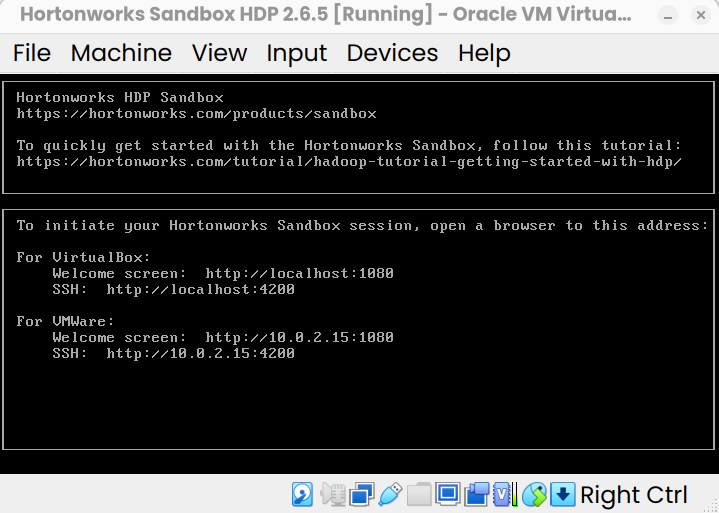
LAB 4

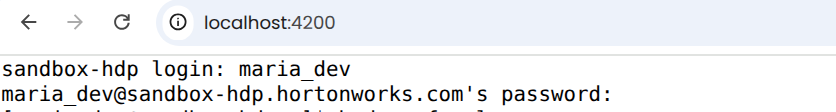
**Aim:** Run a java program (WordCount) on the concept of MapReduce

**Procedure:**

1. Start the HortonWorks it shows the IP address of machine.



1. Navigate to [http://localhost:](http://localhost:1080/)4200
2. Use the username and password: maria\_dev to login.

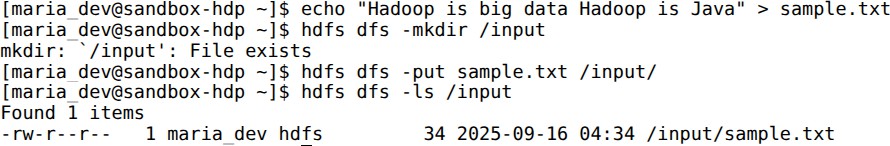


1. Execute the following commands one-by-one:

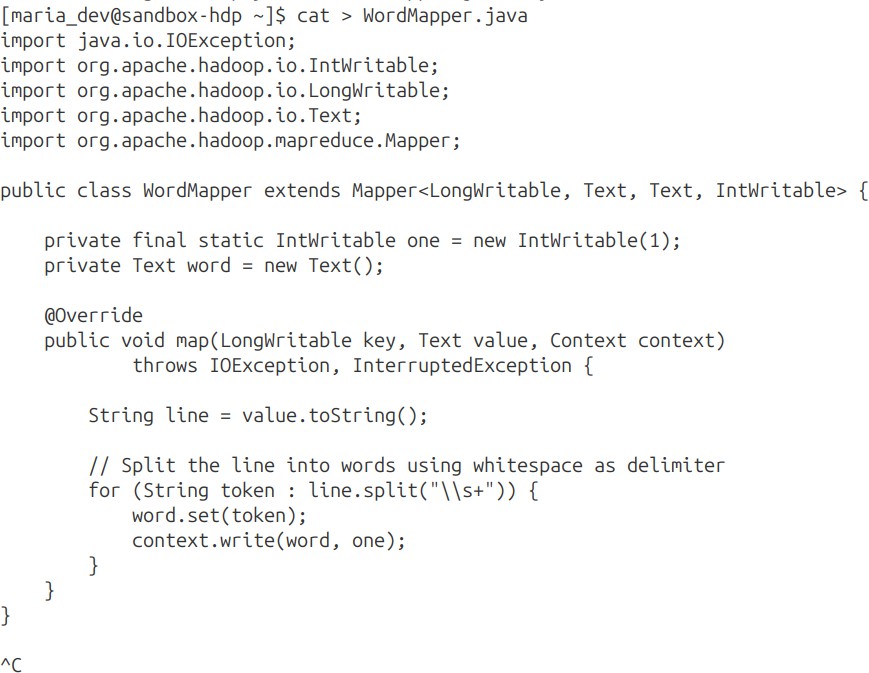
# Prepare Input Data

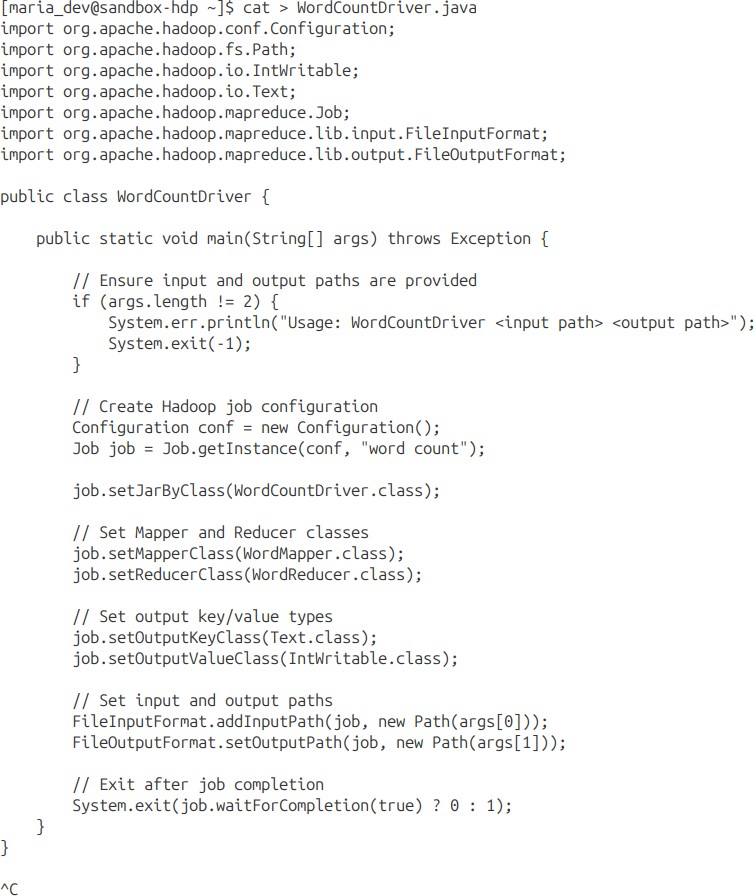
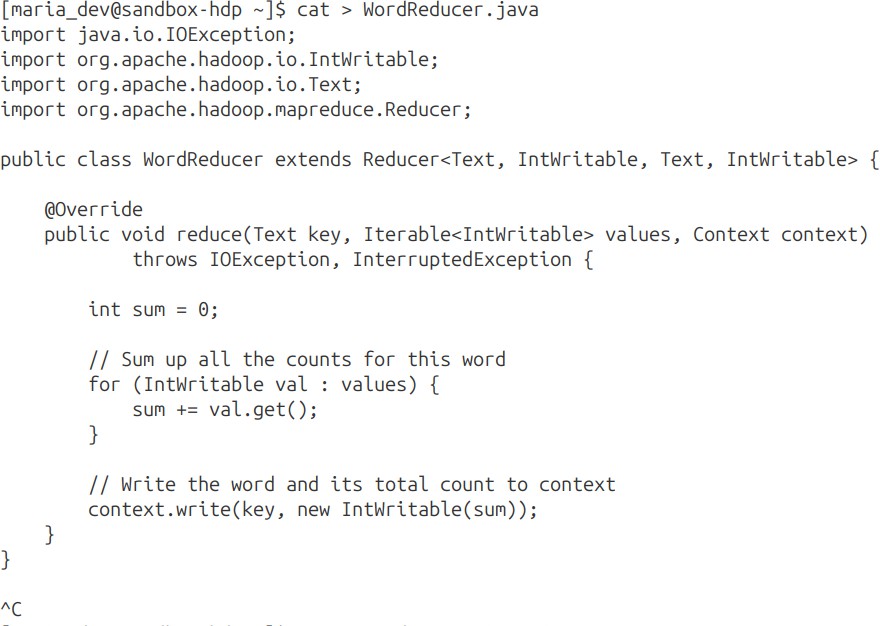
echo "Hadoop is big data Hadoop is Java" > sample.txt hdfs dfs -mkdir /input

hdfs dfs -put sample.txt /input/ hdfs dfs -ls /input



# Create Java Files



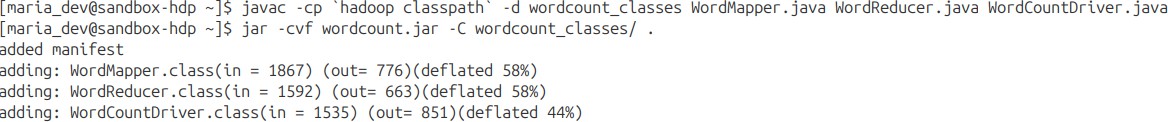
****

* + **Compile Java Files**

javac -cp `hadoop classpath` -d wordcount\_classes WordMapper.java WordReducer.java WordCountDriver.java

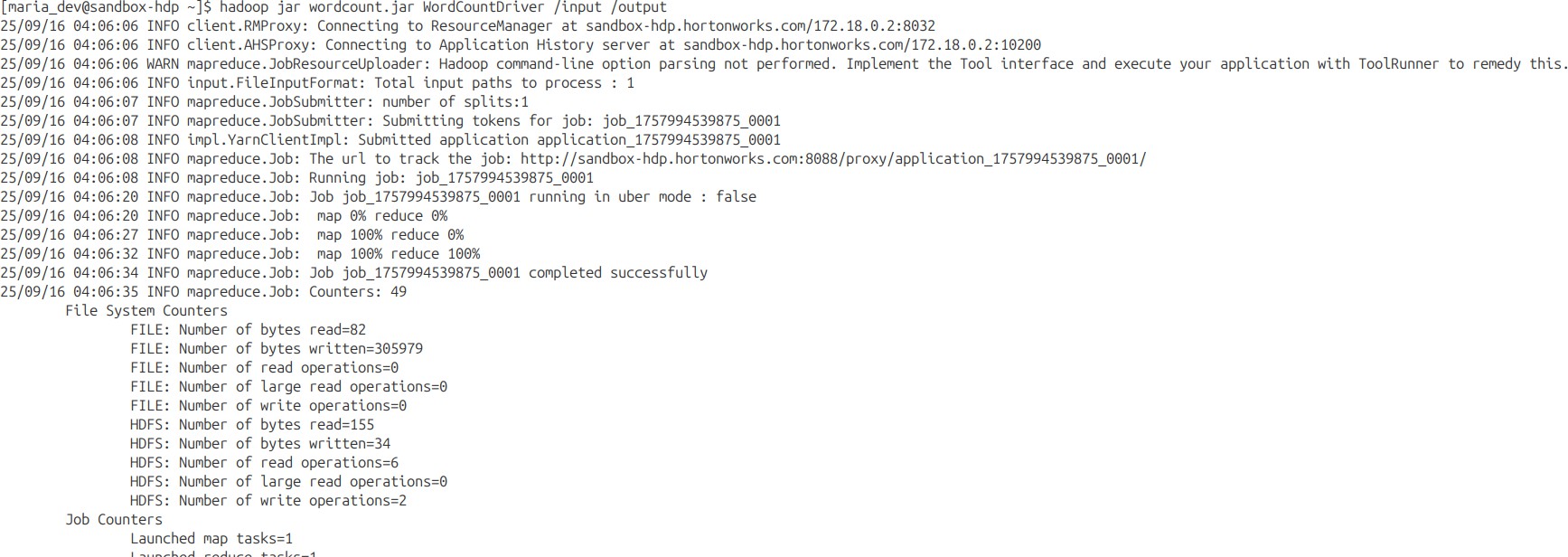
# Create JAR

jar -cvf wordcount.jar -C wordcount\_classes/ .



# Run MapReduceJob

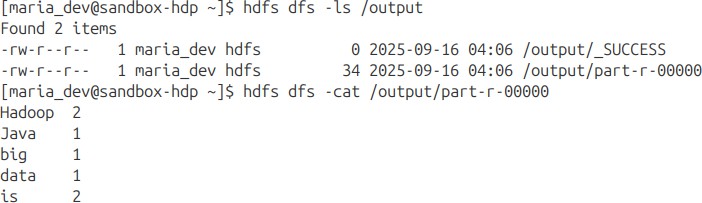
hdfs dfs -rm -r /output

hadoop jar wordcount.jar WordCountDriver /input /output

# View Output

hdfs dfs -ls /output

hdfs dfs -cat /output/part-r-00000



**Result: MapReduce has been used to perform Word Count**