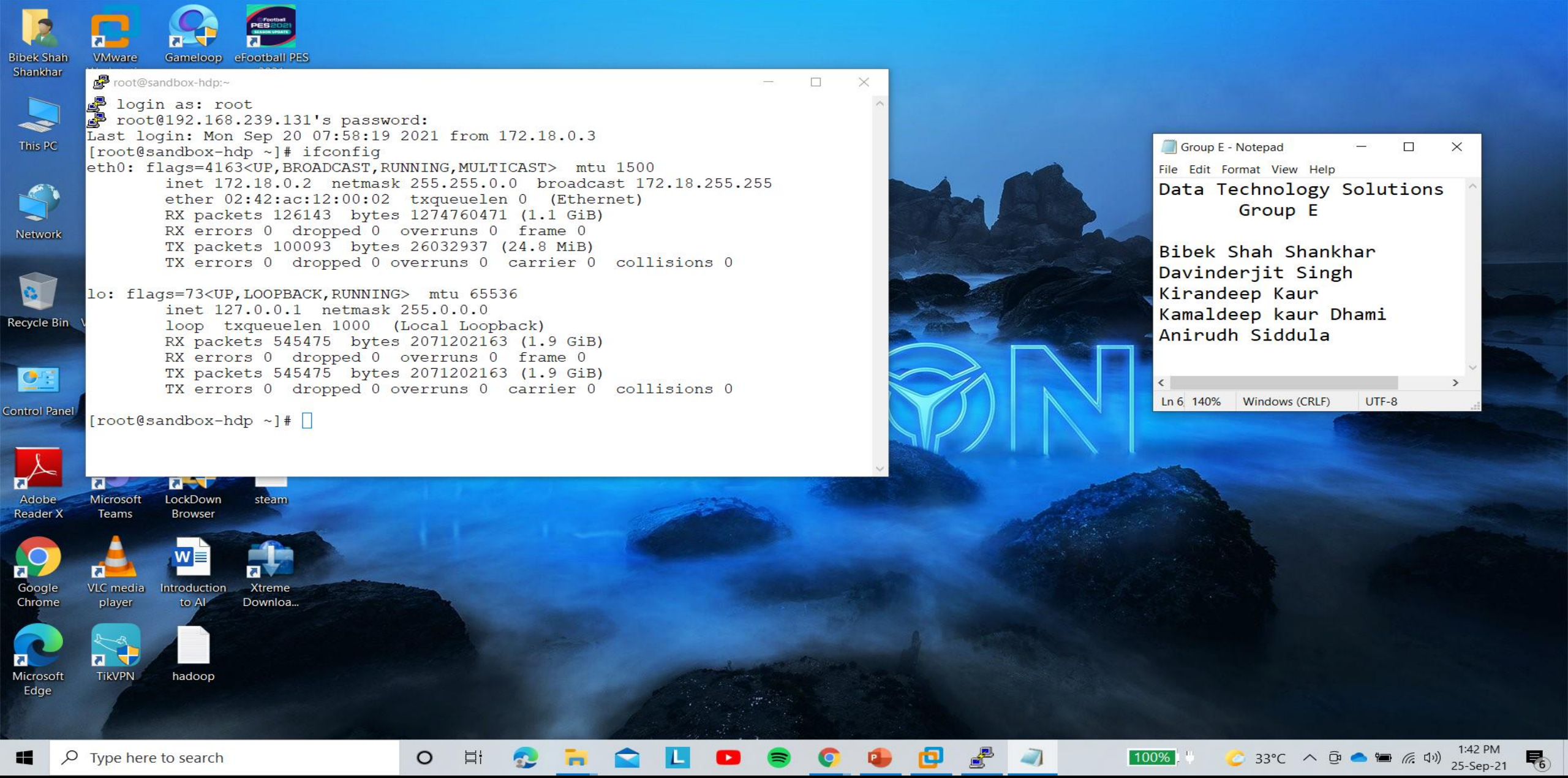
Group E

BDM 1024 - Data Technology Solutions

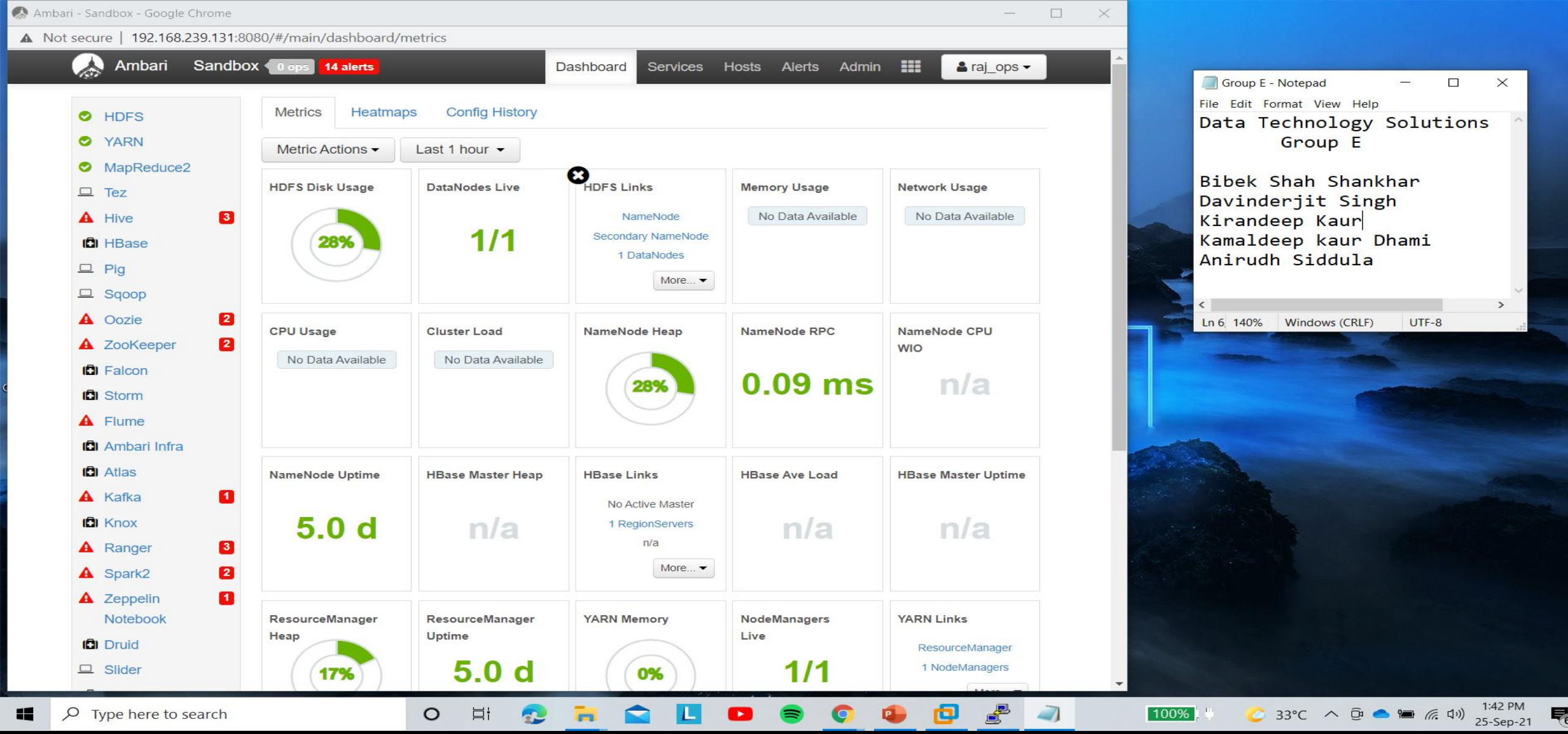
Assignment: Lab 1d -MapReduce

Submitted by:

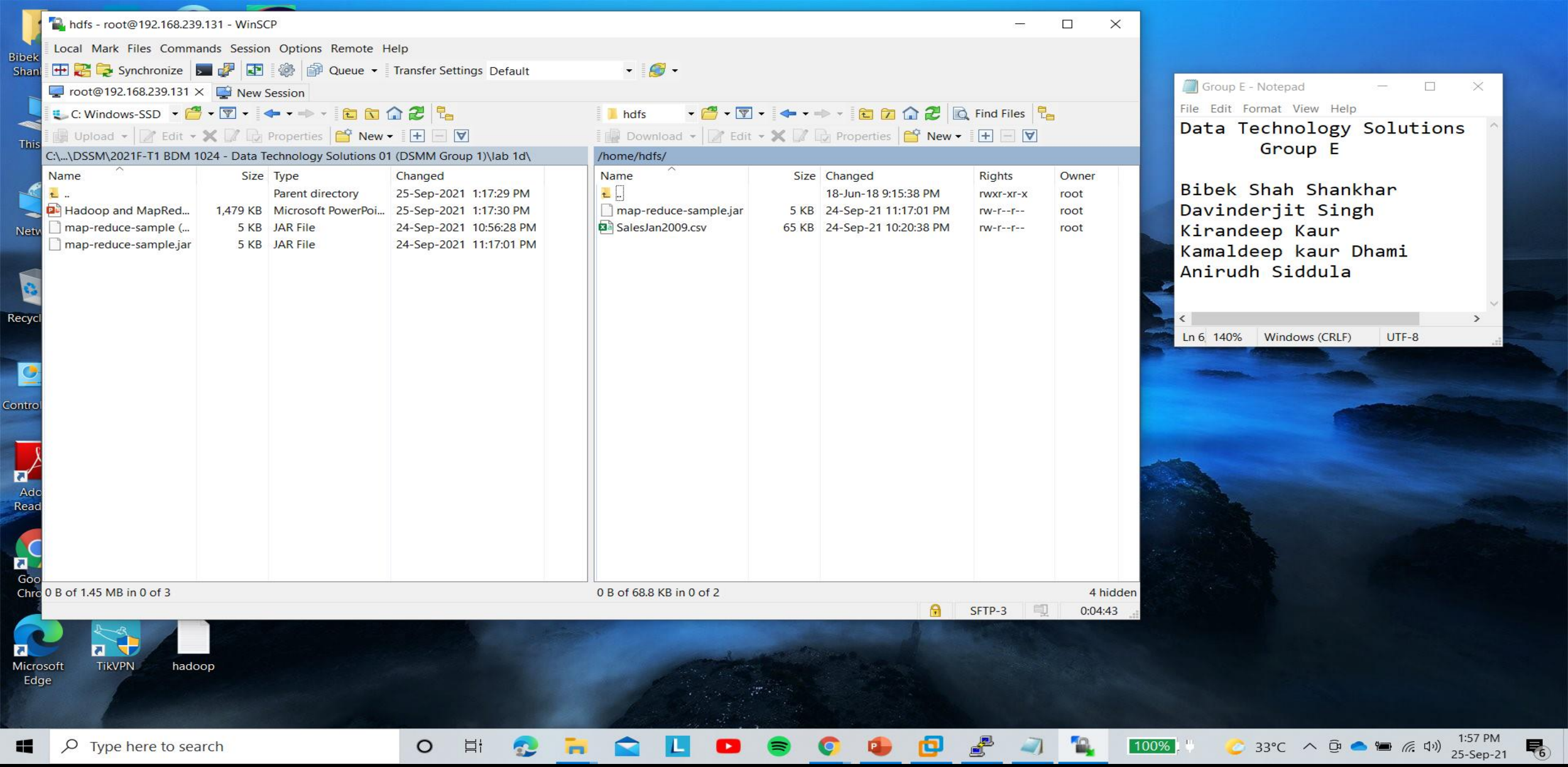
1. Bibek Shah Shankhar
2. Anirudh Siddula
3. Kirandeep Kaur
4. Kamaldeep kaur Dhami
5. Davinderjit Singh



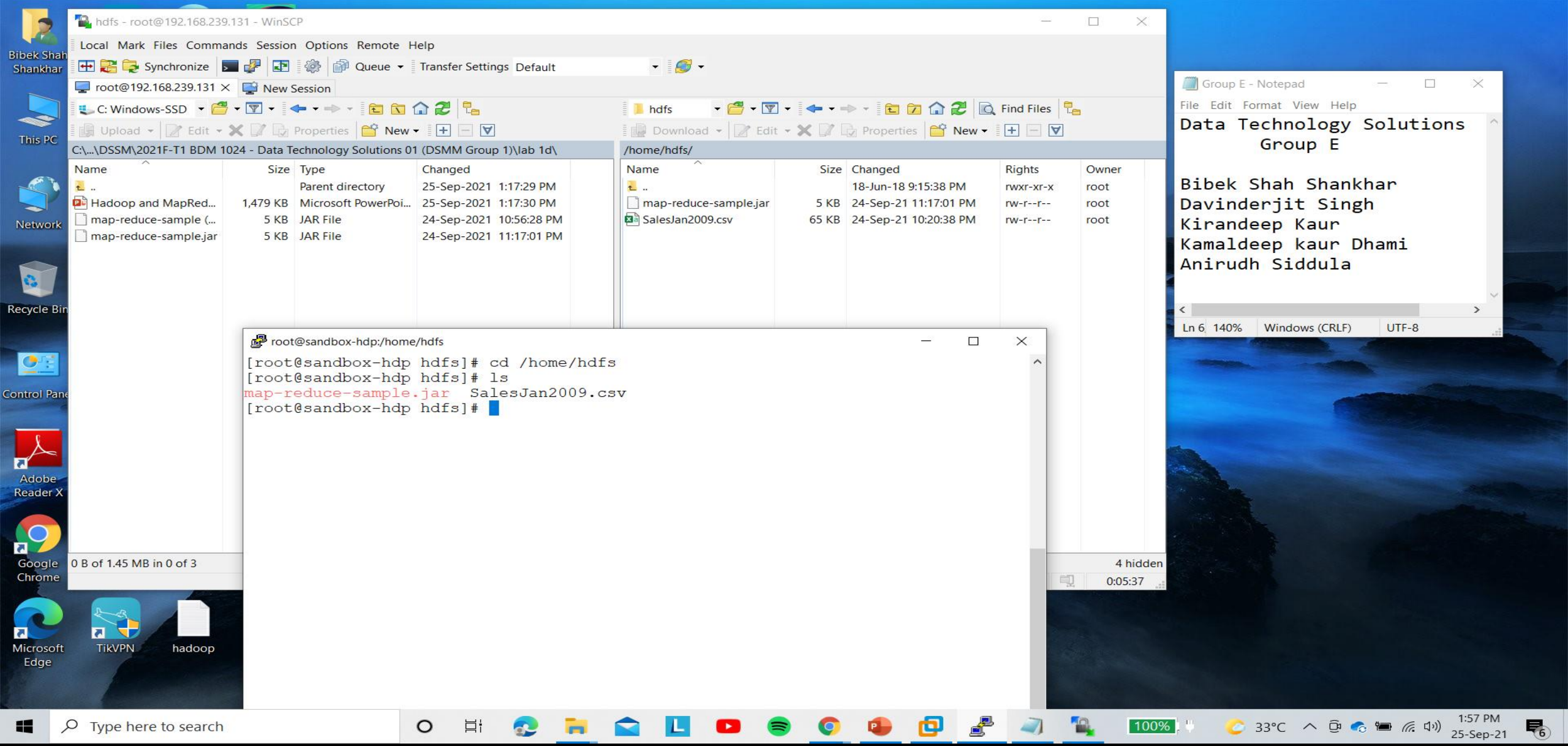
Login To Putty and Checking Ip address.



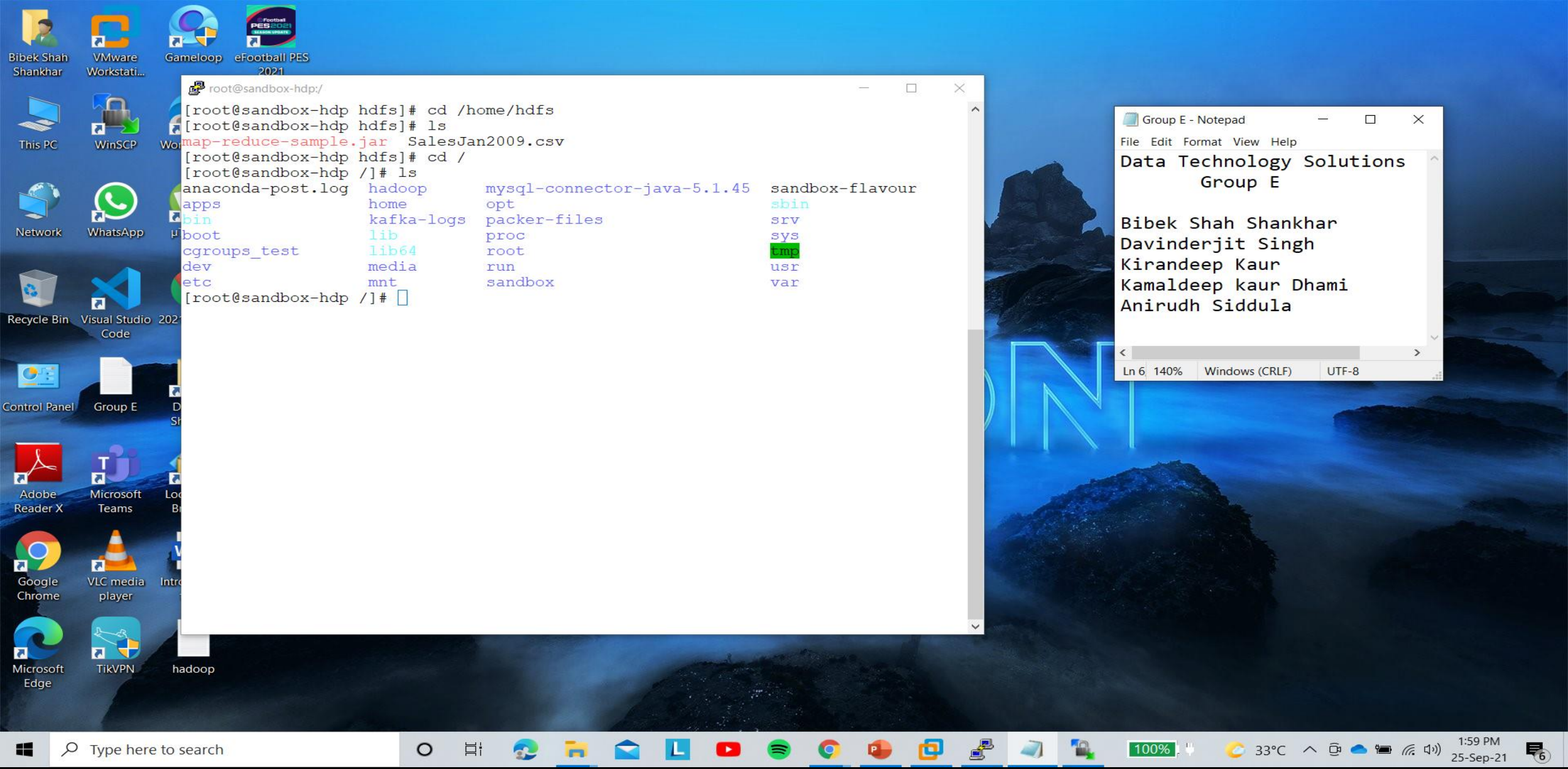
Stopping Unused Services From Dashboard.



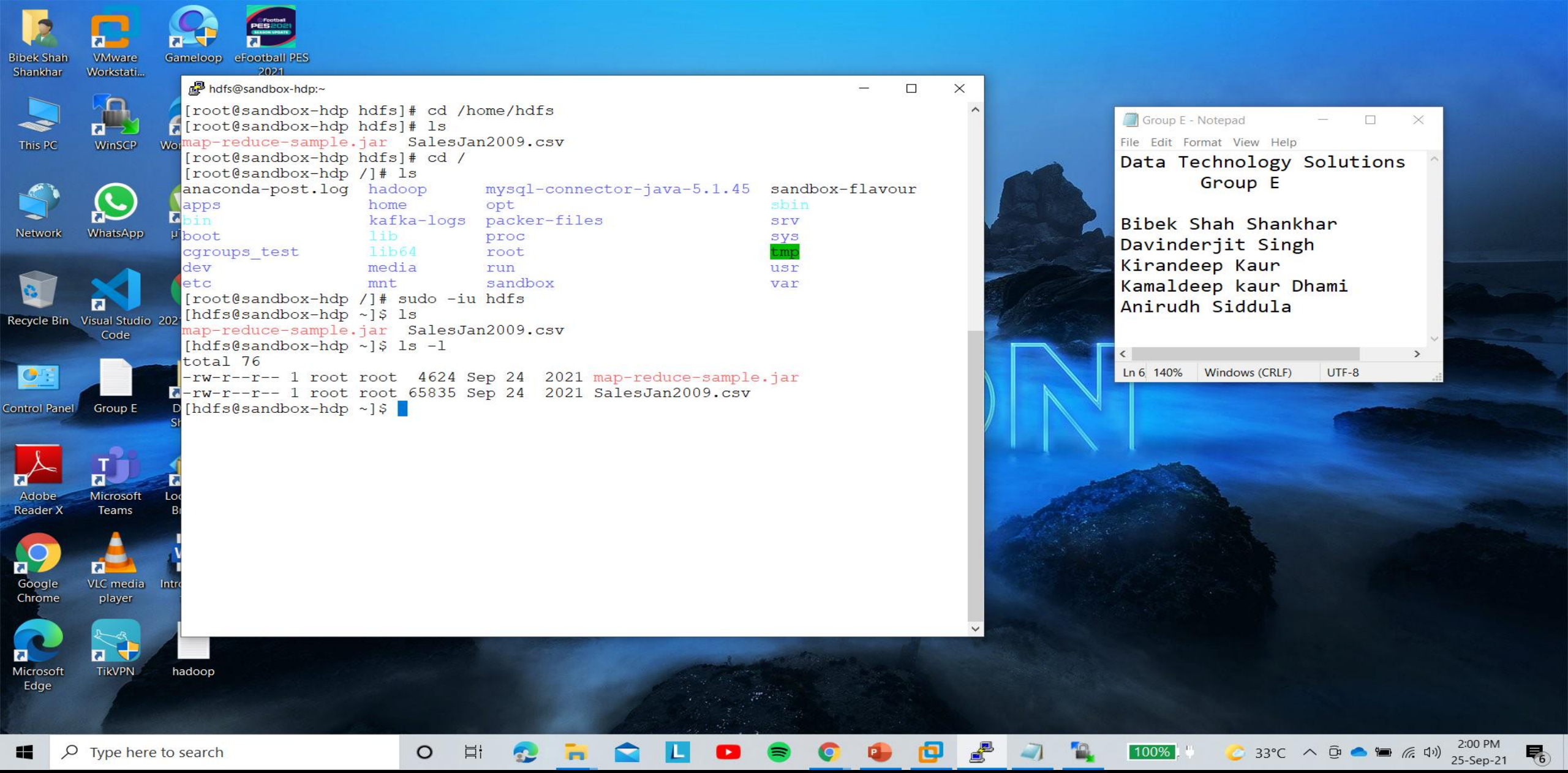
Transferring Local File From System To Hdfs Directory Through WinSCP.



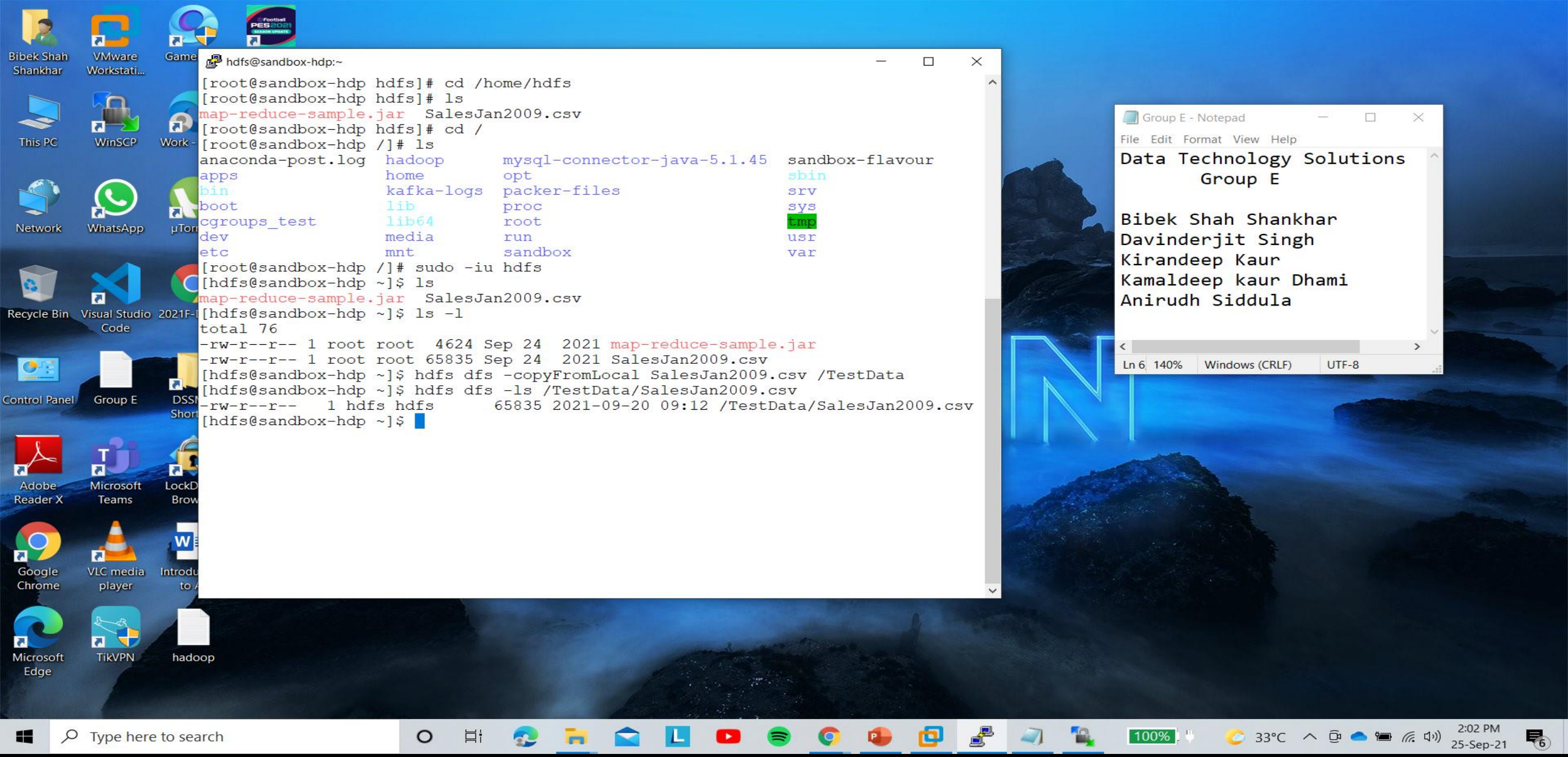
Checking Transferred File from Putty Commands.



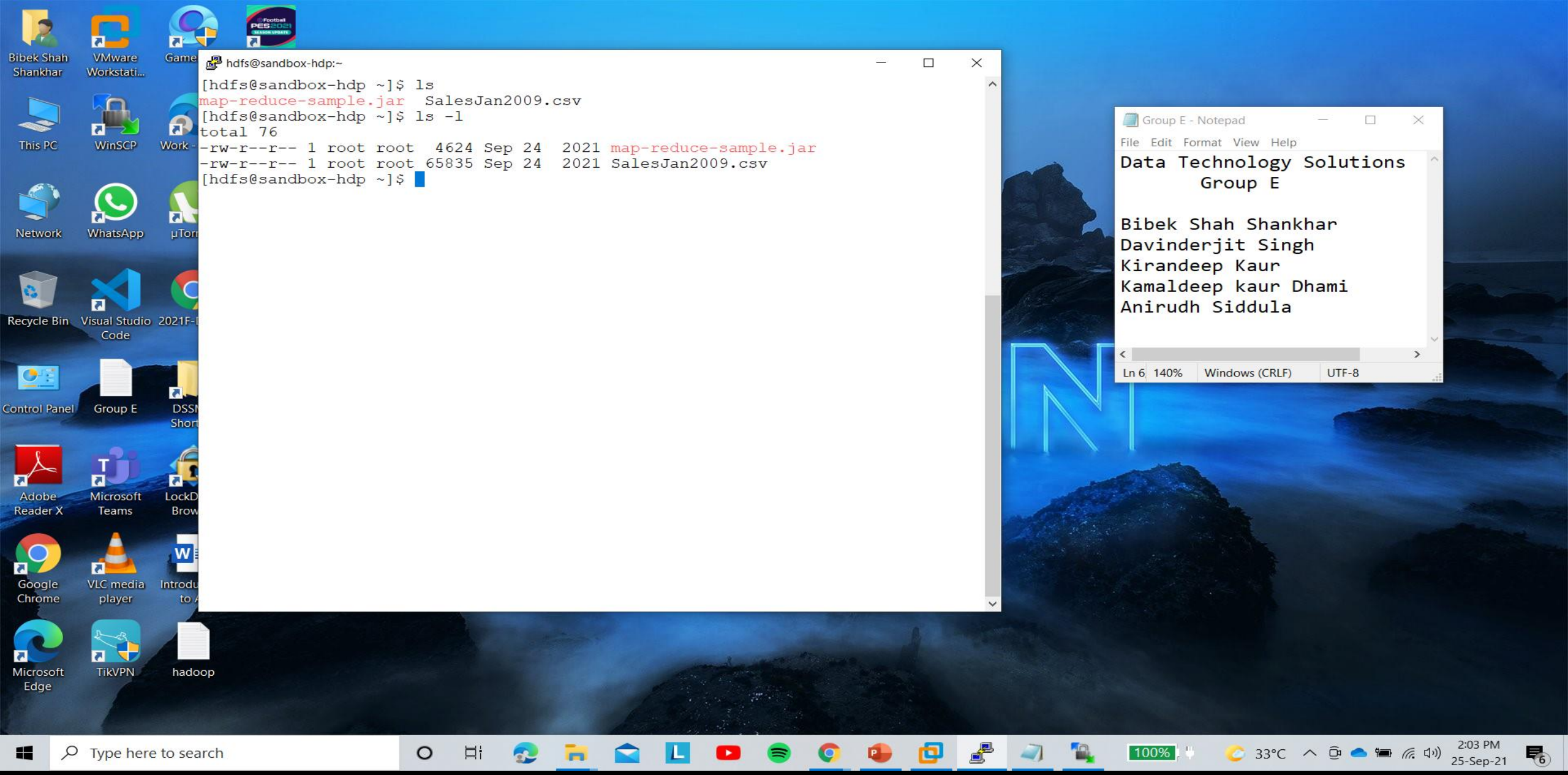
Listing Files in The Directories.



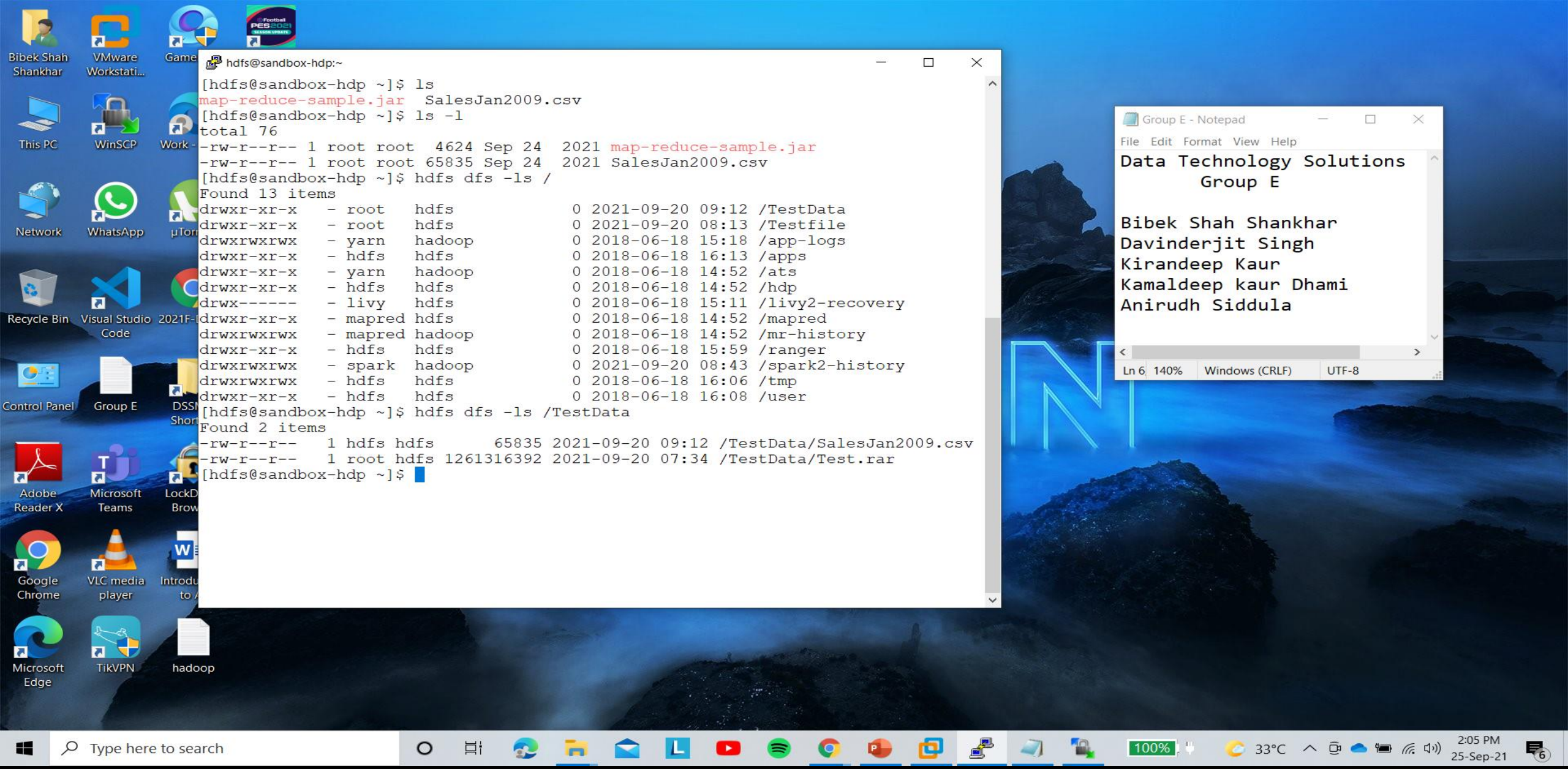
Logging in as hdfs user.



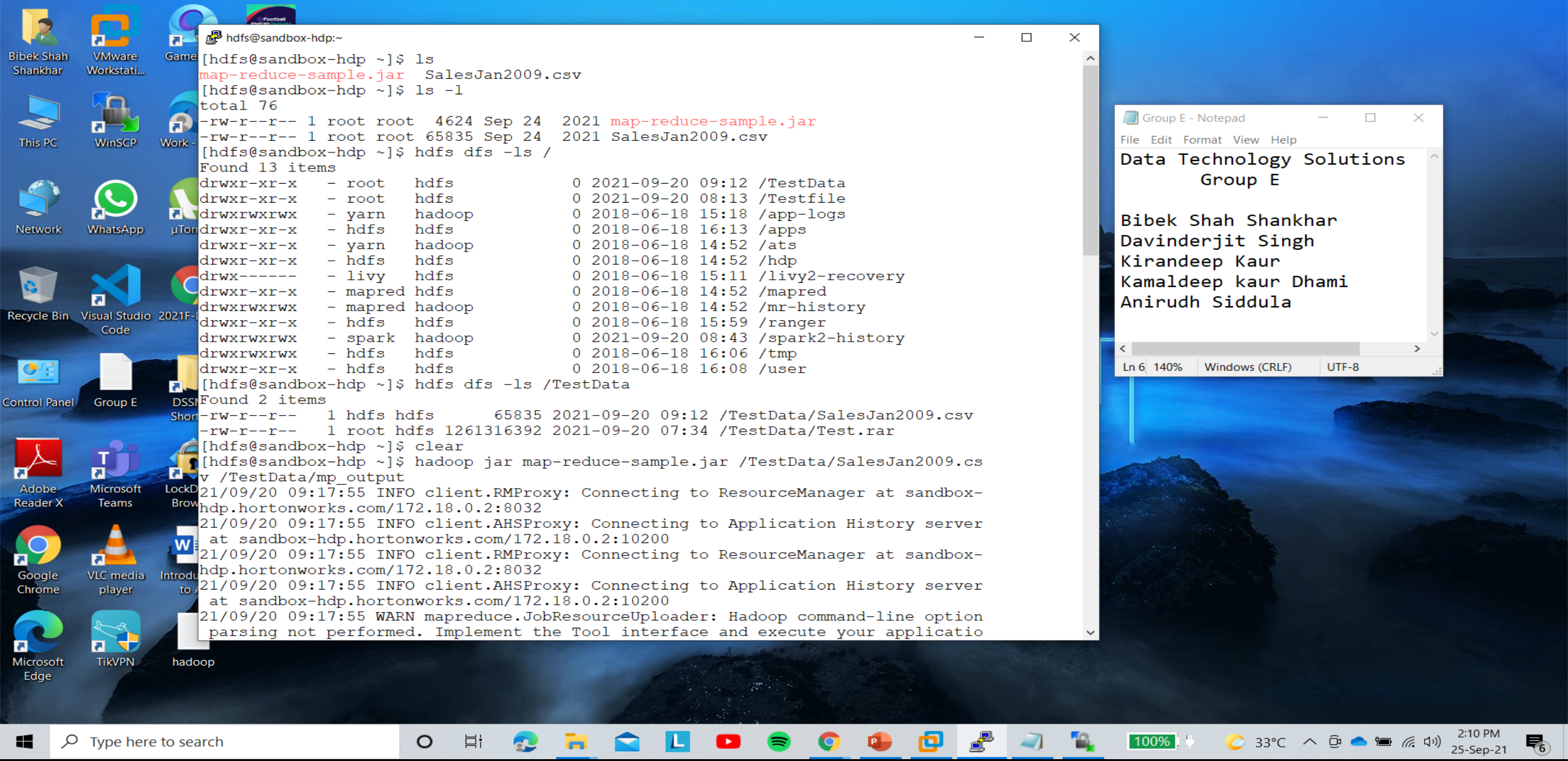
Copying Files from Local System to HDFS and Listing.



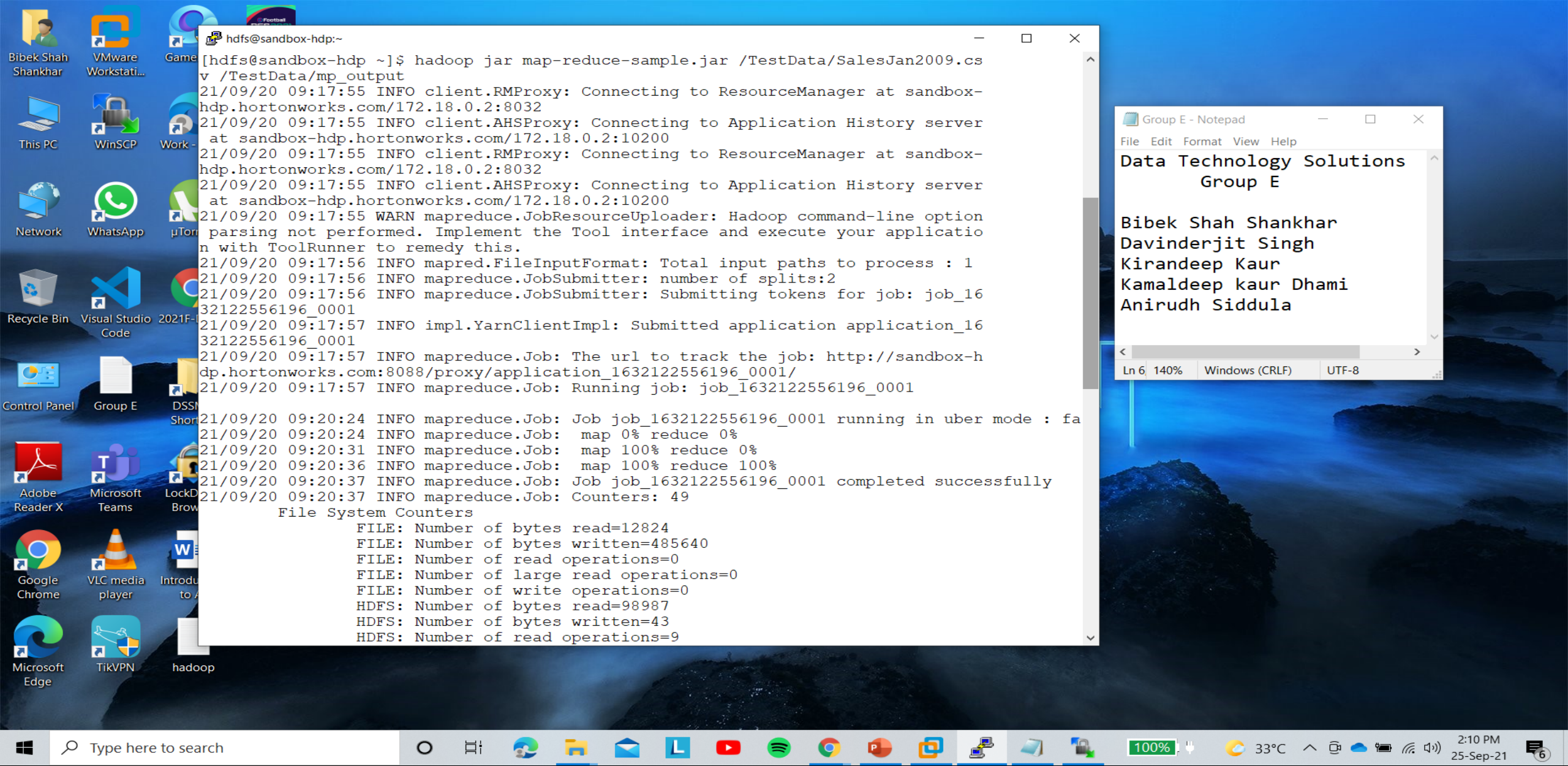
Listing the Copied Files.



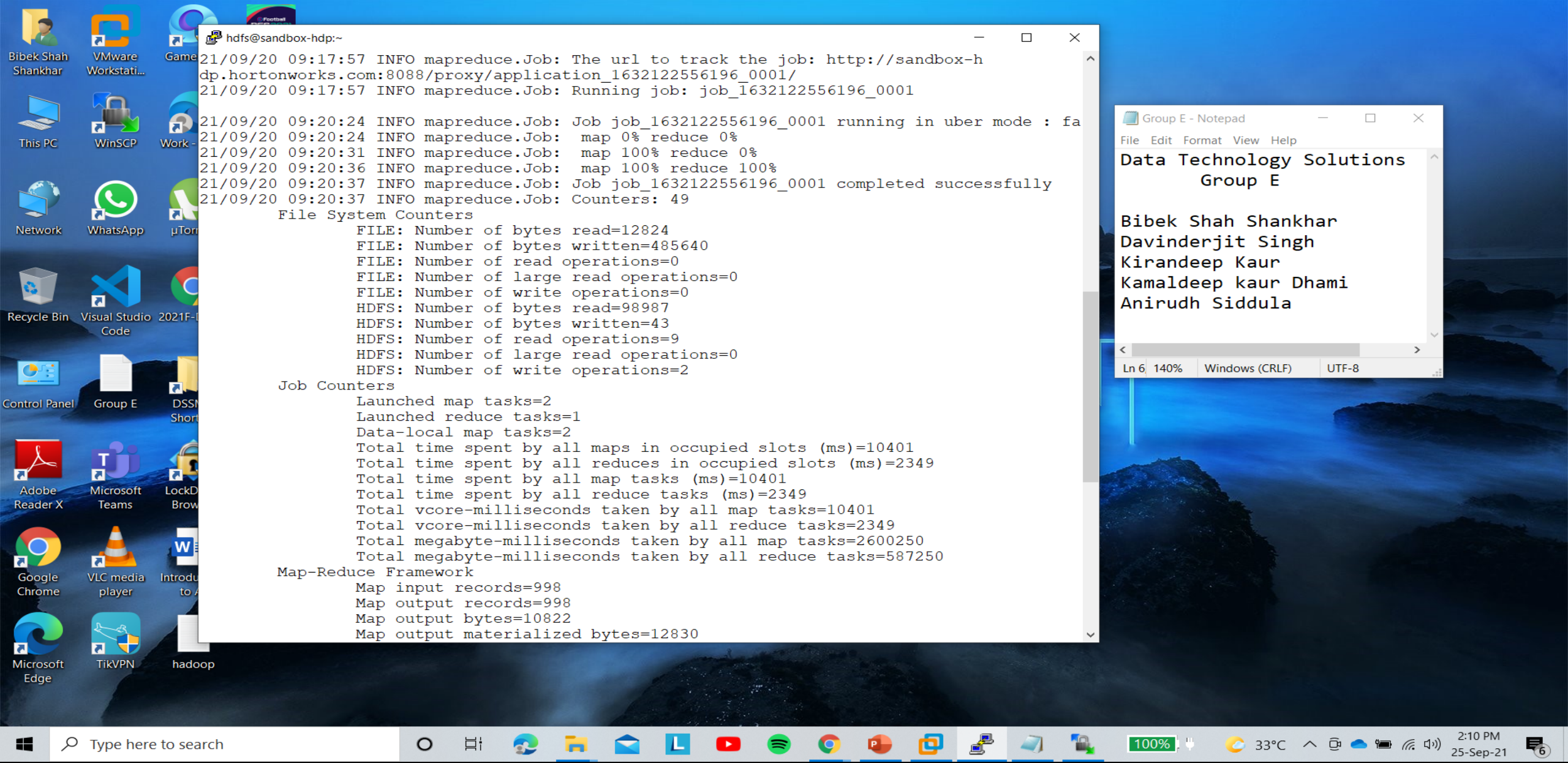
Listing files from Hdfs and From TestData Directory Which includes our Project Files.



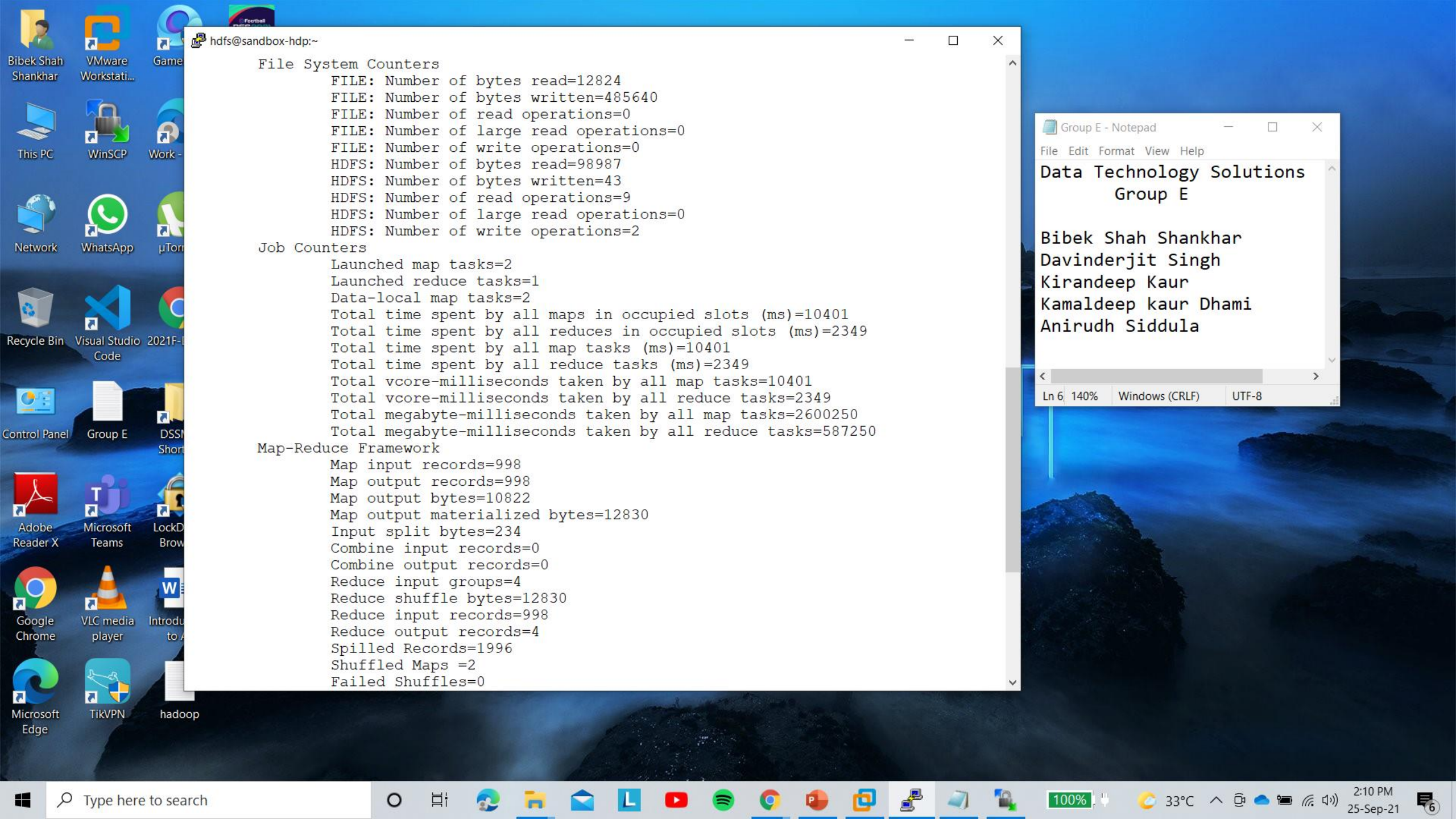
We run the Jar file Which Contains The mapreduce code using Hadoop to execute SalesJan2009.csv and Trying to store Results in TestData/mp_output Directory.

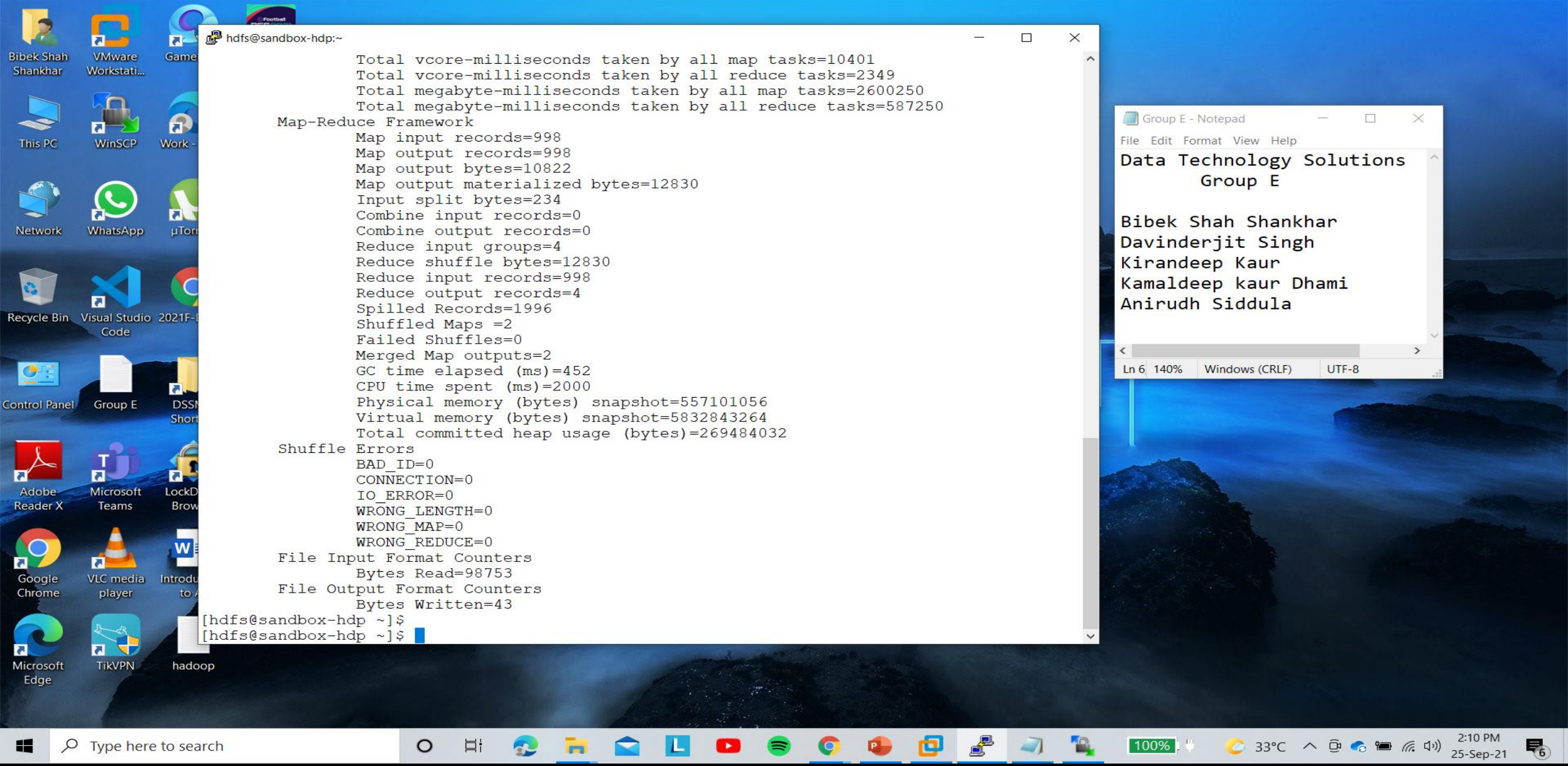


Our map-reduce Code is running and executed Sucessfully 100%.

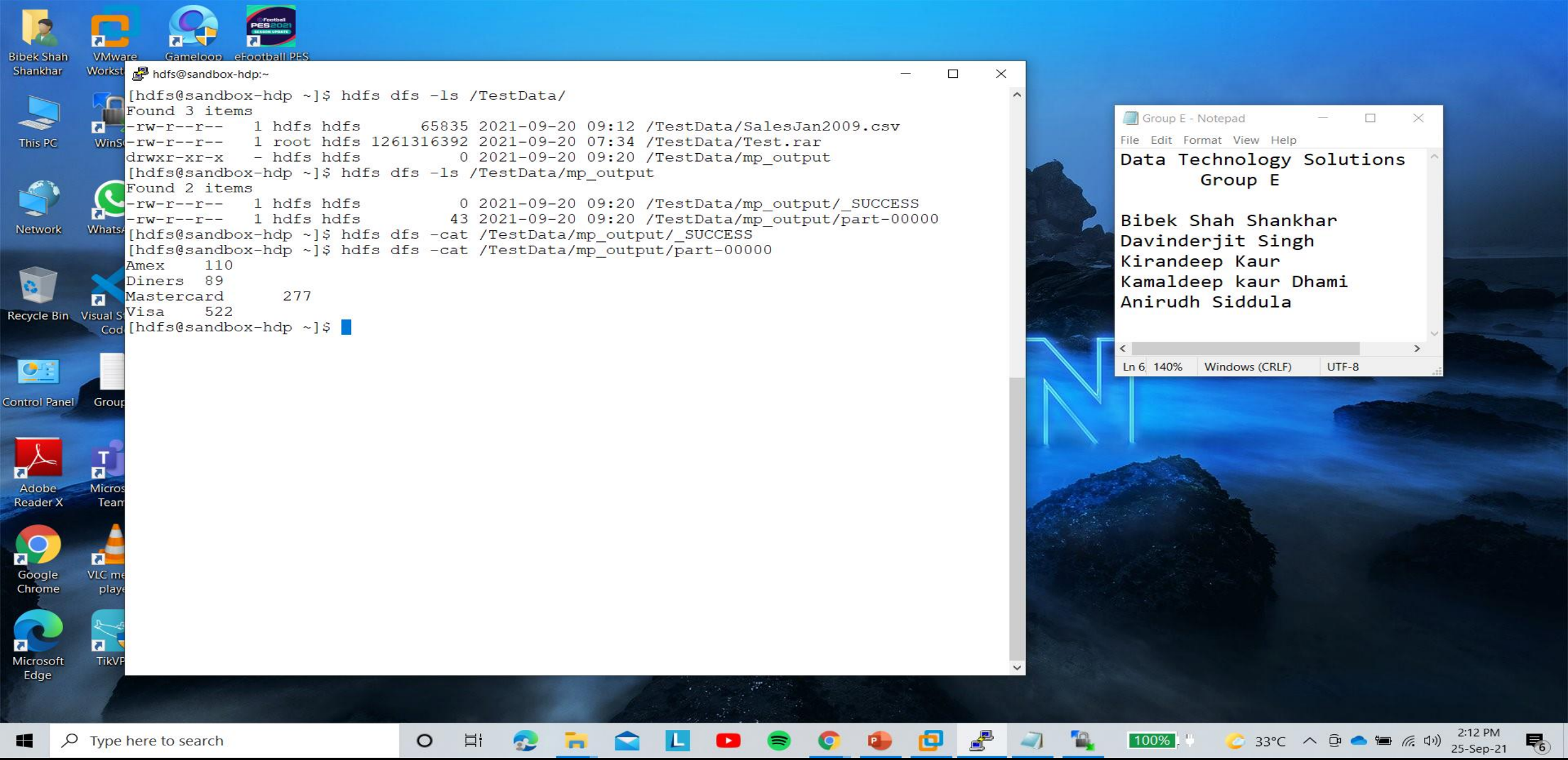


We Have 49 Counters.


hdfs@sandbox-hdp:~
File System Counters
FILE: Number of bytes read=12824
FILE: Number of bytes written=485640
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=98987
HDFS: Number of bytes written=43
HDFS: Number of read operations=9
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=2
Launched reduce tasks=1
Data-local map tasks=2
Total time spent by all maps in occupied slots (ms)=10401
Total time spent by all reduces in occupied slots (ms)=2349
Total time spent by all map tasks (ms)=10401
Total time spent by all reduce tasks (ms)=2349
Total vcore-milliseconds taken by all map tasks=10401
Total vcore-milliseconds taken by all reduce tasks=2349
Total megabyte-milliseconds taken by all map tasks=2600250
Total megabyte-milliseconds taken by all reduce tasks=587250
Map-Reduce Framework
Map input records=998
Map output records=998
Map output bytes=10822
Map output materialized bytes=12830
Input split bytes=234
Combine input records=0
Combine output records=0
Reduce input groups=4
Reduce shuffle bytes=12830
Reduce input records=998
Reduce output records=4
Spilled Records=1996
Shuffled Maps =2
Failed Shuffles=0
Group E - Notepad
File Edit Format View Help
Data Technology Solutions
Group E
Bibek Shah Shankhar
Davinderjit Singh
Kirandeep Kaur
Kamaldeep kaur Dhami
Anirudh Siddula
Ln 6 140% Windows (CRLF) UTF-8
Bibek Shah Shankhar
VMware Workstati...
This PC
WinSCP
Network
WhatsApp
Recycle Bin
Visual Studio Code
Control Panel
Group E
Adobe Reader X
Microsoft Teams
Google Chrome
VLC media player
Microsoft Edge
TikVPN
hadoop
Type here to search
33°C
2:10 PM
25-Sep-21



Code Executed Sucessfully and our result is stored in /TestData/mp_output



We can see we have a result and Analysis of our mapreduce code.