**Group A**

**Lab Assignment 1d**

### **BDM 1024 - Data Technology Solutions**

Install HDP in VM Ware

Running Hortonworks Sandbox on Virtual Machine.

A screenshot of a video game

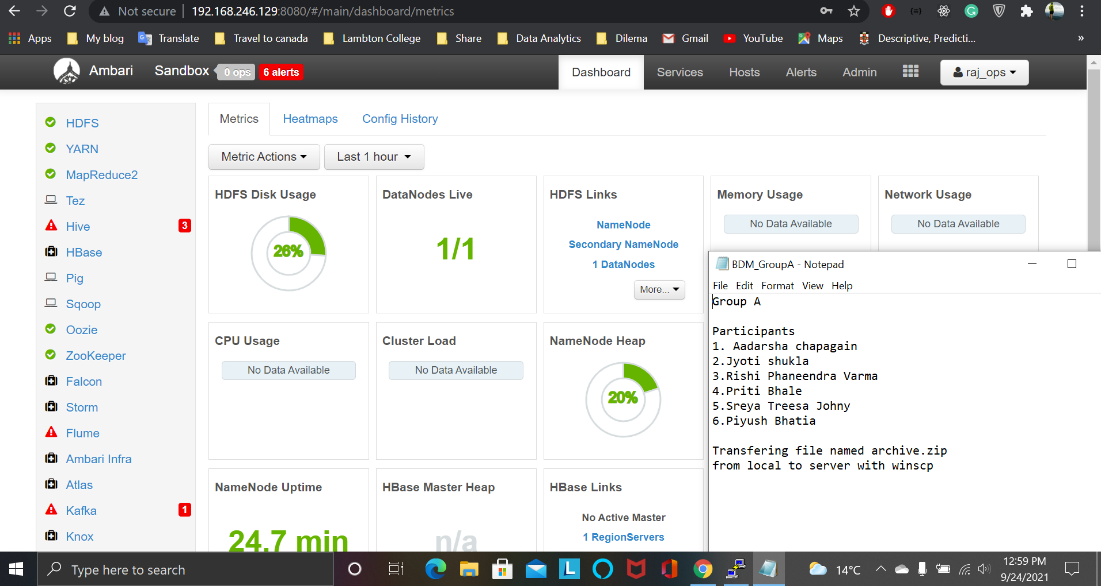
Description automatically generated

Creating Sandbox session using Putty.

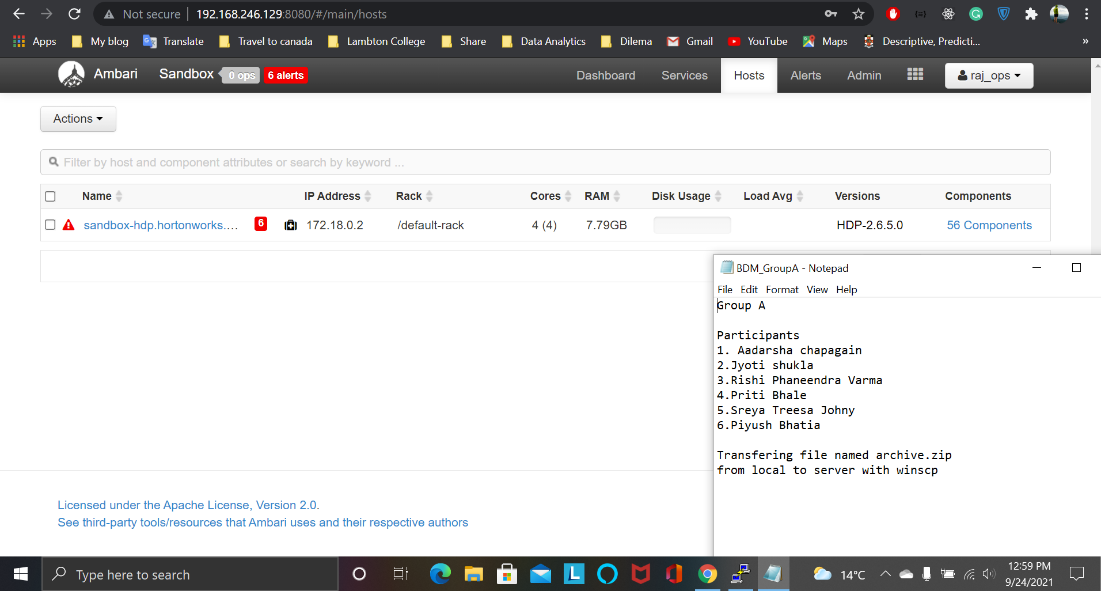
A screenshot of a computer screen

Description automatically generated with low confidence

All the components were monitored in Ambari sandbox available at the local machine at port 8080:



In the second screen shot host is listed as sandbox-hdp.hortonworks

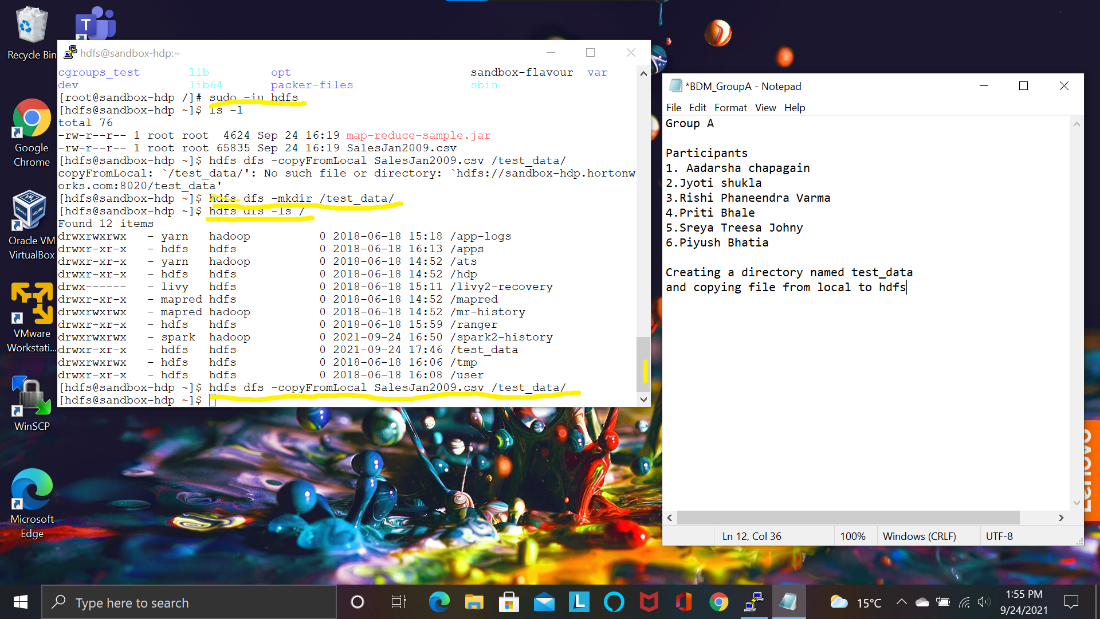


Copying file from local machine to server using winscp. Two files, one jar and and other input files were copied.

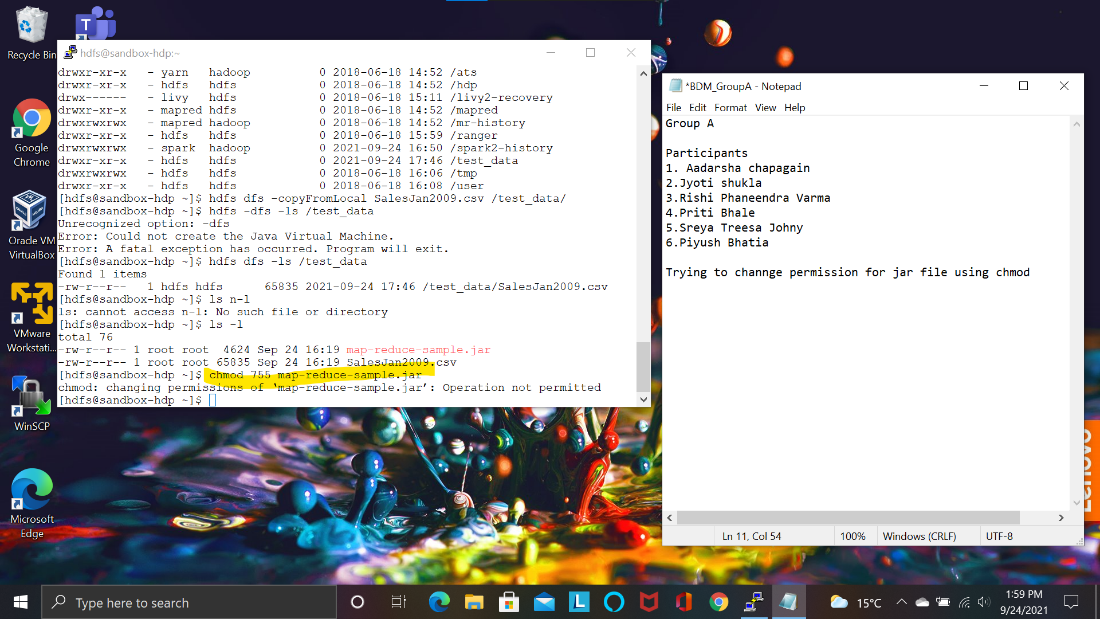
A screenshot of a computer

Description automatically generated

Directory named test\_data was created under hdfs and input csv file was copied to the same directory under hdfs



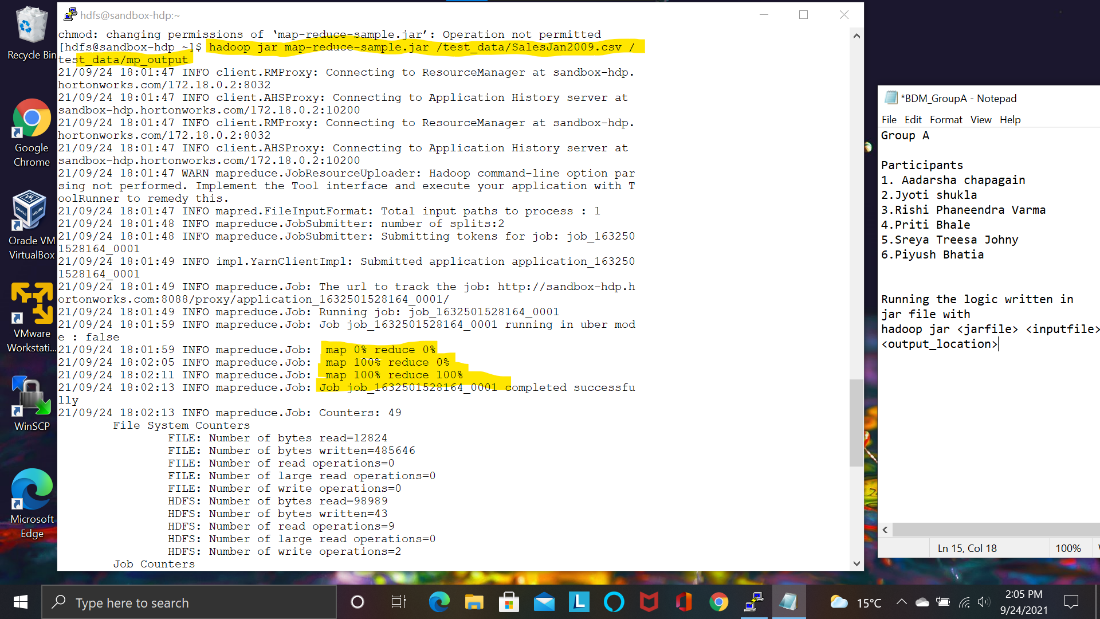
Trying to change permission of the jar file but it was not permitted.



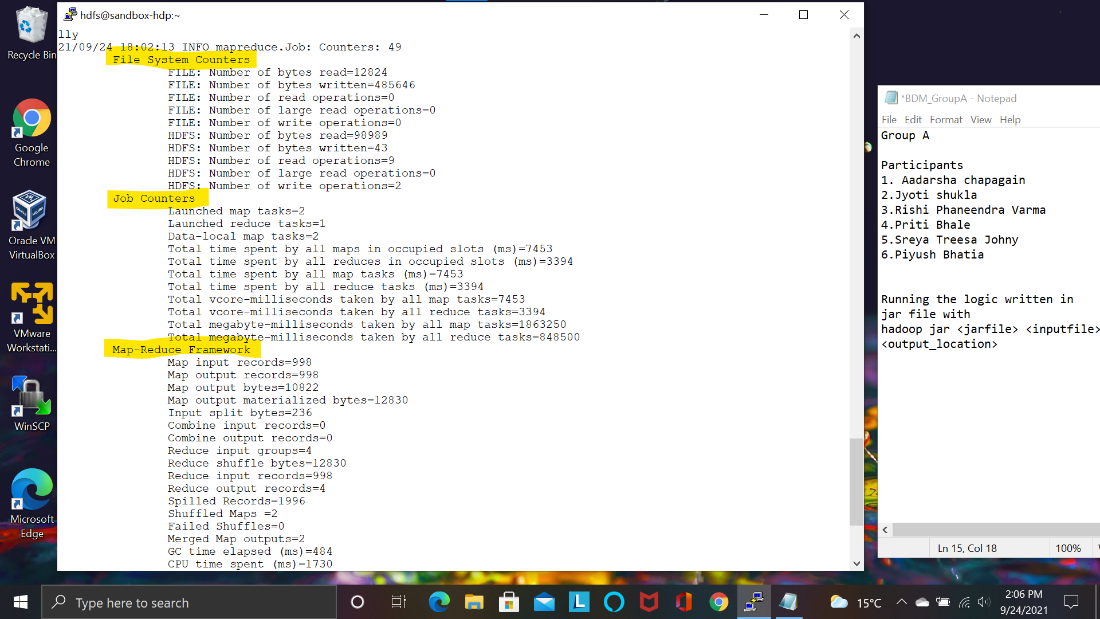
The jar file contained some logic written into it. The input csv file was passed through the jar file using

“hadoop jar <jarfile> <inputfilepathinhdfs> <outputdirectoryinhdfs>

Here directory named mp\_output was given as part of path and the directory was created



Output from the command showing details of execution



The output was written in the file named part-0000 under the directory named mp\_output

as given in command)

It was read using cat command.

Hdfs dfs –cat <filepath>

