**MAP REDUCE**

**COMMANDS to execute on your ubantu terminal :**

**Create a folder called analyzelogs in the home and paste following ﬁles in it:**

**1)** **Link CLASSPATH & MANIFEST**

**2)** **Link log ﬁle**

**3)** **Mapper**

**4)** **Reducer**

**5)** **Driver**

**These ﬁles can be downloaded from the url -**

**<https://www.youtube.com/watch?v=gXQT7vHB59Q>**

1. cd ~/analyzelogs

**######################################**

**convert the accesslog.txt to csv by opening it in libre calc and setting rows 2,**

**separated by other –**

**######################################**

2. ls

3. export CLASSPATH="$HADOOP\_HOME/share/hadoop/mapreduce/

hadoopmapreduce-client-core-2.9.0.jar:$HADOOP\_HOME/share/hadoop/mapreduce/

hadoop-mapreduce-client-common-2.9.0.jar:$HADOOP\_HOME/share/hadoop/

common/hadoop-common-2.9.0.jar:~/analyzelogs/SalesCountry/\*:$HADOOP\_HOME/

lib/\*"

4. javac -d . SalesMapper.java SalesCountryReducer.java SalesCountryDriver.java

5. nano Manifest.txt

**######################################**

**Paste following in the Manifest.txt**

Main-Class: SalesCountry.SalesCountryDriver

**######################################**

6. jar -cfm analyzelogs.jar Manifest.txt SalesCountry/\*.class

7. start-dfs.sh

8. start-yarn.sh

9. mkdir ~/input2000

10. cp access\_log\_short.csv ~/input2000/

11. $HADOOP\_HOME/bin/hdfs dfs -put ~/input2000 /

12. $HADOOP\_HOME/bin/hadoop jar analyzelogs.jar /input2000 /

13. output2000 $HADOOP\_HOME/bin/hdfs dfs -cat /output2000/part-

00000