FINAL PROJECT BIG DATA TECHNOLOGY

**GROUP 20 :**

Tuan Anh Nguyen (614636)

Alagie Nget (614251)

Biniam Abrha Nigusse (614623)

1. **Step 1: setup Kafka**

*Download Kafka:* [*https://archive.apache.org/dist/kafka/2.4.1/kafka\_2.11-2.4.1.tgz*](https://archive.apache.org/dist/kafka/2.4.1/kafka_2.11-2.4.1.tgz)

* Open Cloudera VM & copy file downloaded and extract to any directory.

For example, I extracted at

*/home/cloudera/cs523/kafka\_2.11-2.4.1*

* Go to bin kafka folder:

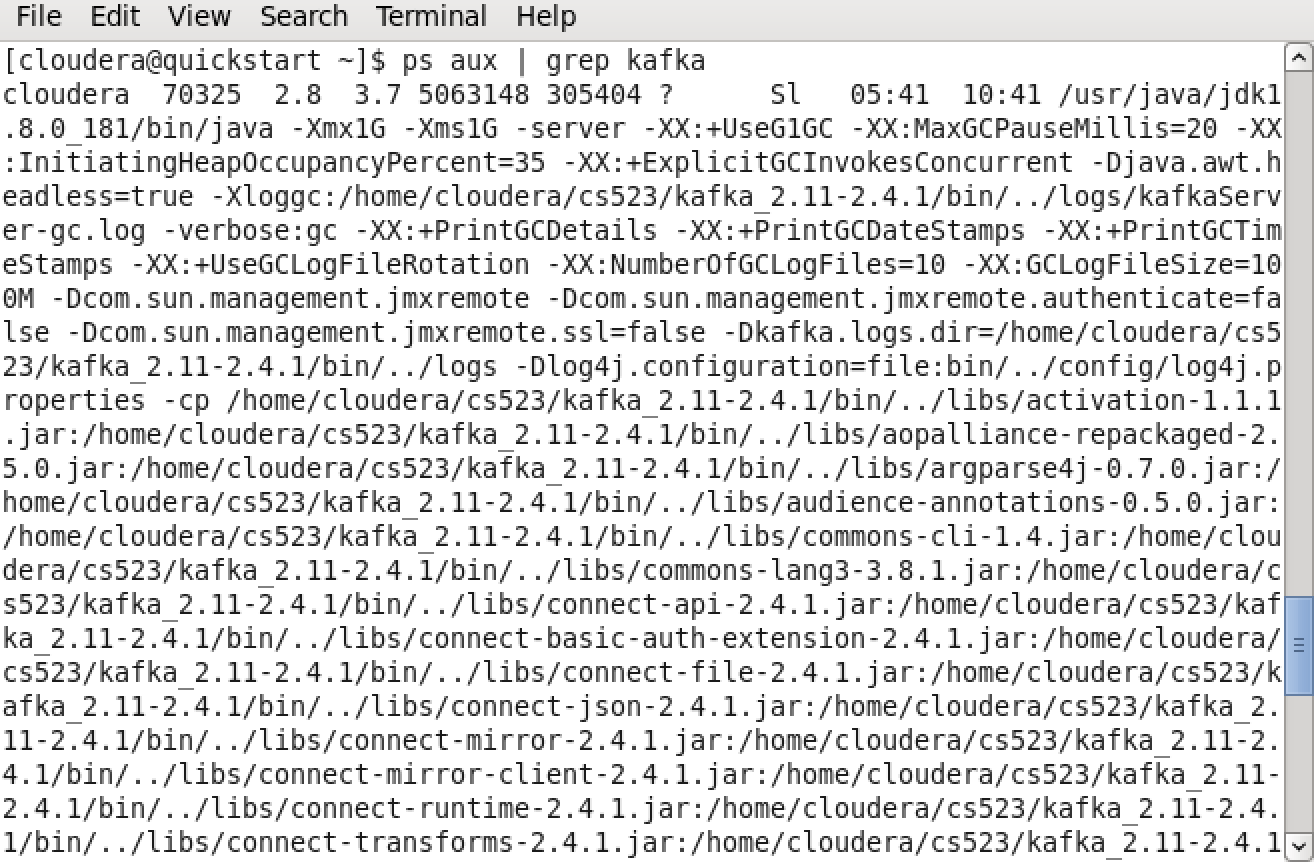
*cd '/home/cloudera/cs523/kafka\_2.11-2.4.1/bin'*

* Start Kafka:

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

* Check Kafka start or not:

*ps aux | grep kafka*



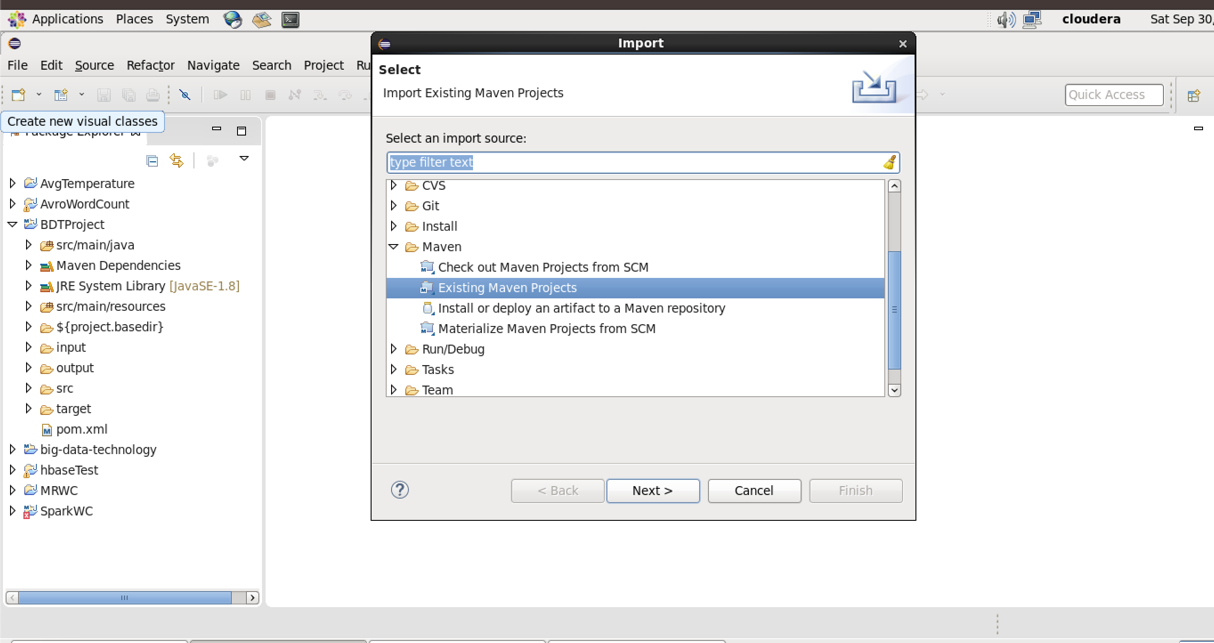
1. **Start HBase:**

*sudo service hbase-master start;*

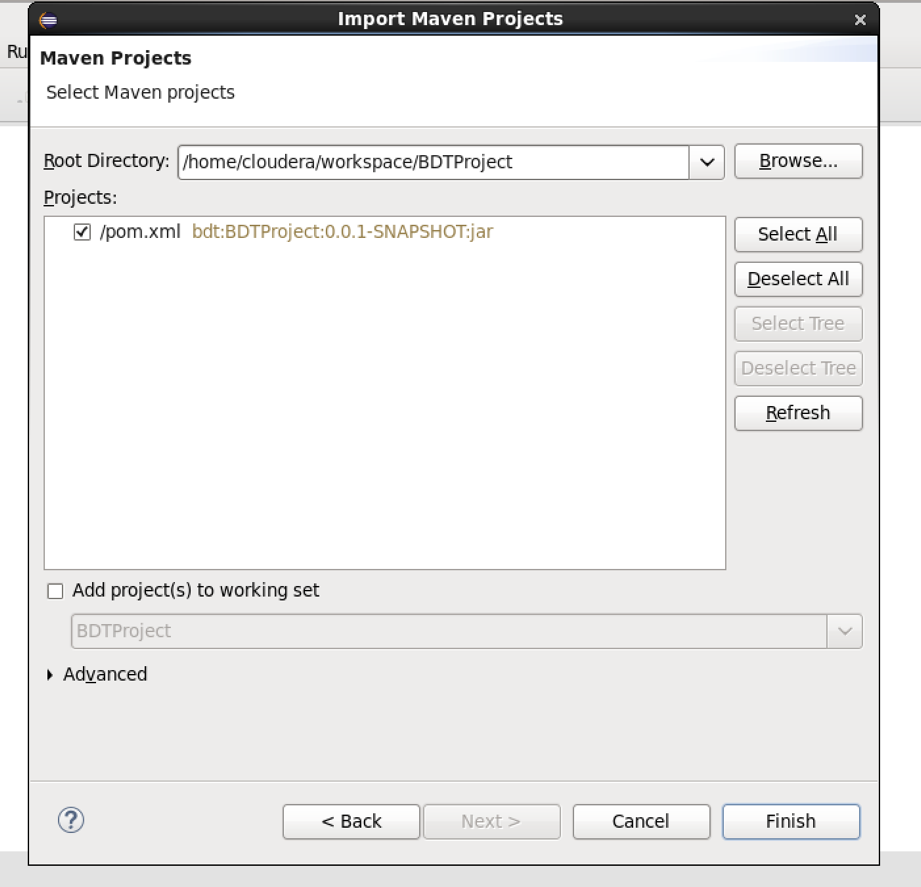
*sudo service hbase-regionserver start;*

1. **Import and run project:**

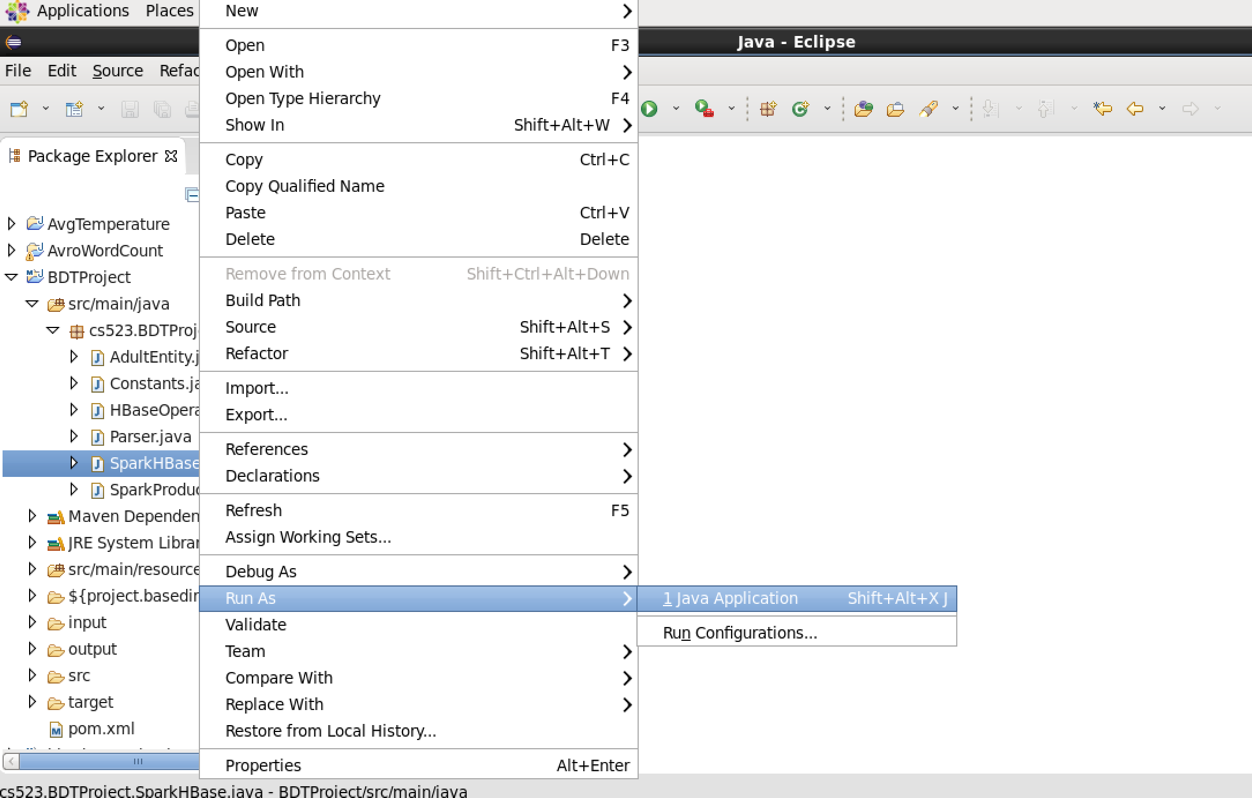
Import Maven project



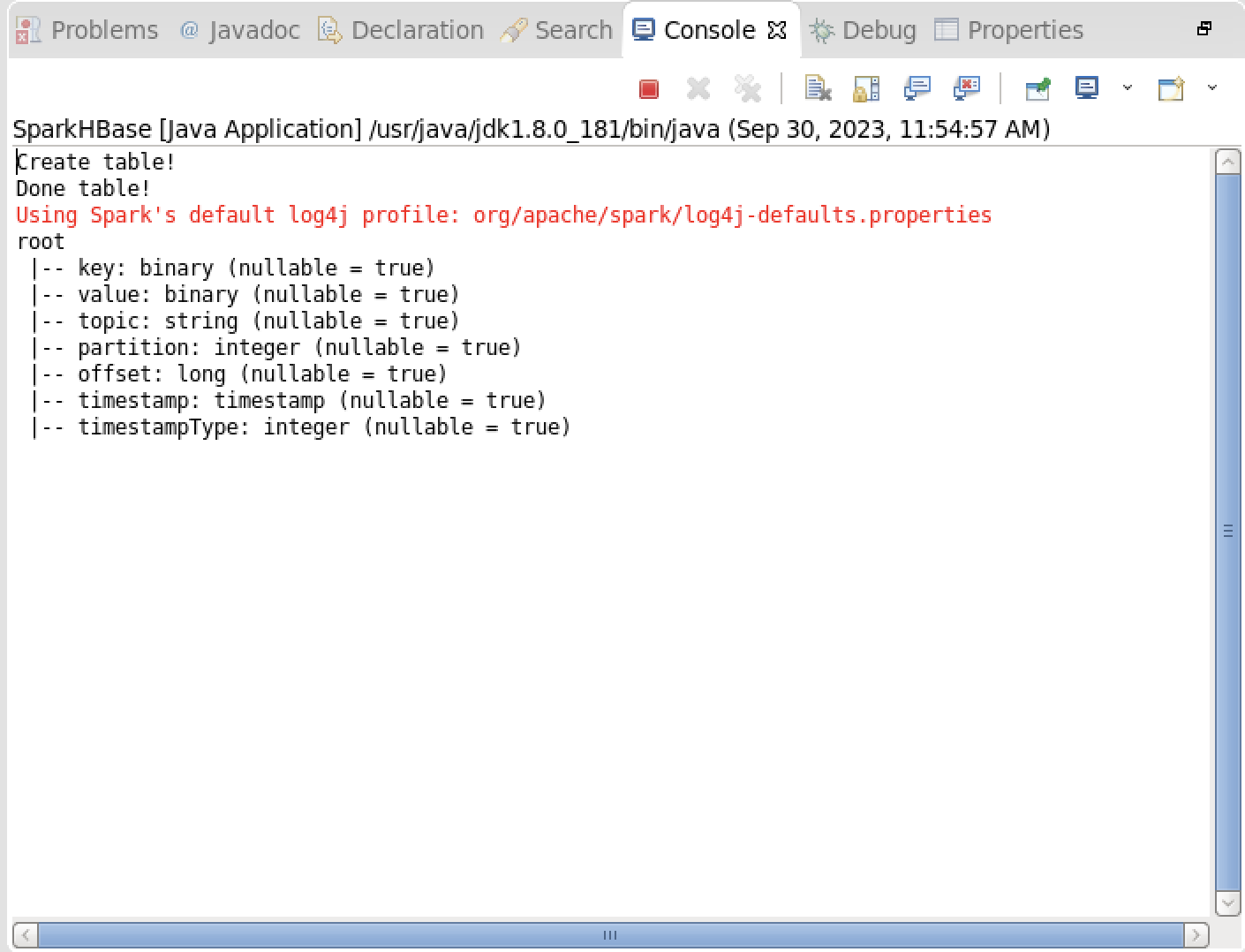
Choose project “BDTProject”



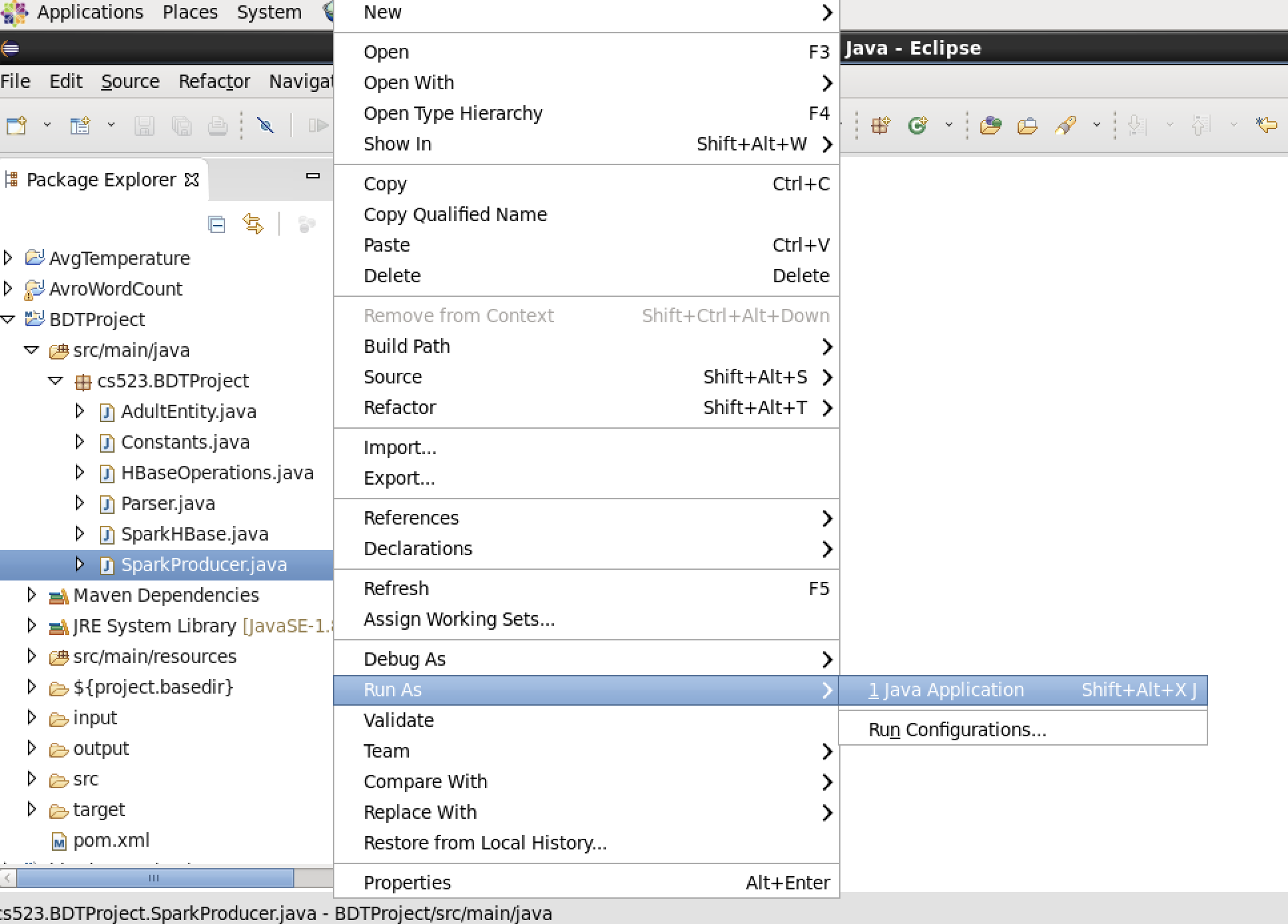
Run SparkHBase.java first



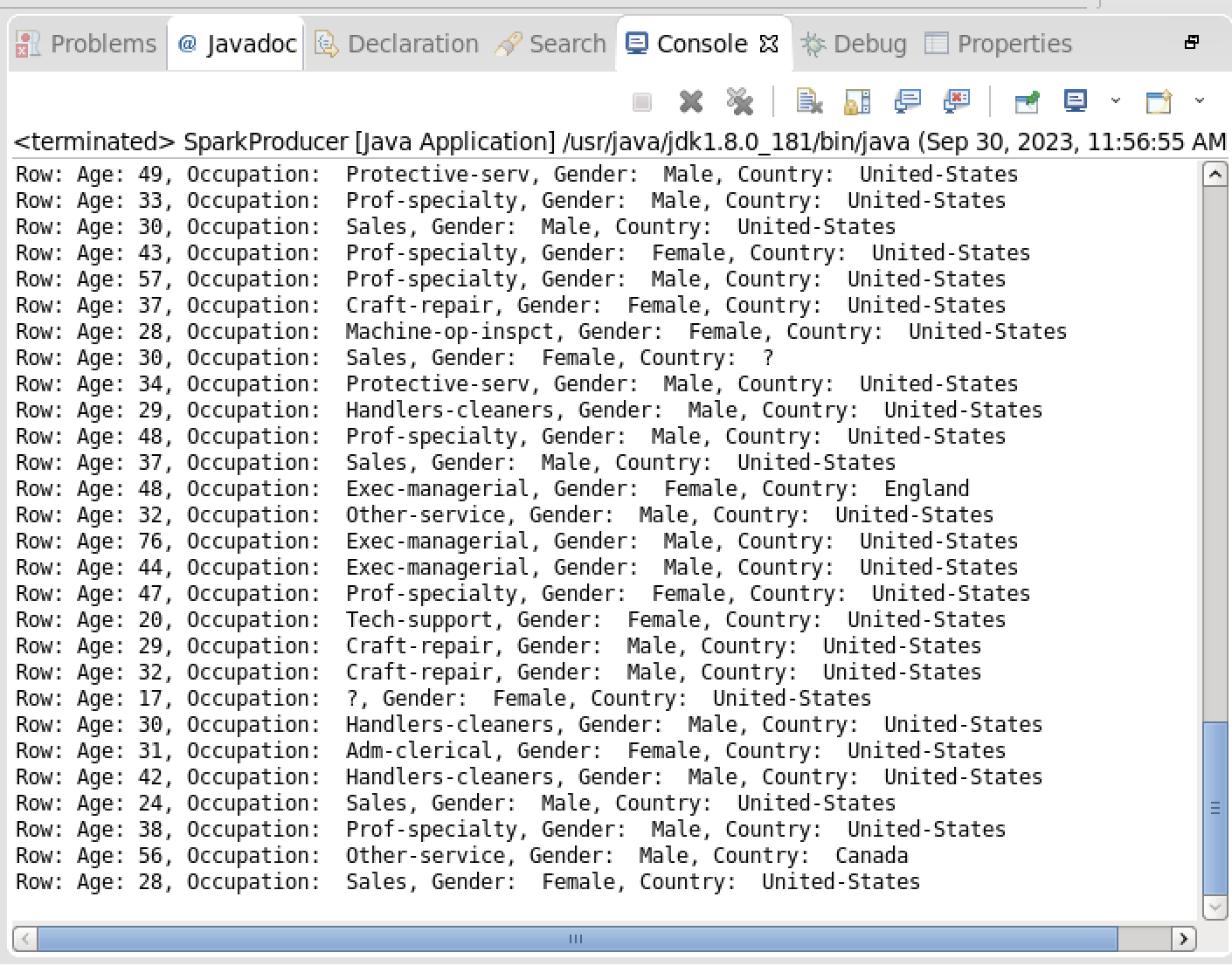
Waiting. If it like this, ready to listen message and add to HBase



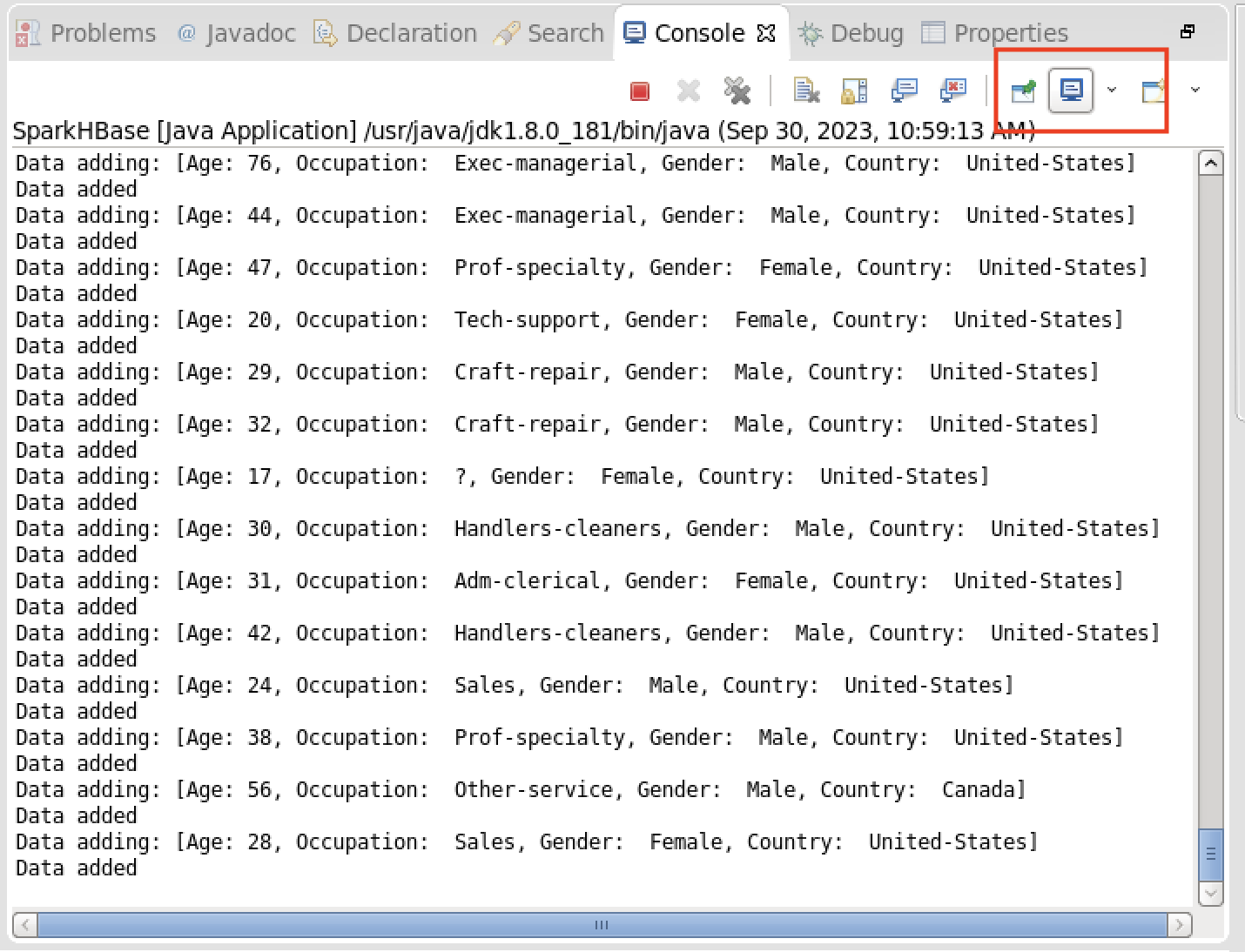
Continue run SparkProducer.java



Look at on Console, will see like this



Click “Display Selected Console”, you will data added like this



1. **Result**

Check on HBase, you will see table “Adult” have data



Check on browser

