2018-05-27 14:32:21,860] INFO Session establishment complete on server localhost/127.0.0.1:2181, sessionid = 0x163a0d3fc700000, negoti
ookeeper.ClientCnxn)
[2018-05-27 14:32:21,862] INFO zookeeper state changed (SyncConnected) (org.IOItec.zkclient.ZkClient)
[2018-05-27 14:32:22,144] INFO Cluster ID = pv7NEl2tSqiFD2G93O_wxQ (kafka.server.KafkaServer)
[2018-05-27 14:32:22,298] INFO Loading logs. (kafka.log.LogManager)
[2018-05-27 14:32:22,380] INFO Completed load of log UserTopic-0 with 1 log segments and log end offset 6 in 51 ms (kafka.log.Log)
[2018-05-27 14:32:22,395] INFO Completed load of log sample_topic-0 with 1 log segments and log end offset 2 in 2 ms (kafka.log.Log)
[2018-05-27 14:32:22,405] INFO Completed load of log KeyedTopic-0 with 1 log segments and log end offset 10 in 3 ms (kafka.log.Log)
[2018-05-27 14:32:22,416] INFO Completed load of log ItemTopic-0 with 1 log segments and log end offset 6 in 5 ms (kafka.log.Log)
[2018-05-27 14:32:22,428] INFO Completed load of log TestTopic1-0 with 1 log segments and log end offset 5 in 3 ms (kafka.log.Log)
[2018-05-27 14:32:22,439] INFO Completed load of log TestTopic-0 with 1 log segments and log end offset 5 in 3 ms (kafka.log.Log)
[2018-05-27 14:32:22,448] INFO Logs loading complete in 150 ms. (kafka.log.LogManager)
[2018-05-27 14:32:22,509] INFO Starting log cleanup with a period of 300000 ms. (kafka.log.LogManager)
[2018-05-27 14:32:22,515] INFO Starting log flusher with a default period of 9223372036854775807 ms. (kafka.log.LogManager)
[2018-05-27 14:32:22,634] INFO Awaiting socket connections on 0.0.0.0:9092. (kafka.network.Acceptor)
[2018-05-27 14:32:22,639] INFO [Socket Server on Broker 0], Started 1 acceptor threads (kafka.network.SocketServer)
[2018-05-27 14:32:22,685] INFO [ExpirationReaper-0], Starting  (kafka.server.DelayedOperationPurgatory$ExpiredOperationReaper)
[2018-05-27 14:32:22,687] INFO [ExpirationReaper-0], Starting  (kafka.server.DelayedOperationPurgatory$ExpiredOperationReaper)
[2018-05-27 14:32:22,781] INFO Creating /controller (is it secure? false) (kafka.utils.ZKCheckedEphemeral)
[2018-05-27 14:32:22,800] INFO Result of znode creation is: OK (kafka.utils.ZKCheckedEphemeral)
[2018-05-27 14:32:22,804] INFO 0 successfully elected as leader (kafka.server.ZookeeperLeaderElector)
[2018-05-27 14:32:23,304] INFO New leader is 0 (kafka.server.ZookeeperLeaderElector$LeaderChangeListener)
[2018-05-27 14:32:23,314] INFO [ExpirationReaper-0], Starting  (kafka.server.DelayedOperationPurgatory$ExpiredOperationReaper)
[2018-05-27 14:32:23,325] INFO [ExpirationReaper-0], Starting  (kafka.server.DelayedOperationPurgatory$ExpiredOperationReaper)
[2018-05-27 14:32:23,334] INFO [ExpirationReaper-0], Starting  (kafka.server.DelayedOperationPurgatory$ExpiredOperationReaper)
[2018-05-27 14:32:23,350] INFO [GroupCoordinator 0]: Starting up. (kafka.coordinator.GroupCoordinator)
[2018-05-27 14:32:23,353] INFO [GroupCoordinator 0]: Startup complete. (kafka.coordinator.GroupCoordinator)
[2018-05-27 14:32:23,365] INFO [Group Metadata Manager on Broker 0]: Removed 0 expired offsets in 4 milliseconds. (kafka.coordinator.Gr
[2018-05-27 14:32:23,403] INFO Will not load MX4J, mx4j-tools.jar is not in the classpath (kafka.utils.Mx4jLoader$)
[2018-05-27 14:32:23,532] INFO Creating /brokers/ids/0 (is it secure? false) (kafka.utils.ZKCheckedEphemeral)
[2018-05-27 14:32:23,550] INFO Result of znode creation is: OK (kafka.utils.ZKCheckedEphemeral)
[2018-05-27 14:32:23,552] INFO Registered broker 0 at path /brokers/ids/0 with addresses: PLAINTEXT -> EndPoint(localhost,9092,PLAINTEX
[2018-05-27 14:32:23,582] INFO Kafka version : 0.10.1.1 (org.apache.kafka.common.utils.AppInfoParser)
[2018-05-27 14:32:23,582] INFO Kafka commitId : f10ef2720b03b247 (org.apache.kafka.common.utils.AppInfoParser)
[2018-05-27 14:32:23,587] INFO [Kafka Server 0], started (kafka.server.KafkaServer)
[2018-05-27 14:32:24,003] INFO [ReplicaFetcherManager on broker 0] Removed fetcher for partitions UserTopic-0,ItemTopic-0,TestTopic-0,K
pic1-0 (kafka.server.ReplicaFetcherManager)
[2018-05-27 14:32:24,098] INFO [ReplicaFetcherManager on broker 0] Removed fetcher for partitions UserTopic-0,ItemTopic-0,TestTopic-0,K
pic1-0 (kafka.server.ReplicaFetcherManager)
```

3) WE CAN CHECK THE ZOOKEEPER AND KAFKA SERVER RUNNING BY ENTERING:-

CODE: jps

EXPLANATION: NOW TO PERFORM THE TASK FIRST CREATED A MAVEN PROJECT FOR ADDING THE NECESSARY DEPENDENCIES FOR THE KAFKA TO WORK IN ECLIPSE ENVIRONMENT. THEN WE CREATED A JAVA PROJECT WITH THE JAVA CLASS AS "MyKafkaProducer". BELOW IS THE DATASET FOR PRODUCER TO WORK ON WHICH IS SAVED IN THE:- **"PATH:/home/acadgild/dataset_producer"**



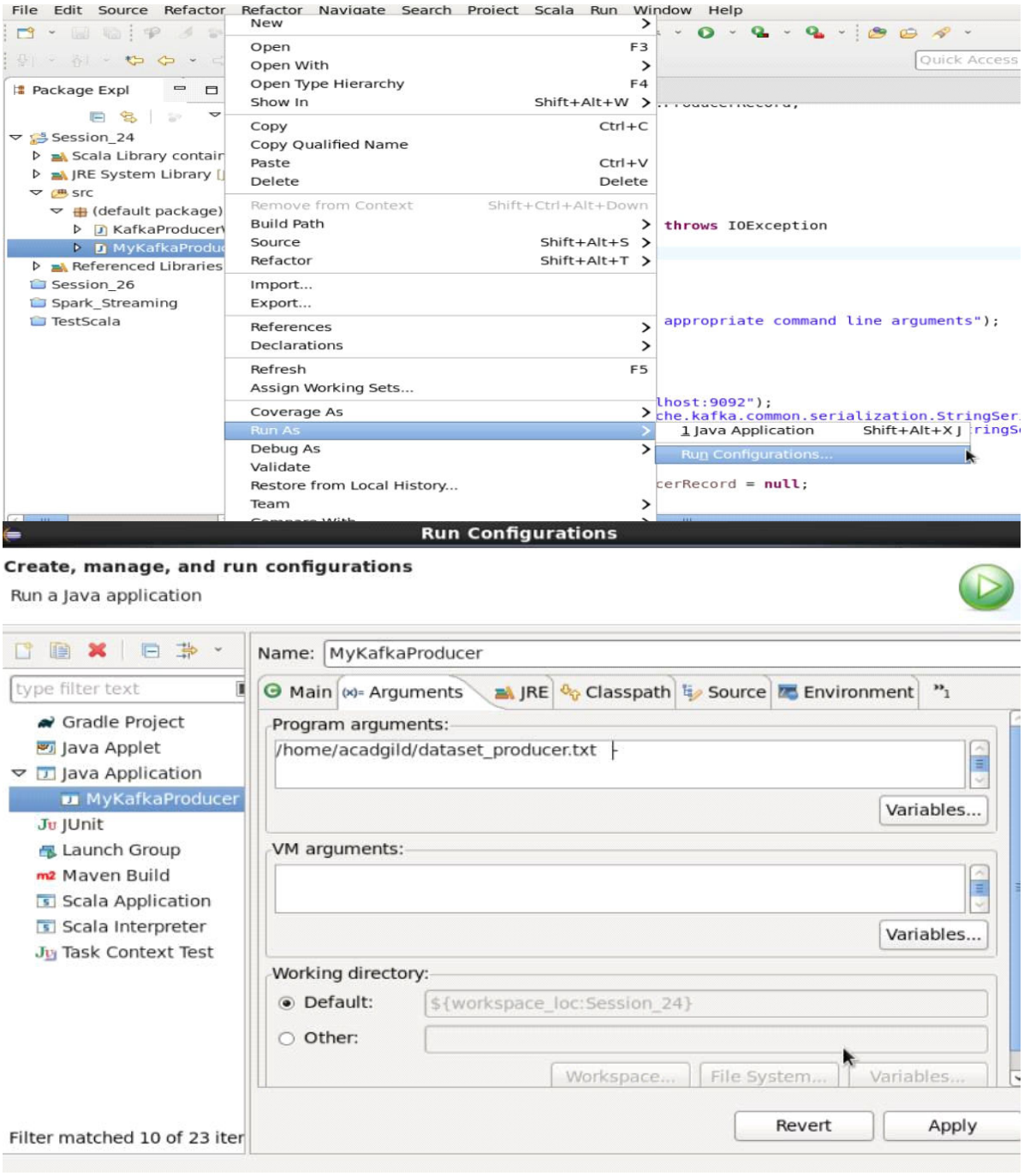
SOLUTION FOR TASK 1:

JAVA CODE FOR MyKafkaProducer

```
package kafkaProjnew;
import org.apache.kafka.clients.producer.KafkaProducer;
import org.apache.kafka.clients.producer.ProducerRecord;

import java.io.BufferedReader;
import java.io.FileReader;
import java.io.IOException;
import java.util.Properties;

public class MyKafkaProducer
{
 public static void main(String[] args) throws IOException
 {
   if (args.length != 2)
   {
      System.out.println("Please provide appropriate command line
argument");
      System.exit(-1);
   }
   Properties props = new Properties();

      props.put("bootstrap.servers", "localhost:9092");
      props.put("key.serializer",
"org.apache.kafka.common.serialization.StringSerializer");
```

```
     props.put("value.serializer",
"org.apache.kafka.common.serialization.StringSerializer");

    KafkaProducer<String, String>  producer = new KafkaProducer<>(props);

    ProducerRecord<String, String> producerRecord = null;


    String fileName = args[0];
    String delimiter = args[1];

    try(BufferedReader br = new BufferedReader(new FileReader(fileName)))
    {
     for(String line; (line = br.readLine()) != null; )
     {
          String[] tempArray = line.split(delimiter);

          String topic = tempArray[0];
          String key = tempArray[1];
          String value = tempArray[2];

     producerRecord = new ProducerRecord<String, String>(topic, key,
value);
     producer.send(producerRecord);

     System.out.printf("Record sent to topic:%s. Key:%s\n, Value:%s\n",
topic ,key, value);
      }
    }
          producer.close();
  }
}
```

EXPLANATION: NOW GO TO **"MyKafkaProducer"** AND THEN RIGHT CLICK ON IT THEN CLICK ON THE OPTION "RUN AS" AND GO TO "RUN CONFIGURATION". THEN GO TO ARGUMENTS OPTION AND GIVE THE FILE PATH AND PASS THE INPUT DATA FILE WITH "-" AS SHOWN BELOW. THEN CLICK "APPLY".



BELOW IS THE CONSOLE OUTPUT:

```
Record sent to topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U101"}
Record sent to topic:UserTopic. Key:{"name":"John"}, Value:{"exp":16}
Record sent to topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U106"}
Record sent to topic:UserTopic. Key:{"name":"Mark"}, Value:{"exp":18}
Record sent to topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U110"}
Record sent to topic:UserTopic. Key:{"name":"Cylin"}, Value:{"exp":15}
Record sent to topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U101"}
Record sent to topic:UserTopic. Key:{"name":"Prod"}, Value:{"exp":14}
Record sent to topic:ItemTopic. Key:{"item_id":"104"}, Value:{"user_id":"U102"}
Record sent to topic:UserTopic. Key:{"name":"Abhay"}, Value:{"exp":17}
Record sent to topic:ItemTopic. Key:{"item_id":"107"}, Value:{"user_id":"U104"}
Record sent to topic:UserTopic. Key:{"name":"Misano"}, Value:{"exp":19}
```

EXPLANATION:RUNNING A KAFKA CONSOLE CONSUMER FOR TOPIC "ItemTopic" TO

PRINT MESSAGE IN KEY VALUE FORMAT AS SHOWN BELOW. FOR TOPIC "UserTopic" DID THE SAME.

CODE: ./bin/kafka-console-consumer.sh --topic ItemTopic --from-beginning --zookeeper localhost:2181 --property print.key=true

CODE: ./bin/kafka-console-consumer.sh --topic UserTopic --from-beginning --zookeeper localhost:2181 --property print.key=true

```
[acadgild@localhost kafka_2.12-0.10.1.1]$ ./bin/kafka-console-consumer.sh --topic ItemTopic --from-beginning \
> --zookeeper localhost:2181 \
> --property print.key=true
Using the ConsoleConsumer with old consumer is deprecated and will be removed in a future major release. Consider
nstead of [zookeeper].
{"item_id":"101"}       {"user_id":"U101"}
{"item_id":"101"}       {"user_id":"U106"}
{"item_id":"102"}       {"user_id":"U110"}
{"item_id":"102"}       {"user_id":"U101"}
{"item_id":"104"}       {"user_id":"U102"}
{"item_id":"107"}       {"user_id":"U104"}
```

```
[acadgild@localhost kafka_2.12-0.10.1.1]$ ./bin/kafka-console-consumer.sh --topic UserTopic --from-beginning \
> --zookeeper localhost:2181 \
> --property print.key=true
Using the ConsoleConsumer with old consumer is deprecated and will be removed in a future major release. Conside
nstead of [zookeeper].
{"name":"John"} {"exp":16}
{"name":"Mark"} {"exp":18}
{"name":"Cylin"}        {"exp":15}
{"name":"Prod"} {"exp":14}
{"name":"Abhay"}        {"exp":17}
{"name":"Misano"}       {"exp":19}
```

EXPLANATION: WE DID THE SAME STEPS FOR THE JAVA PROGRAM "KafkaProducerWithAck"

SOLUTION FOR TASK 2:

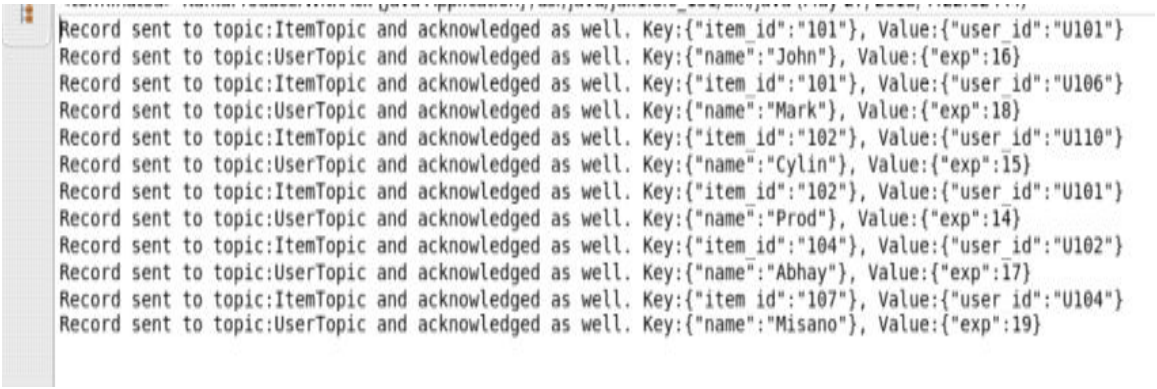
JAVA CODE FOR KafkaProducerWithAck

```
package kafkaProjnew;
import org.apache.kafka.clients.producer.KafkaProducer;
import org.apache.kafka.clients.producer.ProducerRecord;

import java.io.BufferedReader;
import java.io.FileReader;
import java.io.IOException;
import java.util.Properties;
import java.util.concurrent.ExecutionException;


public class KafkaProducerWithAck
{
  public static void main(String[] args) throws IOException,
```

```
InterruptedException, ExecutionException
  {
      if (args.length != 2)
      {
        System.out.println("Please provide appropriate command line
argument");
        System.exit(-1);
      }
      Properties props = new Properties();

      props.put("bootstrap.servers", "localhost:9092");
      props.put("acks",  "all");
      props.put("retries",  "all");

      props.put("key.serializer","org.apache.kafka.common.serialization
.StringSerializer");

     props.put("value.serializer",
"org.apache.kafka.common.serialization.StringSerializer");

      KafkaProducer<String, String>  producer = new
KafkaProducer<>(props);

      ProducerRecord<String, String> producerRecord = null;

      String fileName = args[0];
      String delimiter = args[1];

      try(BufferedReader br = new BufferedReader(new
FileReader(fileName)))
      {
       for(String line; (line = br.readLine()) != null; )
       {
            String[] tempArray = line.split(delimiter);
            String topic = tempArray[0];
            String key = tempArray[1];
            String value = tempArray[2];

        producerRecord = new ProducerRecord<String, String>(topic, key,
value);

        producer.send(producerRecord);

        System.out.printf("Record sent to topic:%s. and Acknowledged as
well. Key:%s\n, Value:%s\n",  topic ,key, value);
       }
      }

      producer.close();

  }

}
```

OUTPUT IN CONSOLE:

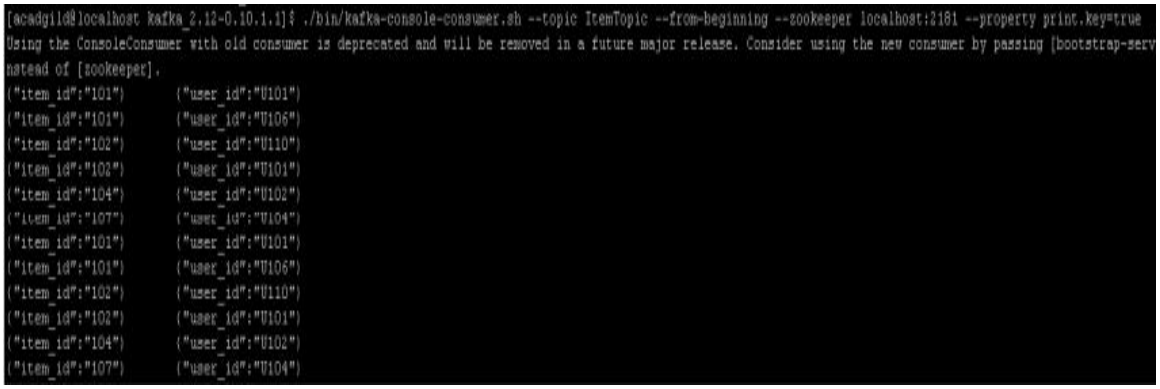
Below screenshot shows the producer program acknowledge first and after receiving acknowledgment from the broker, it print the message **"Acknowledged as well"** and key and value, which has been written to a specified topic



EXPLANATION: RUNNING A KAFKA CONSOLE CONSUMER FOR TOPIC "ItemTopic" TO PRINT MESSAGE IN KEY VALUE FORMAT AS SHOWN BELOW. FOR TOPIC "UserTopic" DID THE SAME.

CODE: ./bin/kafka-console-consumer.sh --topic ItemTopic --from-beginning --zookeeper localhost:2181 --property print.key=true

CODE: ./bin/kafka-console-consumer.sh --topic UserTopic --from-beginning --zookeeper localhost:2181 --property print.key=true



```
[acadgild@localhost kafka_2.12-0.10.1.1]$ ./bin/kafka-console-consumer.sh --topic UserTopic --from-beginning --zookeeper localhost:2181 --property print.key=true
Using the ConsoleConsumer with old consumer is deprecated and will be removed in a future major release. Consider using the new consumer by passing [bootstrap-ser
nstead of [zookeeper].
{"name":"John"} {"exp":16}
{"name":"Mark"} {"exp":18}
{"name":"Cylin"}        {"exp":15}
{"name":"Prod"} {"exp":14}
{"name":"Abhay"}        {"exp":17}
{"name":"Misano"}       {"exp":19}
{"name":"John"} {"exp":16}
{"name":"Mark"} {"exp":18}
{"name":"Cylin"}        {"exp":15}
{"name":"Prod"} {"exp":14}
{"name":"Abhay"}        {"exp":17}
{"name":"Misano"}       {"exp":19}
```