Group B

# refer https://youtu.be/6sK3LDY7Pp4

#install hadoop and java

#run $ in the terminal

$ hadoop version

$ javac -version

#create a folder name WordCount1

#create a file with name WordCount.java and paste the java code inside it

#now create a folder in the WordCount1 with name input\_data and create a document with name input.txt and type anything inside it

#now in WordCount1 create a new folder to hold the java class files

$ export HADOOP\_CLASSPATH=$(hadoop classpath)

$echo $HADOOP\_CLASSPATH

#create a directory on HDFS $ hadoop fs -mkdir <DIRECTORY NAME>

$ hadoop fs -mkdir /WordCount1

#create a directory inside it for the input hadoop fs -mkdir <HDFS\_INPUT\_DIRECTORY>

$hadoop fs -mkdir /WordCount1/Input

#now open a browser and go to localhost:50070

#upload the input file to that directory: $ hadoop fs -put <input\_file> <hdfs\_input\_directory>

#change the current directory to the WordCount1 directory

#compile the java code: javac -classpath ${HADOOP\_CLASSPATH} -d <CLASSES\_FOLDER> <JAVA\_FILE>

#put the output files in one jar file: jar-cvf <JAR\_FILE\_NAME> -C <CLASSES\_FOLDER>

#now we have the jar file

#now run the jar file on hadoop jar

#output: hadoop dfs -cal <HDFS\_OUTPUT\_DIRECTORY>