

[root@sandbox-hdp ~]# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/hive --username root --P --table TBLS --target-dir /user/bibek/sqoop/tbls --fields-terminated-by '|' --num-mapper 8 --delete-target-dir  
Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.  
Please set \$ACCUMULO\_HOME to the root of your Accumulo installation.  
21/10/02 13:48:06 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292  
Enter password:  
21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Error parsing arguments for import:  
21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: --num-mapper  
21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: 8  
21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: --delete-target-dir  
  
Try --help for usage instructions.  
[root@sandbox-hdp ~]# clear  
[root@sandbox-hdp ~]# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/hive --username root --P --table TBLS --target-dir /user/bibek/sqoop/tbls --fields-terminated-by '|' --num-mappers 8 --delete-target-dir  
Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.  
Please set \$ACCUMULO\_HOME to the root of your Accumulo installation.  
21/10/02 13:49:35 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292  
Enter password:  
21/10/02 13:49:42 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.  
21/10/02 13:49:42 INFO tool.CodeGenTool: Beginning code generation  
Sat Oct 02 13:49:42 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.  
21/10/02 13:49:42 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `TBLS` AS t LIMIT 1  
21/10/02 13:49:42 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `TBLS` AS t LIMIT 1  
21/10/02 13:49:42 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/hdp/2.6.5.0-292/hadoop-mapreduce  
  
Note: /tmp/sqoop-root/compile/77e5968757b63b13332d74ff51196eld/TBLS.java uses or overrides a deprecated API.  
  
[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 8 150% Windows (CRLF) UTF-8  
  
7:35 PM 02-Oct-21

root@sandbox-hdp:~# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/hive --username root --P --table TBLS --target-dir /user/bibek/sqoop/tbls --fields-terminated-by '|' --num-mapper 8 --delete --target-dir

Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.

Please set \$ACCUMULO\_HOME to the root of your Accumulo installation.

21/10/02 13:48:06 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292

Enter password:

21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Error parsing arguments for import:

21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: --num-mapper

21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: 8

21/10/02 13:48:11 ERROR tool.BaseSqoopTool: Unrecognized argument: --delete-target-dir

Try --help for usage instructions.

[root@sandbox-hdp ~]# clear

[root@sandbox-hdp ~]# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/hive --username root --P --table TBLS --target-dir /user/bibek/sqoop/tbls --fields-terminated-by '|' --num-mappers 8 --delete-target-dir

Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.

Please set \$ACCUMULO\_HOME to the root of your Accumulo installation.

21/10/02 13:49:35 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292

Enter password:

21/10/02 13:49:42 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

21/10/02 13:49:42 INFO tool.CodeGenTool: Beginning code generation

Sat Oct 02 13:49:42 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.

21/10/02 13:49:42 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `TBLS` AS t LIMIT 1

21/10/02 13:49:42 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `TBLS` AS t LIMIT 1

21/10/02 13:49:42 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/hdp/2.6.5.0-292/hadoop-mapreduce

Note: /tmp/sqoop-root/compile/77e5968757b63b13332d74ff51196eld/TBLS.java uses or overrides a deprecated API.

Group E - Notepad

Data Technology Solutions

Group E

Bibek Shah Shankhar

Davinderjit Singh

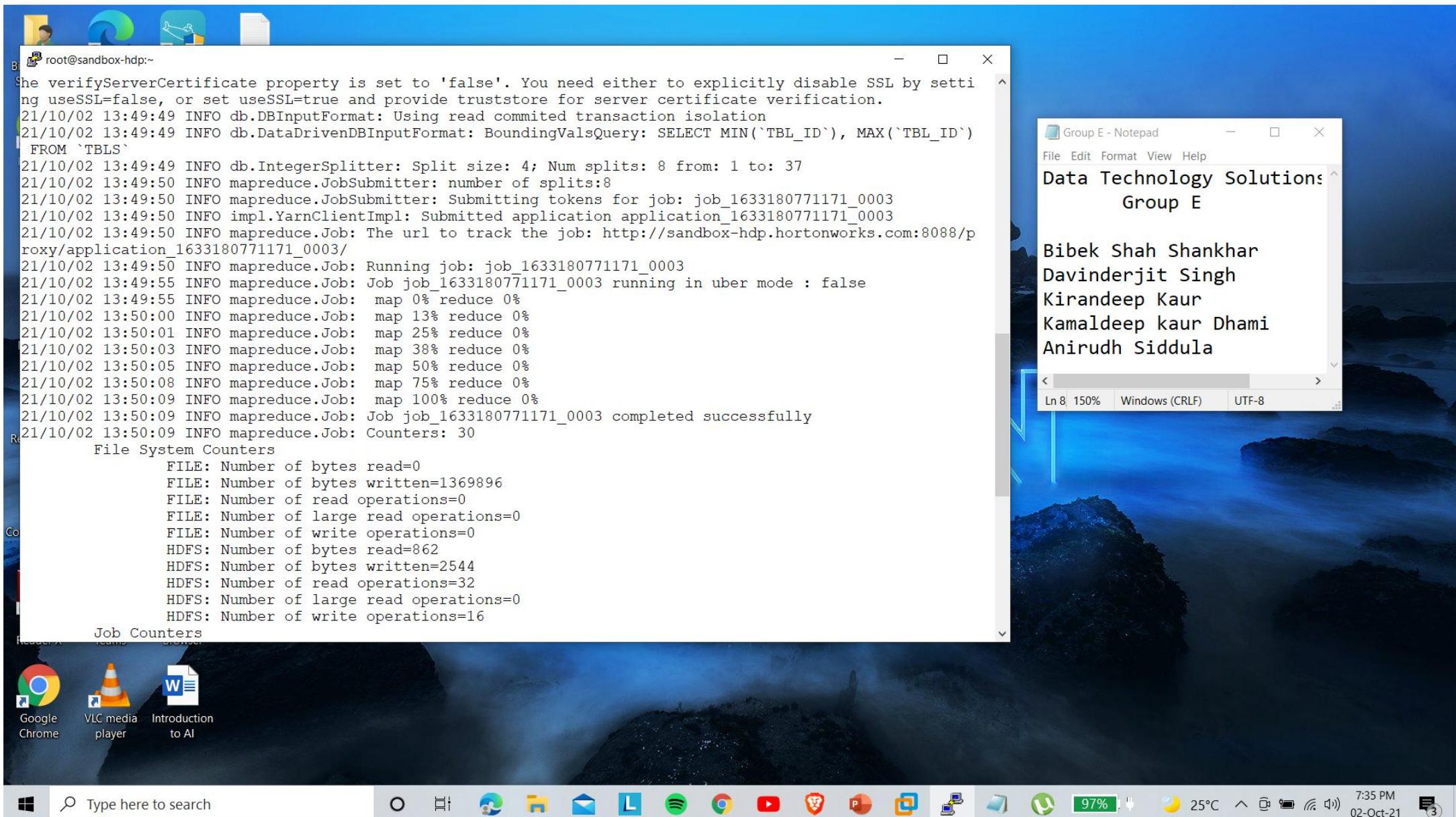
Kirandeep Kaur

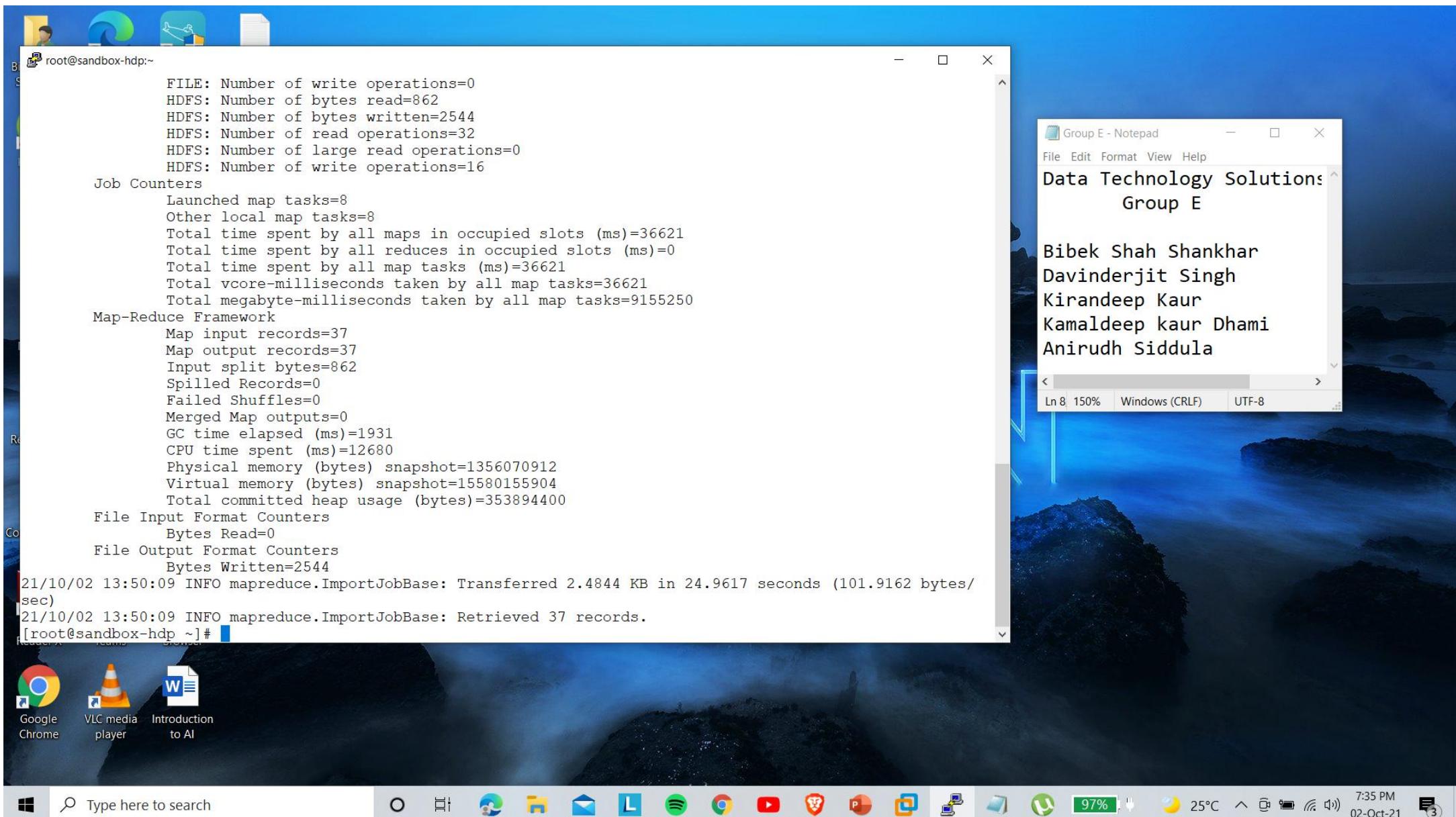
Kamaldeep kaur Dhami

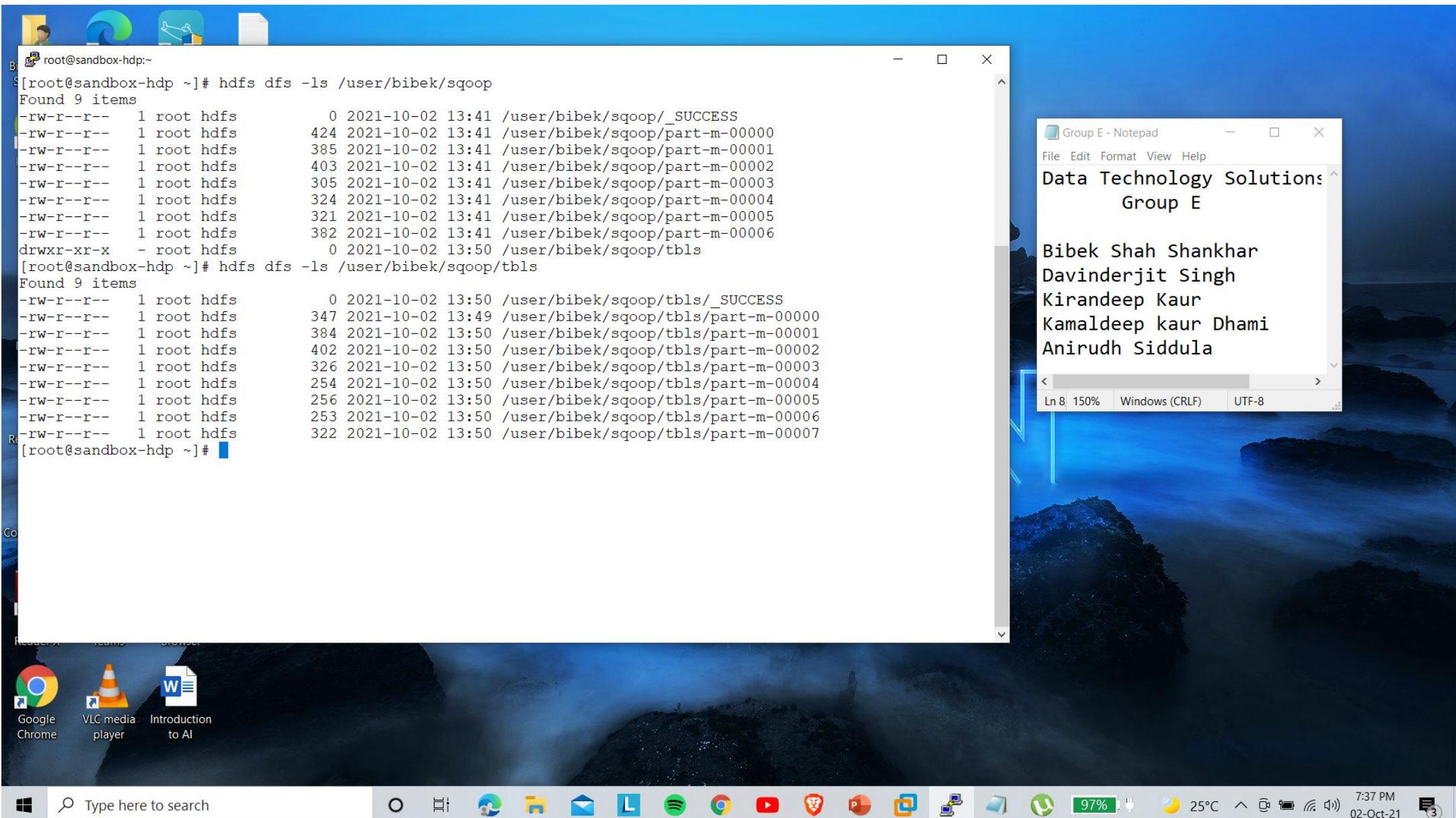
Anirudh Siddula

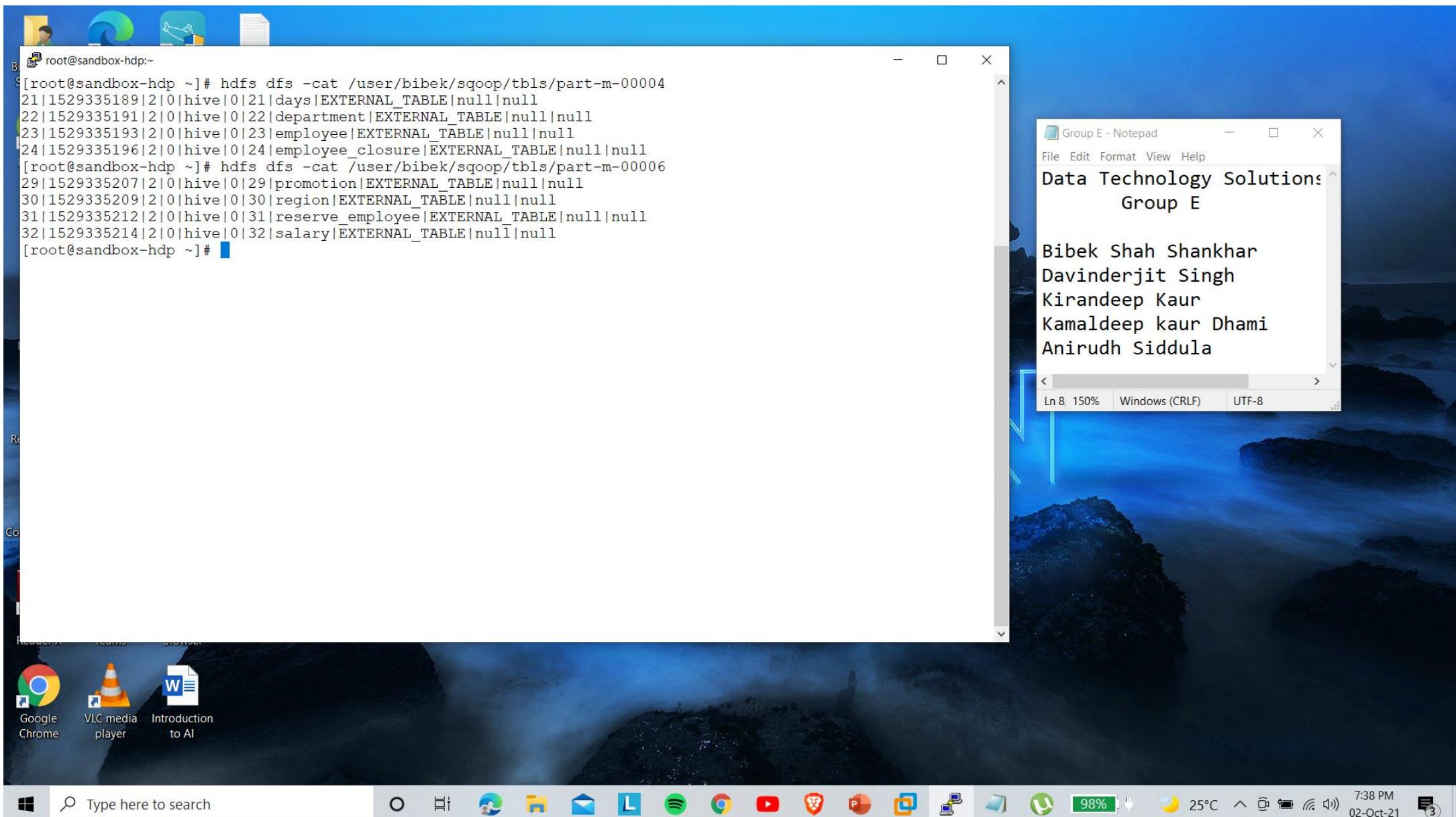
Ln 8 150% Windows (CRLF) UTF-8

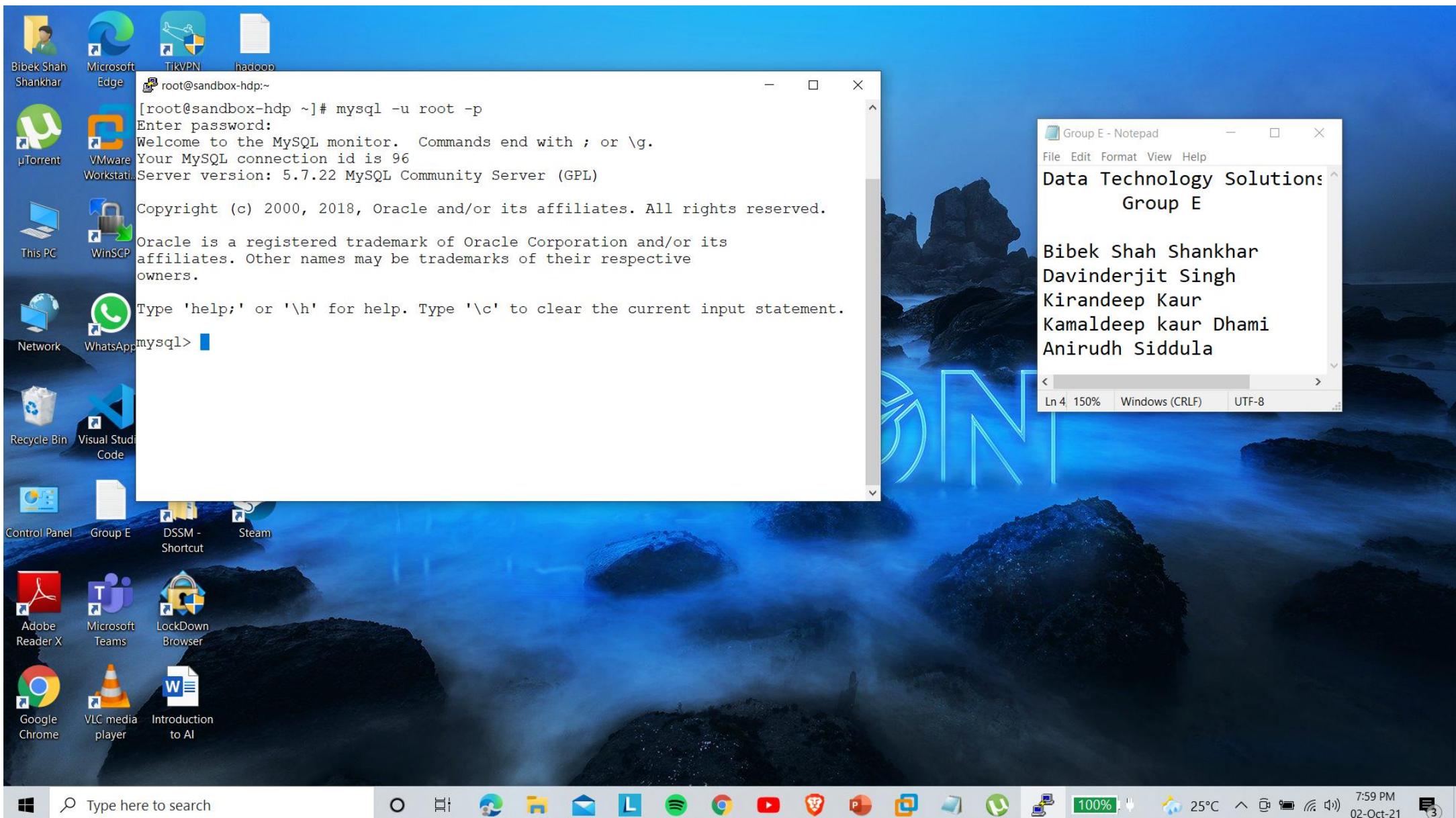
7:35 PM 02-Oct-21 3

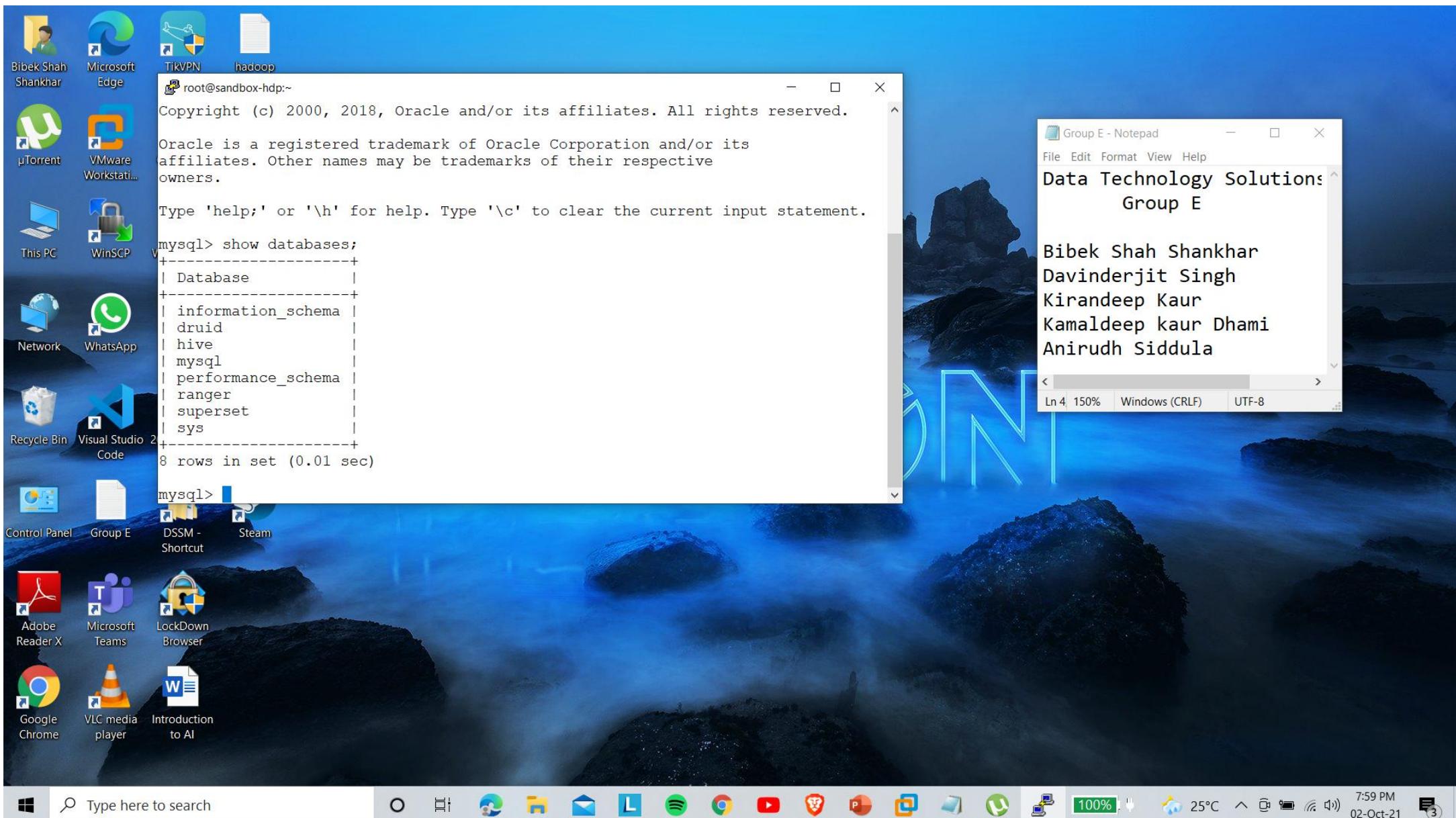


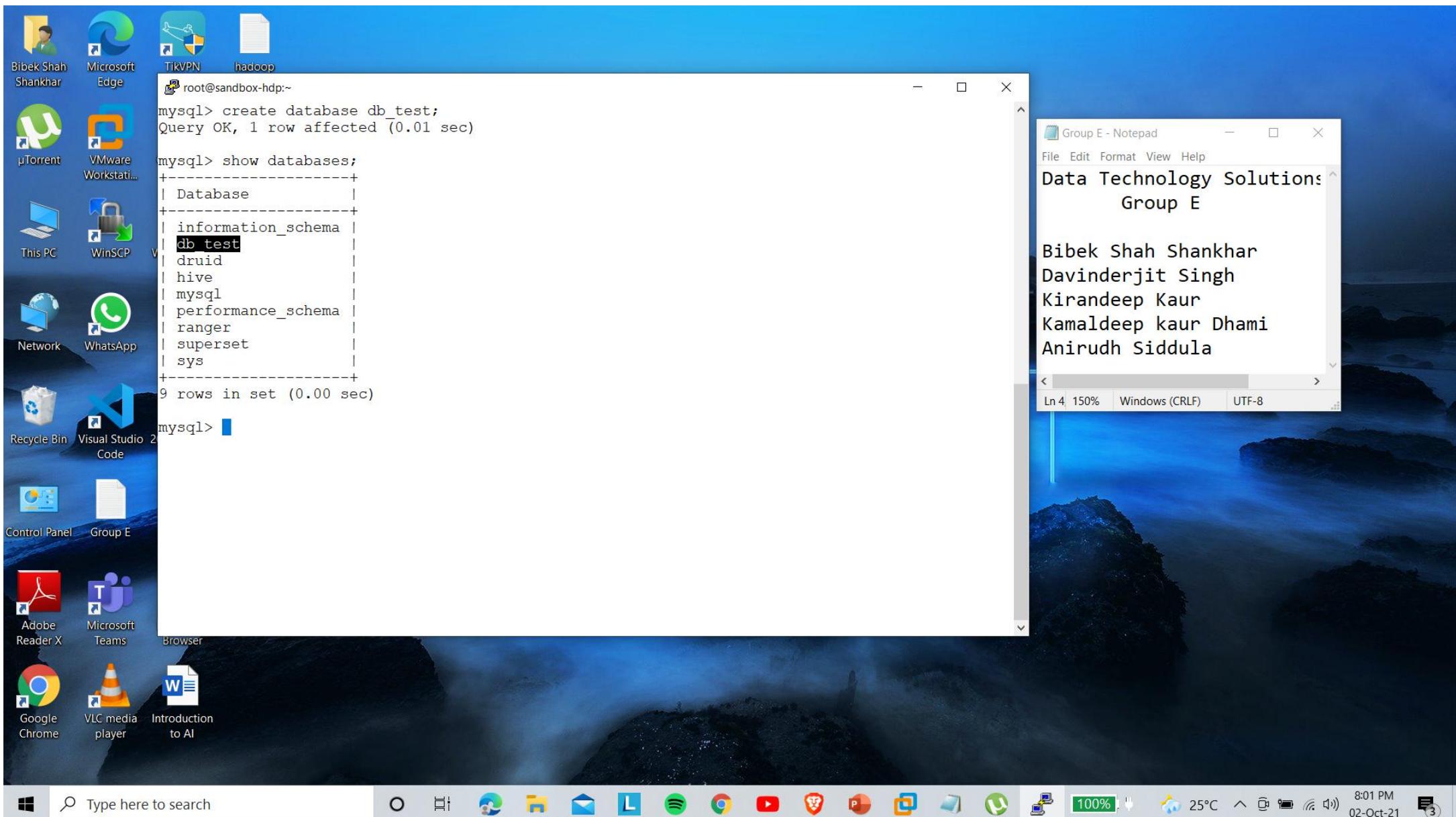


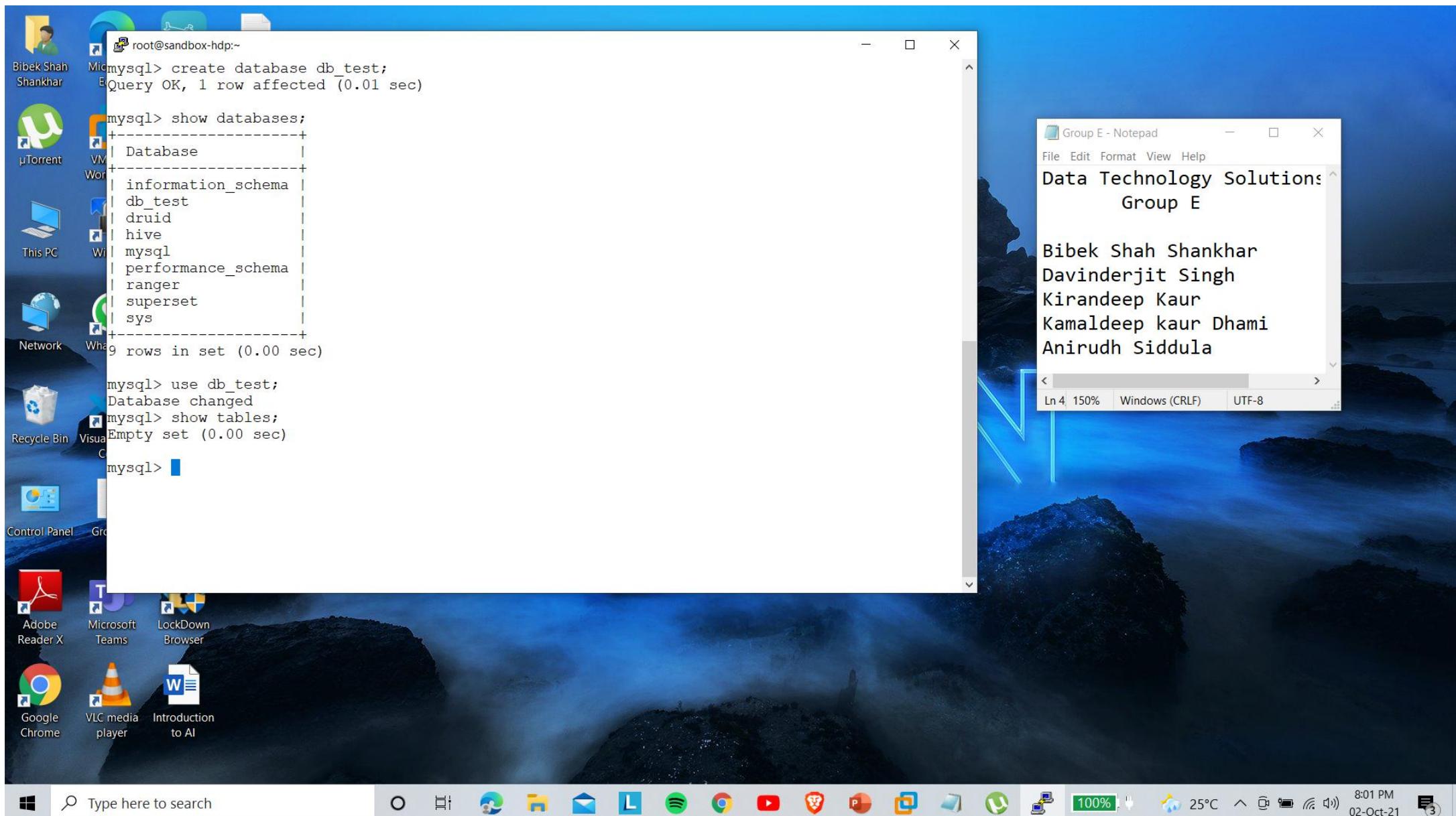


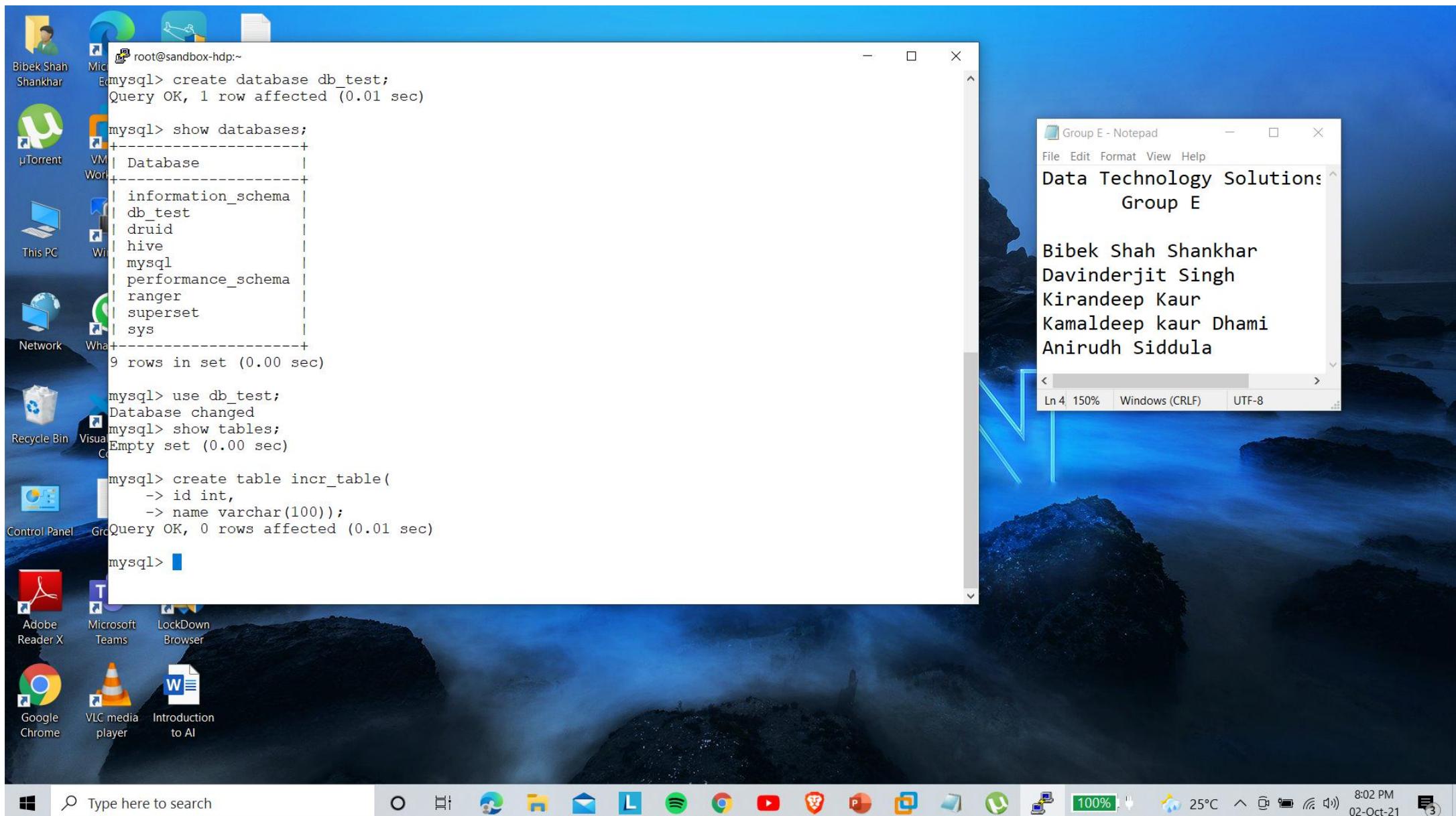


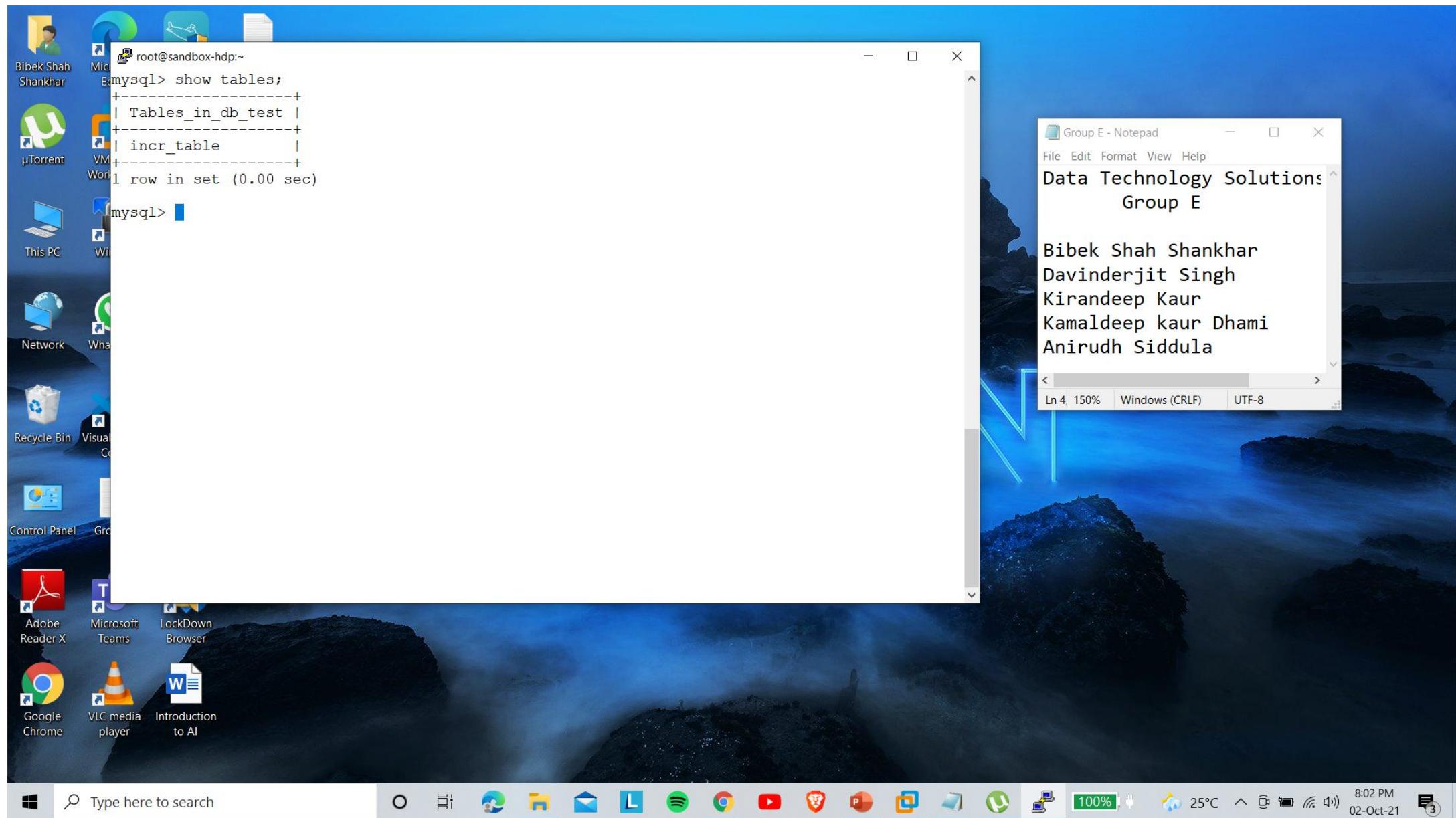


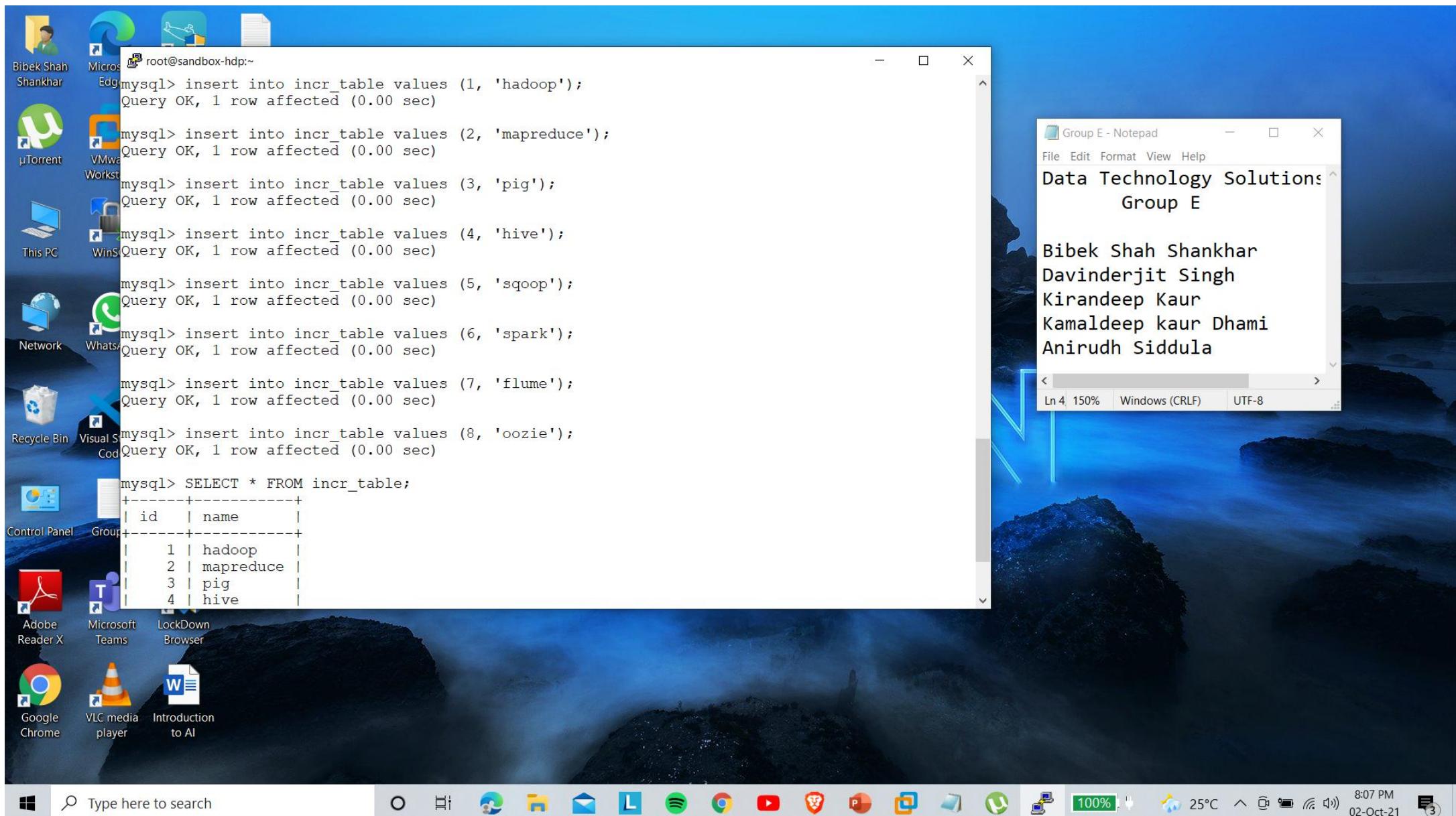












```
root@sandbox-hdp:~  
mysql> insert into incr_table values (1, 'hadoop');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (2, 'mapreduce');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (3, 'pig');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (4, 'hive');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (5, 'sqoop');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (6, 'spark');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (7, 'flume');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> insert into incr_table values (8, 'oozie');  
Query OK, 1 row affected (0.00 sec)  
  
mysql> SELECT * FROM incr_table;  
+----+-----+  
| id | name |  
+----+-----+  
| 1  | hadoop |  
| 2  | mapreduce |  
| 3  | pig |  
| 4  | hive |  
| 5  | sqoop |  
| 6  | spark |  
| 7  | flume |  
| 8  | oozie |  
+----+-----+  
8 rows in set (0.00 sec)  
  
mysql>
```

Group E - Notepad

File Edit Format View Help

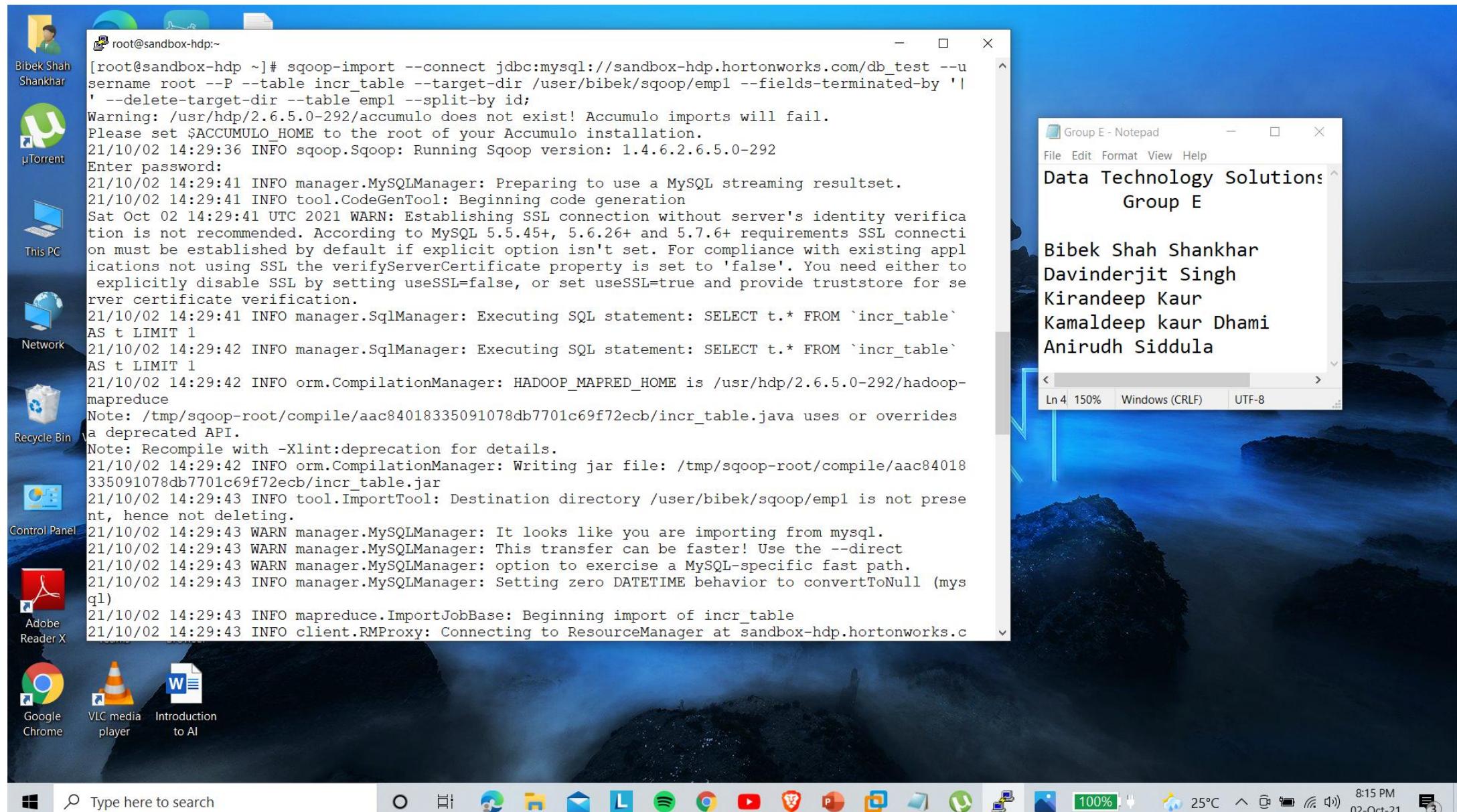
Data Technology Solutions

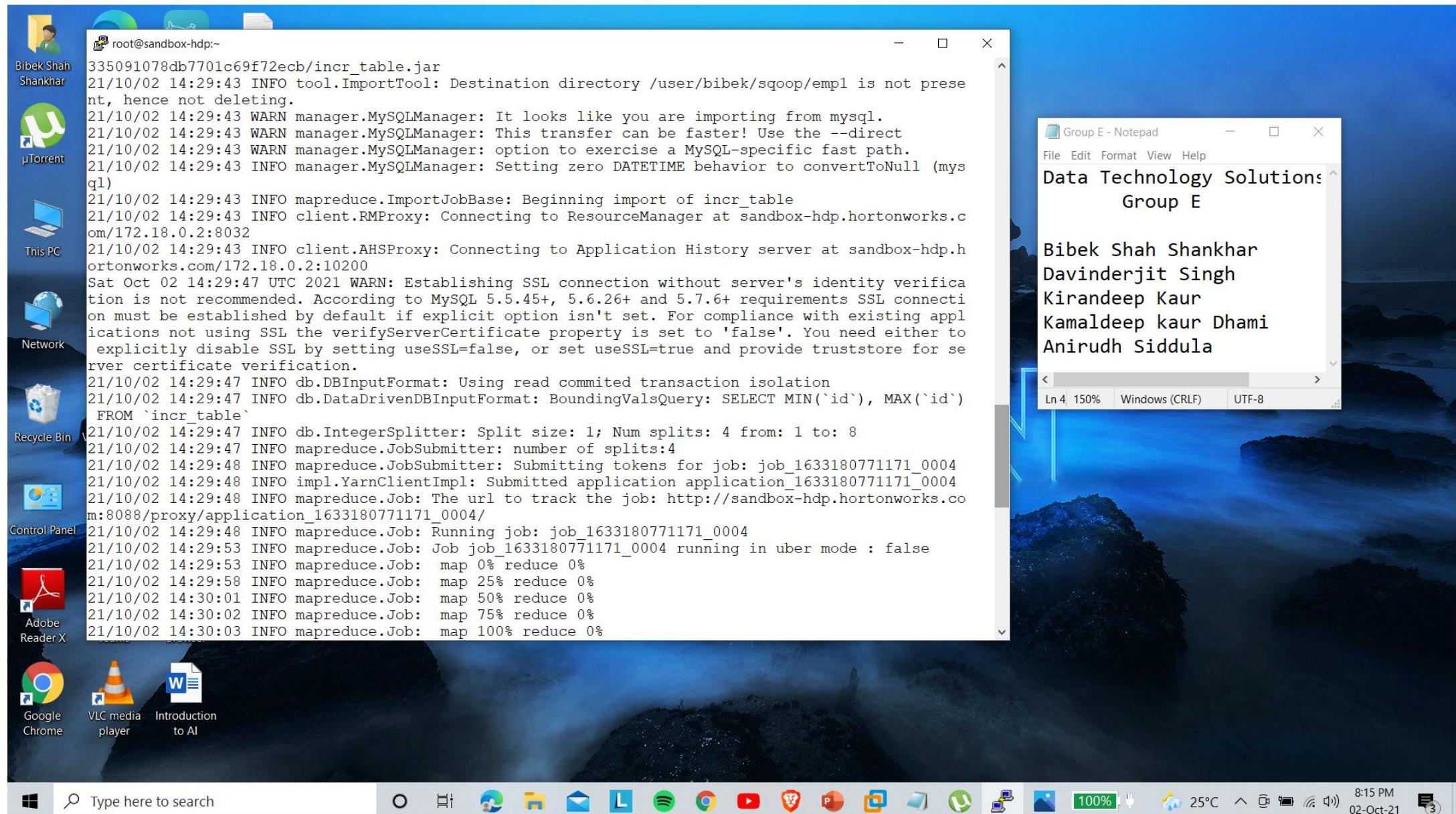
Group E

Bibek Shah Shankhar  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula

Ln 4 150% Windows (CRLF) UTF-8

The image shows a Windows desktop environment. On the left, there is a vertical taskbar with various icons for applications like Microsoft Edge, File Explorer, Mail, Spotify, Google Chrome, and VLC media player. Below the taskbar is a search bar with the placeholder "Type here to search". The main desktop area has a blue-themed wallpaper. A terminal window titled "root@sandbox-hdp:~" is open, showing a series of MySQL commands being run against a table named "incr\_table". The commands involve inserting data into the table and then selecting all rows from it. To the right of the terminal, a Notepad window titled "Group E - Notepad" is open, displaying a list of names under the heading "Data Technology Solutions Group E". The names listed are Bibek Shah Shankhar, Davinderjit Singh, Kirandeep Kaur, Kamaldeep kaur Dhami, and Anirudh Siddula. The Notepad window also shows standard file menu options like File, Edit, Format, View, and Help, along with status bar information like "Ln 4 150% Windows (CRLF) UTF-8". The system tray at the bottom right shows the date and time as "02-Oct-21 8:07 PM" and the battery level as "100%".





```
root@sandbox-hdp:~  
21/10/02 14:30:03 INFO mapreduce.Job: map 100% reduce 0%  
21/10/02 14:30:04 INFO mapreduce.Job: Job job_1633180771171_0004 completed successfully  
21/10/02 14:30:04 INFO mapreduce.Job: Counters: 30  
    File System Counters  
        FILE: Number of bytes read=0  
        FILE: Number of bytes written=684496  
        FILE: Number of read operations=0  
        FILE: Number of large read operations=0  
        FILE: Number of write operations=0  
        HDFS: Number of bytes read=393  
        HDFS: Number of bytes written=66  
        HDFS: Number of read operations=16  
        HDFS: Number of large read operations=0  
        HDFS: Number of write operations=8  
    Job Counters  
        Launched map tasks=4  
        Other local map tasks=4  
        Total time spent by all maps in occupied slots (ms)=14390  
        Total time spent by all reduces in occupied slots (ms)=0  
        Total time spent by all map tasks (ms)=14390  
        Total vcore-milliseconds taken by all map tasks=14390  
        Total megabyte-milliseconds taken by all map tasks=3597500  
    Map-Reduce Framework  
        Map input records=8  
        Map output records=8  
        Input split bytes=393  
        Spilled Records=0  
        Failed Shuffles=0  
        Merged Map outputs=0  
        GC time elapsed (ms)=672  
        CPU time spent (ms)=7020  
        Physical memory (bytes) snapshot=695828480  
        Virtual memory (bytes) snapshot=7817027584  
        Total committed heap usage (bytes)=178257920  
File Input Format Counters
```

```
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
```



root@sandbox-hdp:~

```
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=393
HDFS: Number of bytes written=66
HDFS: Number of read operations=16
HDFS: Number of large read operations=0
HDFS: Number of write operations=8

Job Counters
    Launched map tasks=4
    Other local map tasks=4
    Total time spent by all maps in occupied slots (ms)=14390
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=14390
    Total vcore-milliseconds taken by all map tasks=14390
    Total megabyte-milliseconds taken by all map tasks=3597500
```

```
Map-Reduce Framework
    Map input records=8
    Map output records=8
    Input split bytes=393
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=672
    CPU time spent (ms)=7020
    Physical memory (bytes) snapshot=695828480
    Virtual memory (bytes) snapshot=7817027584
    Total committed heap usage (bytes)=178257920
```

```
File Input Format Counters
```

```
    Bytes Read=0
```

```
File Output Format Counters
```

```
    Bytes Written=66
```

```
21/10/02 14:30:04 INFO mapreduce.ImportJobBase: Transferred 66 bytes in 20.4622 seconds (3.2255 byte/s)
```

```
21/10/02 14:30:04 INFO mapreduce.ImportJobBase: Retrieved 8 records.
```

```
[root@sandbox-hdp ~]#
```



Bibek Shah  
Shankhar



μTorrent



This PC



Network



Recycle Bin



Control Panel



Adobe Reader X



Google Chrome



VLC media player



Introduction to AI

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 4 150% Windows (CRLF) UTF-8



Type here to search



100%



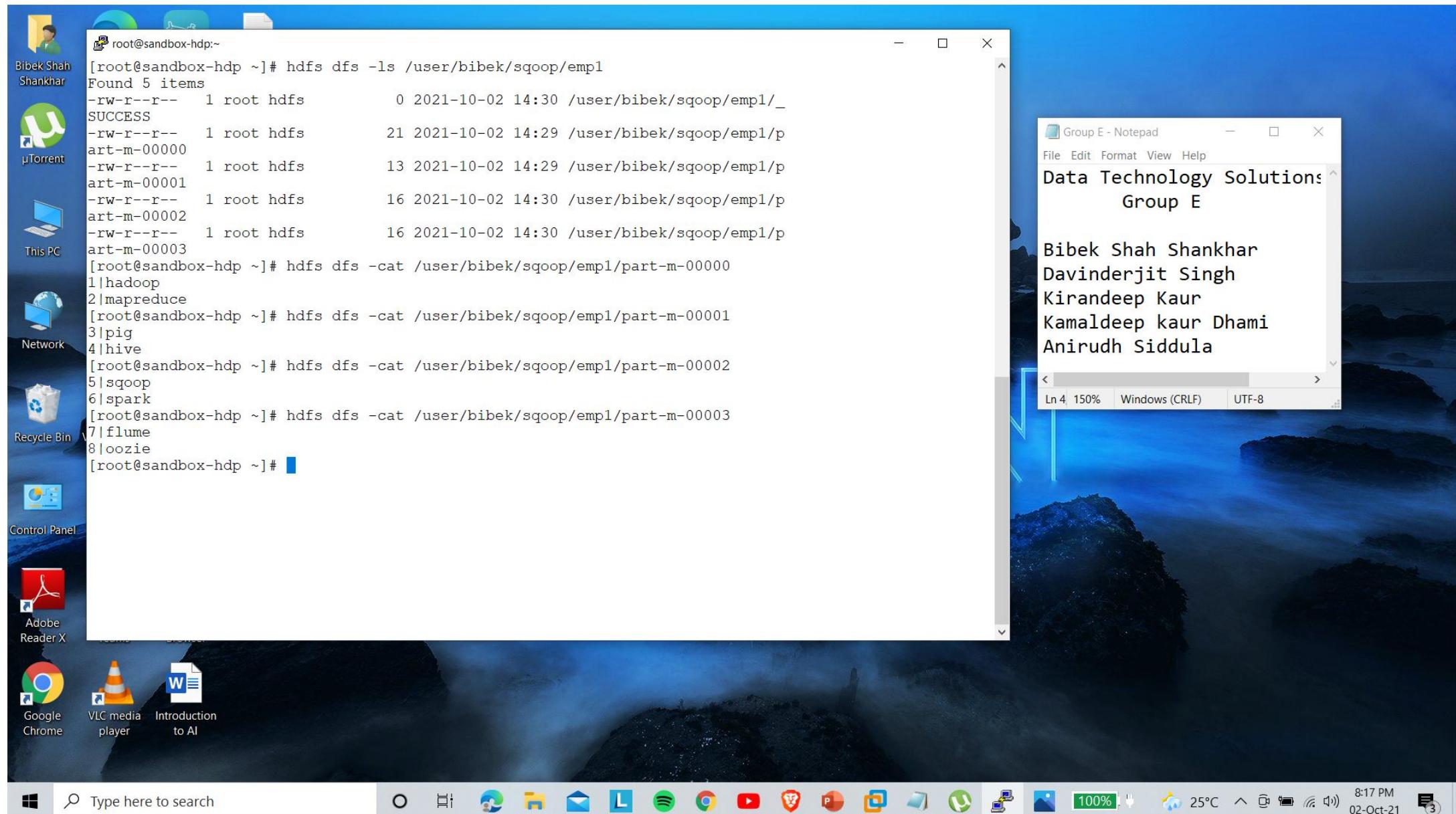
25°C

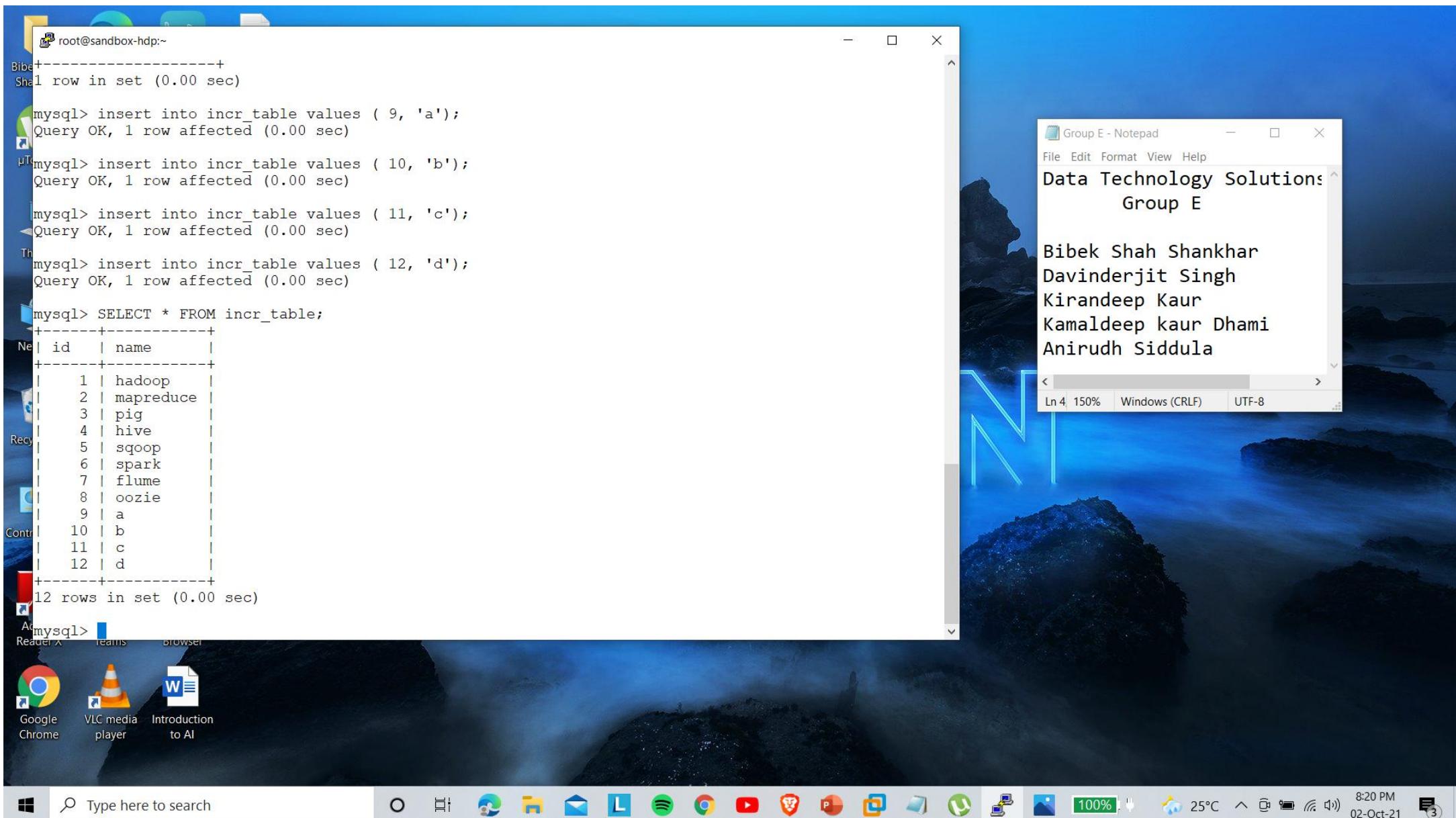


8:15 PM

02-Oct-21







```
root@sandbox-hdp:~# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/db_test --user name root --P --table incr_table --target-dir /user/bibek/sqoop/empl --fields-terminated-by '|' --table empl --split-by id --incremental append --check-column id --last-value 8;
Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
21/10/02 14:46:25 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292
Enter password:
21/10/02 14:46:29 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
21/10/02 14:46:29 INFO tool.CodeGenTool: Beginning code generation
Sat Oct 02 14:46:29 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
21/10/02 14:46:29 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `incr_table` AS t LIMIT 1
21/10/02 14:46:30 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `incr_table` AS t LIMIT 1
21/10/02 14:46:30 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/hdp/2.6.5.0-292/hadoop-mapreduce
Note: /tmp/sqoop-root/compile/a639a8ca5eee695dad17c3632a564546/incr_table.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
21/10/02 14:46:31 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-root/compile/a639a8ca5eee695dad17c3632a564546/incr_table.jar
21/10/02 14:46:31 INFO tool.ImportTool: Maximal id query for free form incremental import: SELECT MAX(`id`) FROM `incr_table`
21/10/02 14:46:31 INFO tool.ImportTool: Incremental import based on column `id`
21/10/02 14:46:31 INFO tool.ImportTool: Lower bound value: 8
21/10/02 14:46:31 INFO tool.ImportTool: Upper bound value: 12
21/10/02 14:46:31 WARN manager.MySQLManager: It looks like you are importing from mysql.
21/10/02 14:46:31 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
21/10/02 14:46:31 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
21/10/02 14:46:31 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
21/10/02 14:46:31 INFO mapreduce.ImportJobBase: Beginning import of incr_table
21/10/02 14:46:31 INFO client.RMProxy: Connecting to ResourceManager at sandbox-hdp.hortonworks.com/172.18.0.2:8032
21/10/02 14:46:31 INFO client.AHSProxy: Connecting to Application History server at sandbox-hdp.hortonworks.com/172.18.0.2:10200
```

Group E - Notepad

File Edit Format View Help

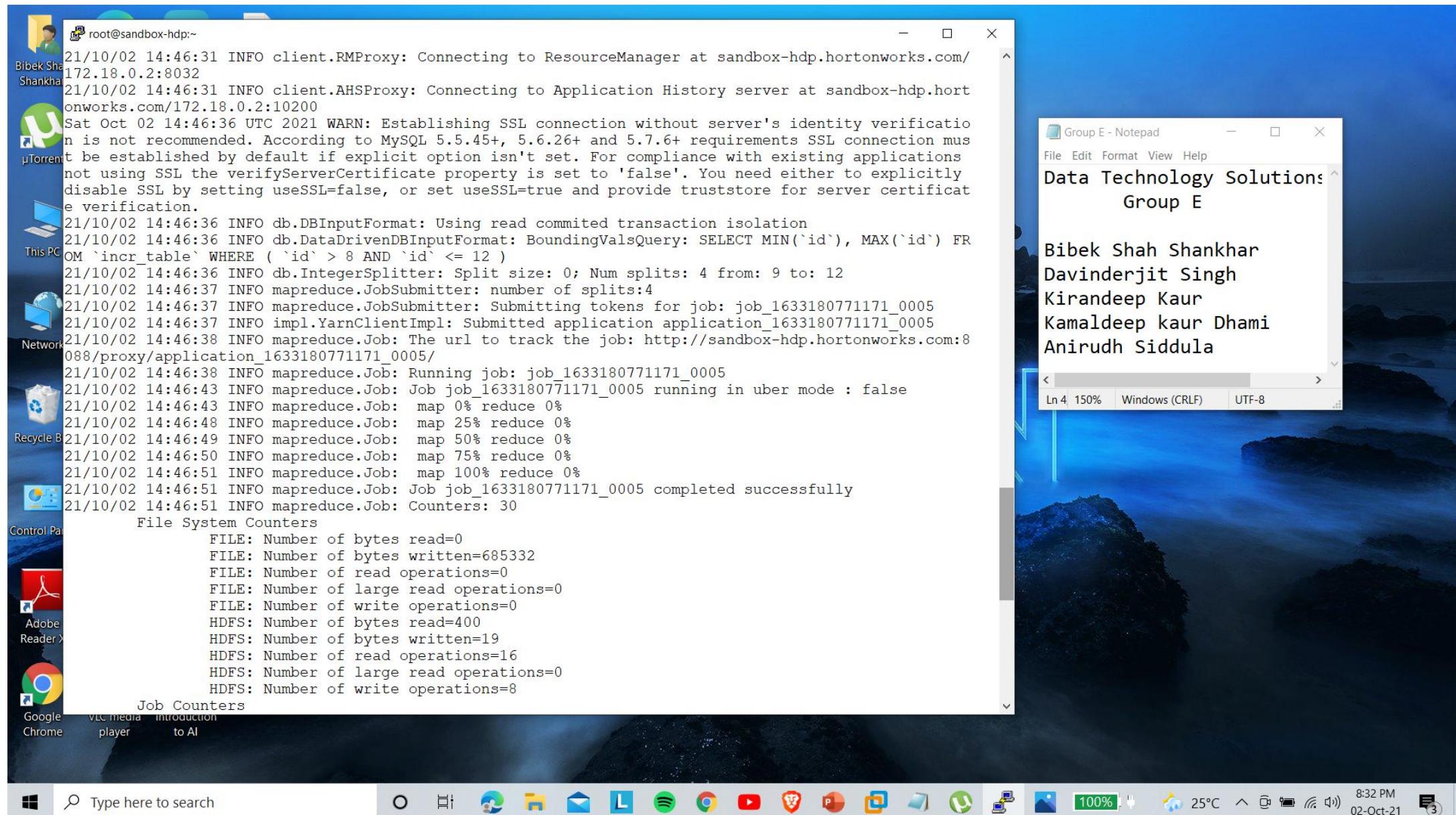
Data Technology Solutions

Group E

Bibek Shah Shankhar  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula

Ln 4 150% Windows (CRLF) UTF-8

Type here to search 8:32 PM 02-Oct-21



```
root@sandbox-hdp:~# 
HDFS: Number of bytes written=19
HDFS: Number of read operations=16
HDFS: Number of large read operations=0
HDFS: Number of write operations=8
Job Counters
    Launched map tasks=4
    Other local map tasks=4
    Total time spent by all maps in occupied slots (ms)=12525
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=12525
    Total vcore-milliseconds taken by all map tasks=12525
    Total megabyte-milliseconds taken by all map tasks=3131250
Map-Reduce Framework
    Map input records=4
    Map output records=4
    Input split bytes=400
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=503
    CPU time spent (ms)=5820
    Physical memory (bytes) snapshot=654598144
    Virtual memory (bytes) snapshot=7787544576
    Total committed heap usage (bytes)=181927936
File Input Format Counters
    Bytes Read=0
File Output Format Counters
    Bytes Written=19
21/10/02 14:46:51 INFO mapreduce.ImportJobBase: Transferred 19 bytes in 19.4182 seconds (0.9785 byte/s/sec)
21/10/02 14:46:51 INFO mapreduce.ImportJobBase: Retrieved 4 records.
21/10/02 14:46:51 INFO util.AppendUtils: Appending to directory empl
21/10/02 14:46:51 INFO util.AppendUtils: Using found partition 4
21/10/02 14:46:51 INFO tool.ImportTool: Incremental import complete! To run another incremental import of all data following this import, supply the following arguments:
21/10/02 14:46:51 INFO tool.ImportTool: --incremental append
21/10/02 14:46:51 INFO tool.ImportTool: --check-column id
21/10/02 14:46:51 INFO tool.ImportTool: --last-value 12
21/10/02 14:46:51 INFO tool.ImportTool: (Consider saving this with 'sqoop job --create')
[root@sandbox-hdp ~]# 
```

Group E - Notepad

File Edit Format View Help

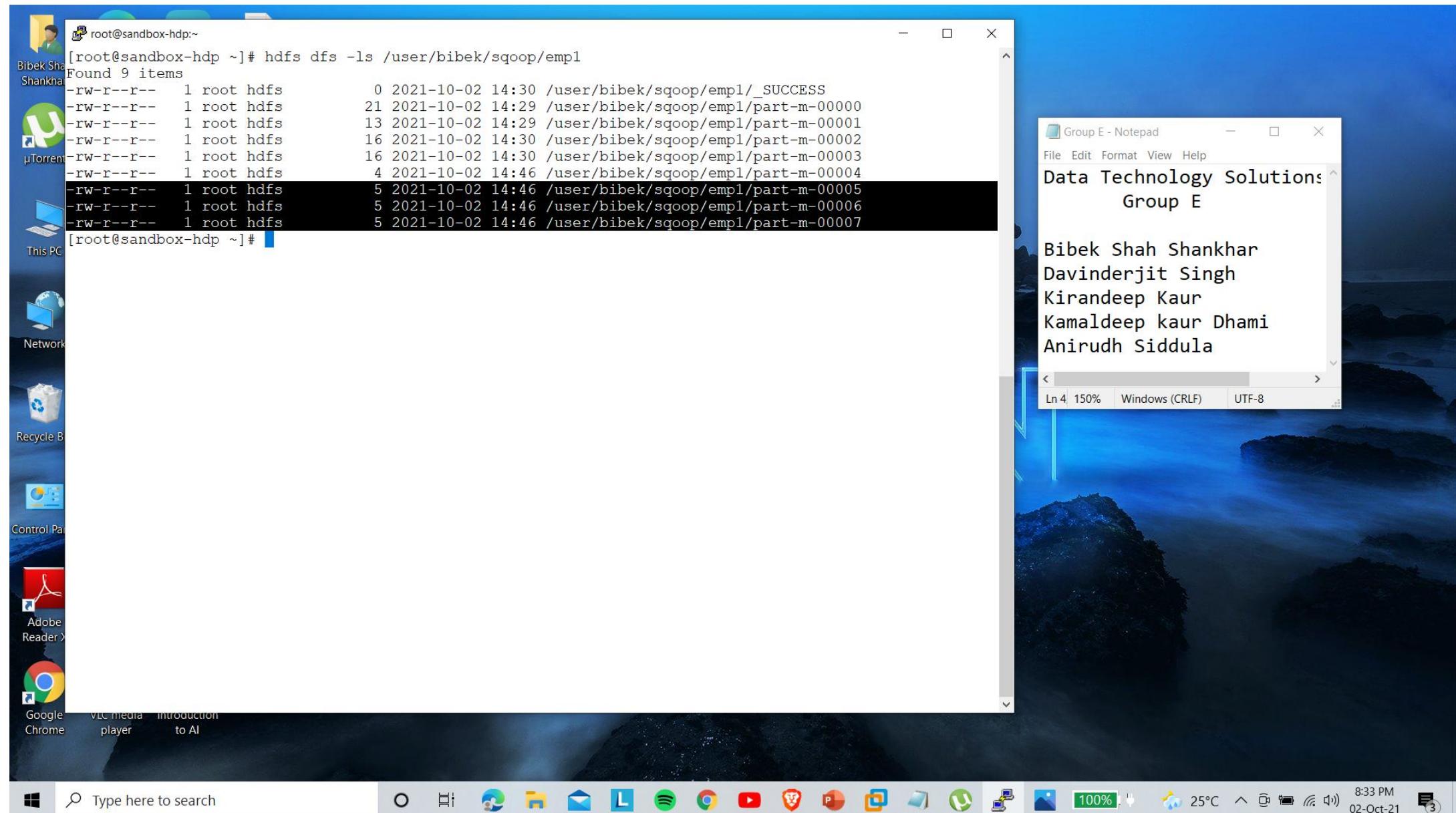
Data Technology Solutions

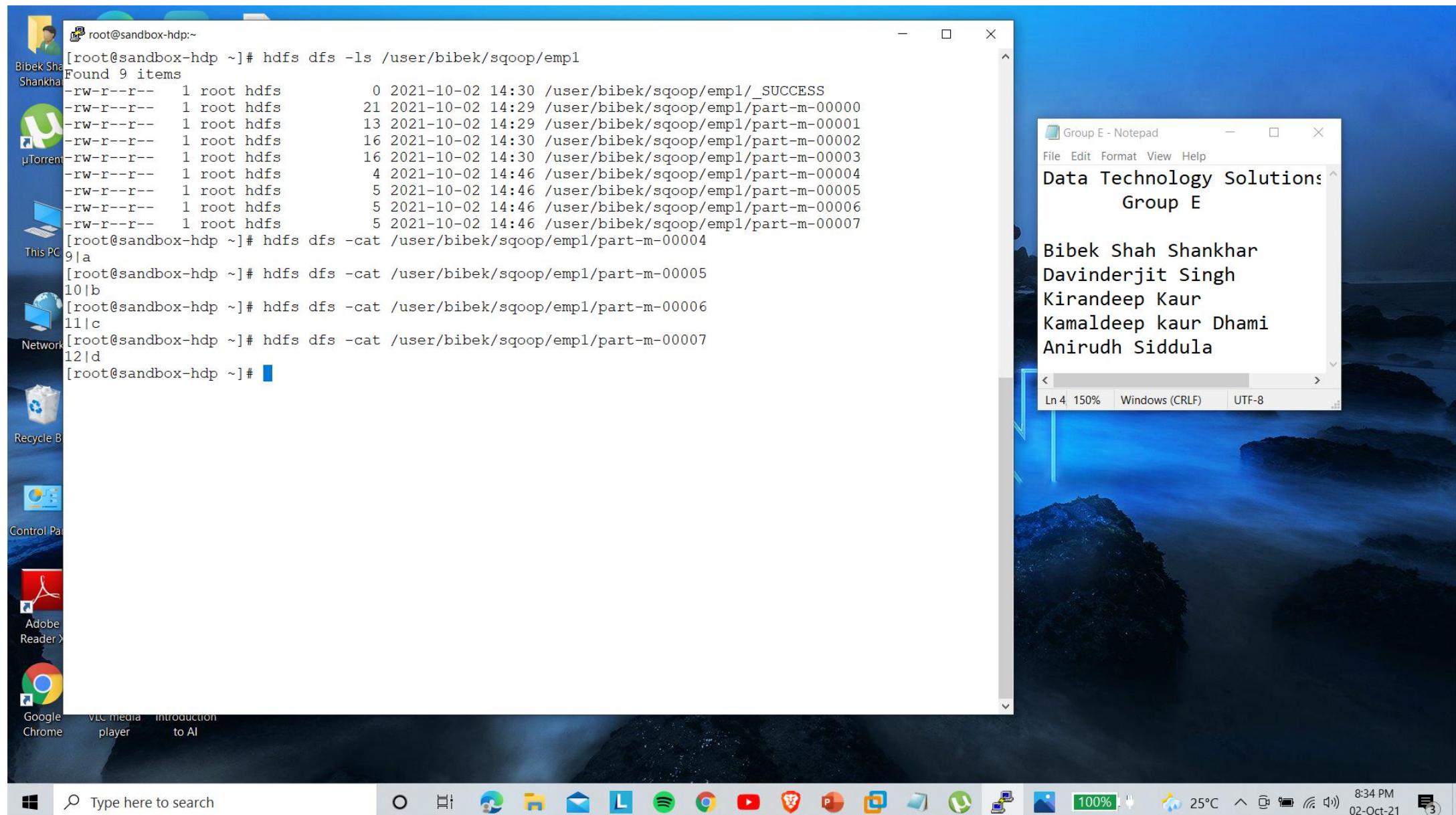
Group E

Bibek Shah Shankhar  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula

Ln 4 150% Windows (CRLF) UTF-8

Type here to search 8:32 PM 25°C 100% 02-Oct-21





```
Bibek Sh
Shankha
Database changed
mysql> SELECT * FROM incr_table;
+----+-----+
| id | name |
+----+-----+
| 1  | hadoop |
| 2  | mapreduce |
| 3  | pig |
| 4  | hive |
| 5  | sqoop |
| 6  | spark |
| 7  | flume |
| 8  | oozie |
| 9  | a |
| 10 | b |
| 11 | c |
| 12 | d |
+----+-----+
12 rows in set (0.00 sec)

mysql> select max(id) from incr_table;
+----+
| max(id) |
+----+
|      12 |
+----+
1 row in set (0.00 sec)

mysql> update incr_table set name='b1' where id=10;
Query OK, 1 row affected (0.01 sec)
Rows matched: 1  Changed: 1  Warnings: 0

mysql> SELECT * FROM incr_table;
+----+-----+
| id | name |
+----+-----+
| 1  | hadoop |
| 2  | mapreduce |
| 3  | pig |
|   | player   |
|   | to AI    |
+----+-----+
```

Group E - Notepad

Data Technology Solutions  
Group E

Bibek Shah Shankha  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula

Ln 4 150% Windows (CRLF) UTF-8

8:39 PM 25°C 100% 02-Oct-21

```
root@sandbox-hdp:~
```

```
Bibek Shah  
Shankhar  
| 8 | oozie |  
| 9 | a |  
| 10 | b |  
| 11 | c |  
| 12 | d |  
+-----+  
12 rows in set (0.00 sec)
```

```
mysql> select max(id) from incr_table;  
+-----+  
| max(id) |  
+-----+  
| 12 |  
+-----+  
1 row in set (0.00 sec)
```

```
mysql> update incr_table set name='b1' where id=10;  
Query OK, 1 row affected (0.01 sec)  
Rows matched: 1 Changed: 1 Warnings: 0
```

```
mysql> SELECT * FROM incr_table;  
+-----+  
| id | name |  
+-----+  
| 1 | hadoop |  
| 2 | mapreduce |  
| 3 | pig |  
| 4 | hive |  
| 5 | sqoop |  
| 6 | spark |  
| 7 | flume |  
| 8 | oozie |  
| 9 | a |  
| 10 | b1 |  
| 11 | c |  
| 12 | d |  
+-----+  
12 rows in set (0.01 sec)
```

```
Google Chrome player to AI
```

```
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 4 150% Windows (CRLF) UTF-8

```
[root@sandbox-hdp ~]# sqoop-import --connect jdbc:mysql://sandbox-hdp.hortonworks.com/db_test --username root --password bibek --table incr_table --target-dir /user/bibek/sqoop/empl --fields-terminated-by '|' --table emp1 --split-by id --incremental append --check-column id --last-value 8;
Warning: /usr/hdp/2.6.5.0-292/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
21/10/02 14:58:19 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292
Enter password:
21/10/02 14:58:25 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
21/10/02 14:58:25 INFO tool.CodeGenTool: Beginning code generation
Sat Oct 02 14:58:25 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
21/10/02 14:58:25 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `incr_table` AS t LIMIT 1
21/10/02 14:58:25 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `incr_table` AS t LIMIT 1
21/10/02 14:58:26 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/hdp/2.6.5.0-292/hadoop-mapreduce
Note: /tmp/sqoop-root/compile/9c040ac97c1ac5c3985f8be0033a70af/incr_table.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
21/10/02 14:58:27 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-root/compile/9c040ac97c1ac5c3985f8be0033a70af/incr_table.jar
21/10/02 14:58:27 INFO tool.ImportTool: Maximal id query for free form incremental import: SELECT MAX(`id`) FROM `incr_table`
21/10/02 14:58:27 INFO tool.ImportTool: Incremental import based on column `id`
21/10/02 14:58:27 INFO tool.ImportTool: Lower bound value: 8
21/10/02 14:58:27 INFO tool.ImportTool: Upper bound value: 12
21/10/02 14:58:27 WARN manager.MySQLManager: It looks like you are importing from mysql.
21/10/02 14:58:27 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
21/10/02 14:58:27 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
21/10/02 14:58:27 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
21/10/02 14:58:27 INFO mapreduce.ImportJobBase: Beginning import of incr_table
21/10/02 14:58:27 INFO client.RMProxy: Connecting to ResourceManager at sandbox-hdp.hortonworks.com/172.18.0.2:8032
21/10/02 14:58:28 INFO client.AHSProxy: Connecting to Application History server at sandbox-hdp.hortonworks.com/172.18.0.2:10200
Sat Oct 02 14:58:32 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
```

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 4 150% Windows (CRLF) UTF-8

root@sandbox-hdp:~

Note: /tmp/sqoop-root/compile/9c040ac97c1ac5c3985f8be0033a70af/incr\_table.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

```
21/10/02 14:58:27 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-root/compile/9c040ac97c1ac5c3985f8be0033a70af/incr_table.jar
21/10/02 14:58:27 INFO tool.ImportTool: Maximal id query for free form incremental import: SELECT MAX(`id`) FROM `incr_table`
21/10/02 14:58:27 INFO tool.ImportTool: Incremental import based on column `id`
21/10/02 14:58:27 INFO tool.ImportTool: Lower bound value: 8
21/10/02 14:58:27 INFO tool.ImportTool: Upper bound value: 12
21/10/02 14:58:27 WARN manager.MySQLManager: It looks like you are importing from mysql.
21/10/02 14:58:27 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
21/10/02 14:58:27 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
21/10/02 14:58:27 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
21/10/02 14:58:27 INFO mapreduce.ImportJobBase: Beginning import of incr_table
21/10/02 14:58:27 INFO client.RMProxy: Connecting to ResourceManager at sandbox-hdp.hortonworks.com/172.18.0.2:8032
21/10/02 14:58:28 INFO client.AHSProxy: Connecting to Application History server at sandbox-hdp.hortonworks.com/172.18.0.2:10200
Sat Oct 02 14:58:32 UTC 2021 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
21/10/02 14:58:32 INFO db.DBInputFormat: Using read committed transaction isolation
21/10/02 14:58:32 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `incr_table` WHERE (`id` > 8 AND `id` <= 12)
21/10/02 14:58:32 INFO db.IntegerSplitter: Split size: 0; Num splits: 4 from: 9 to: 12
21/10/02 14:58:32 INFO mapreduce.JobSubmitter: number of splits:4
21/10/02 14:58:32 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1633180771171_0006
21/10/02 14:58:32 INFO impl.YarnClientImpl: Submitted application application_1633180771171_0006
21/10/02 14:58:32 INFO mapreduce.Job: The url to track the job: http://sandbox-hdp.hortonworks.com:8088/proxy/application_1633180771171_0006/
21/10/02 14:58:32 INFO mapreduce.Job: Running job: job_1633180771171_0006
21/10/02 14:58:38 INFO mapreduce.Job: Job job_1633180771171_0006 running in uber mode : false
21/10/02 14:58:38 INFO mapreduce.Job: map 0% reduce 0%
21/10/02 14:58:44 INFO mapreduce.Job: map 50% reduce 0%
21/10/02 14:58:46 INFO mapreduce.Job: map 75% reduce 0%
21/10/02 14:58:47 INFO mapreduce.Job: map 100% reduce 0%
21/10/02 14:58:48 INFO mapreduce.Job: Job job_1633180771171_0006 completed successfully
```

Bibek Shah  
Shankha

μTorrent

This PC

Network

Recycle bin

Control Panel

Adobe Reader

Google Chrome

VLC media player

Introduction to AI

Type here to search

8:44 PM 02-Oct-21

Group E - Notepad

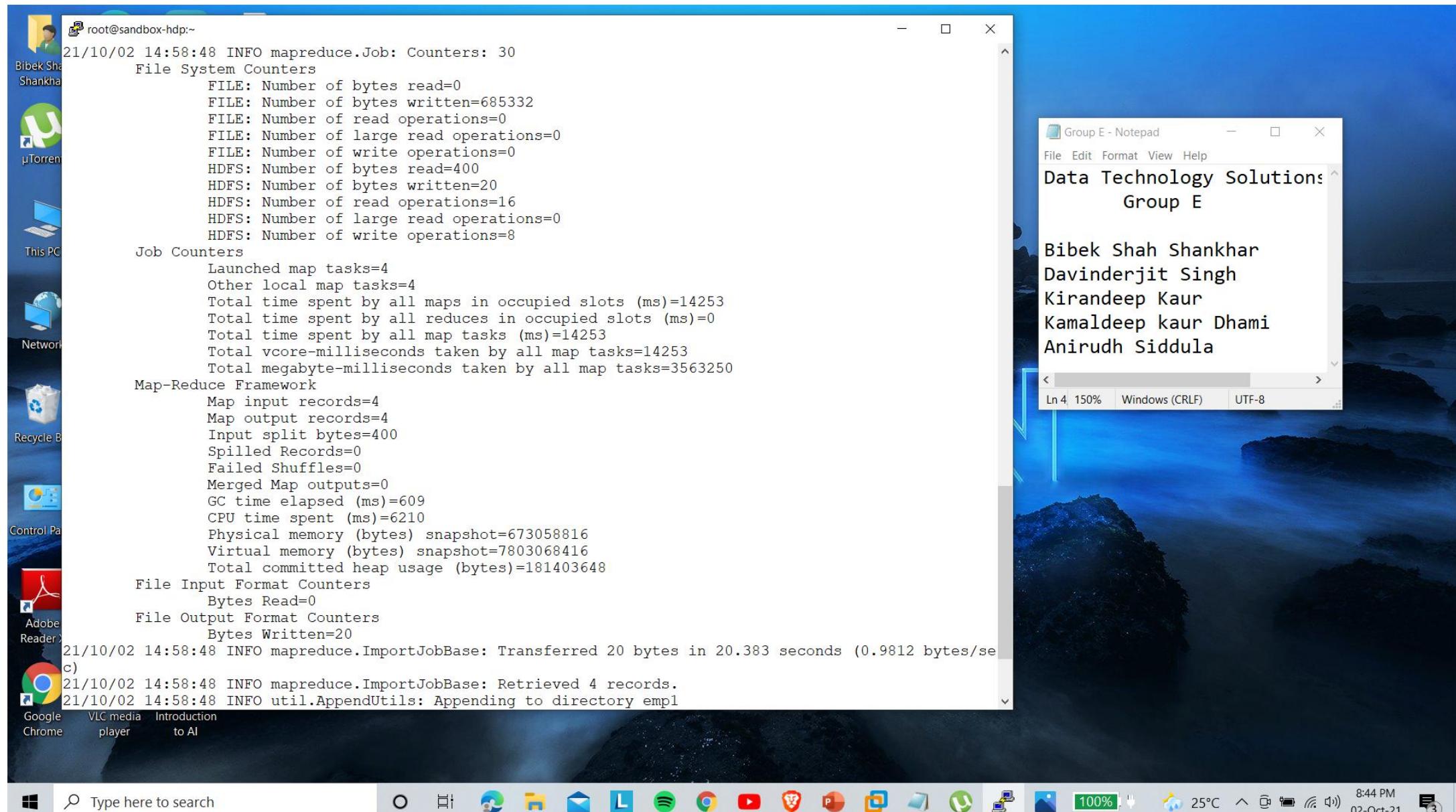
File Edit Format View Help

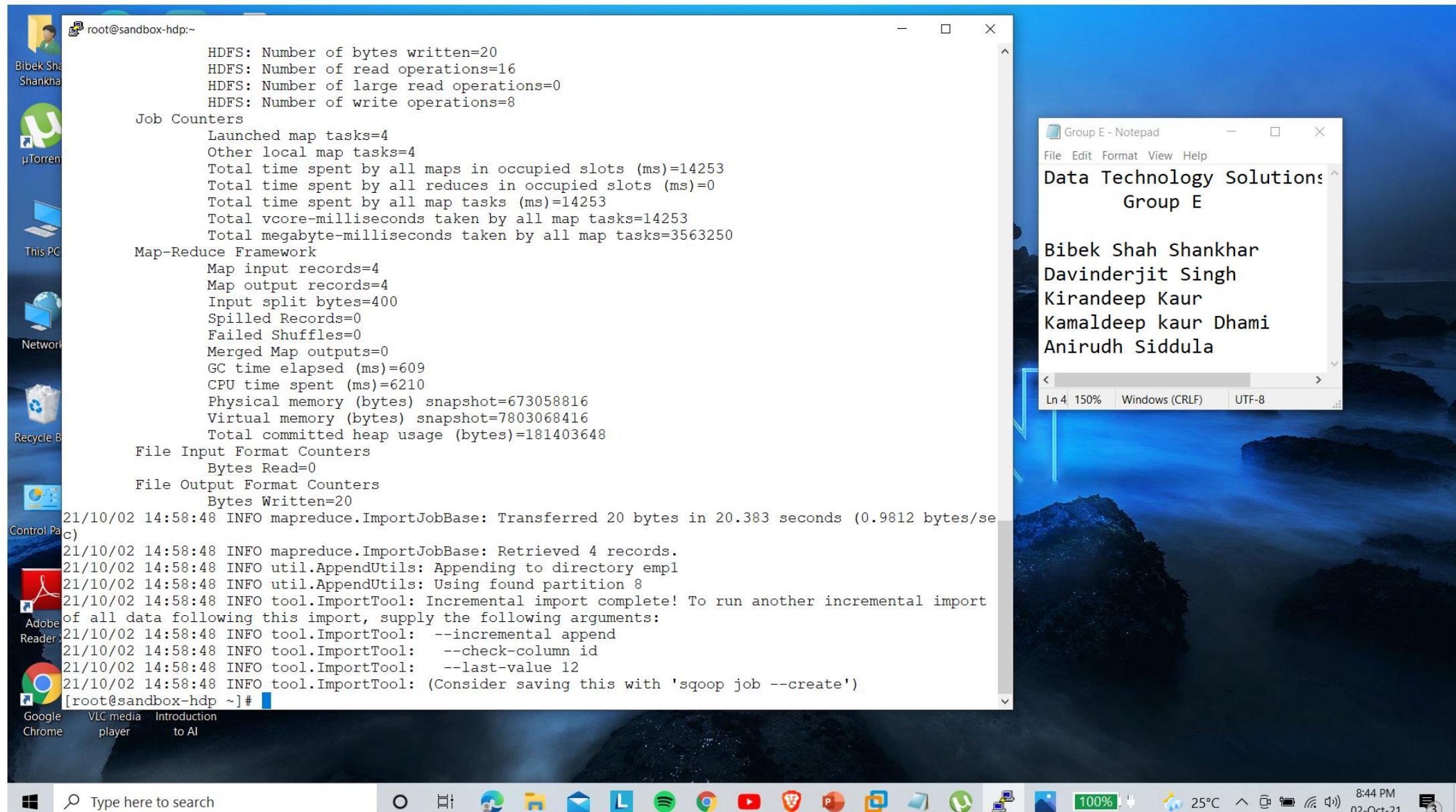
# Data Technology Solutions

## Group E

Bibek Shah  
Shankha  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula

Ln 4 150% Windows (CRLF) UTF-8





```
Bibek Sh [root@sandbox-hdp ~]# hdfs dfs -ls /user/bibek/sqoop/empl
Shankha Found 13 items
-rw-r--r-- 1 root hdfs          0 2021-10-02 14:30 /user/bibek/sqoop/empl/_SUCCESS
-rw-r--r-- 1 root hdfs         21 2021-10-02 14:29 /user/bibek/sqoop/empl/part-m-00000
-rw-r--r-- 1 root hdfs         13 2021-10-02 14:29 /user/bibek/sqoop/empl/part-m-00001
-rw-r--r-- 1 root hdfs         16 2021-10-02 14:30 /user/bibek/sqoop/empl/part-m-00002
-rw-r--r-- 1 root hdfs         16 2021-10-02 14:30 /user/bibek/sqoop/empl/part-m-00003
-rw-r--r-- 1 root hdfs          4 2021-10-02 14:46 /user/bibek/sqoop/empl/part-m-00004
-rw-r--r-- 1 root hdfs          5 2021-10-02 14:46 /user/bibek/sqoop/empl/part-m-00005
-rw-r--r-- 1 root hdfs          5 2021-10-02 14:46 /user/bibek/sqoop/empl/part-m-00006
-rw-r--r-- 1 root hdfs          5 2021-10-02 14:46 /user/bibek/sqoop/empl/part-m-00007
-rw-r--r-- 1 root hdfs          4 2021-10-02 14:58 /user/bibek/sqoop/empl/part-m-00008
-rw-r--r-- 1 root hdfs          6 2021-10-02 14:58 /user/bibek/sqoop/empl/part-m-00009
-rw-r--r-- 1 root hdfs          5 2021-10-02 14:58 /user/bibek/sqoop/empl/part-m-00010
-rw-r--r-- 1 root hdfs          5 2021-10-02 14:58 /user/bibek/sqoop/empl/part-m-00011
[root@sandbox-hdp ~]# hdfs dfs -cat /user/bibek/sqoop/empl/00000
cat: `/user/bibek/sqoop/empl/00000': No such file or directory
[root@sandbox-hdp ~]# hdfs dfs -cat /user/bibek/sqoop/empl/part-m-00000
1|hadoop
2|mapreduce
[root@sandbox-hdp ~]# hdfs dfs -cat /user/bibek/sqoop/empl/part-m-00009
10|b1
[root@sandbox-hdp ~]#
```

Group E - Notepad

Data Technology Solutions

Group E

Bibek Shah Shankha

Davinderjit Singh

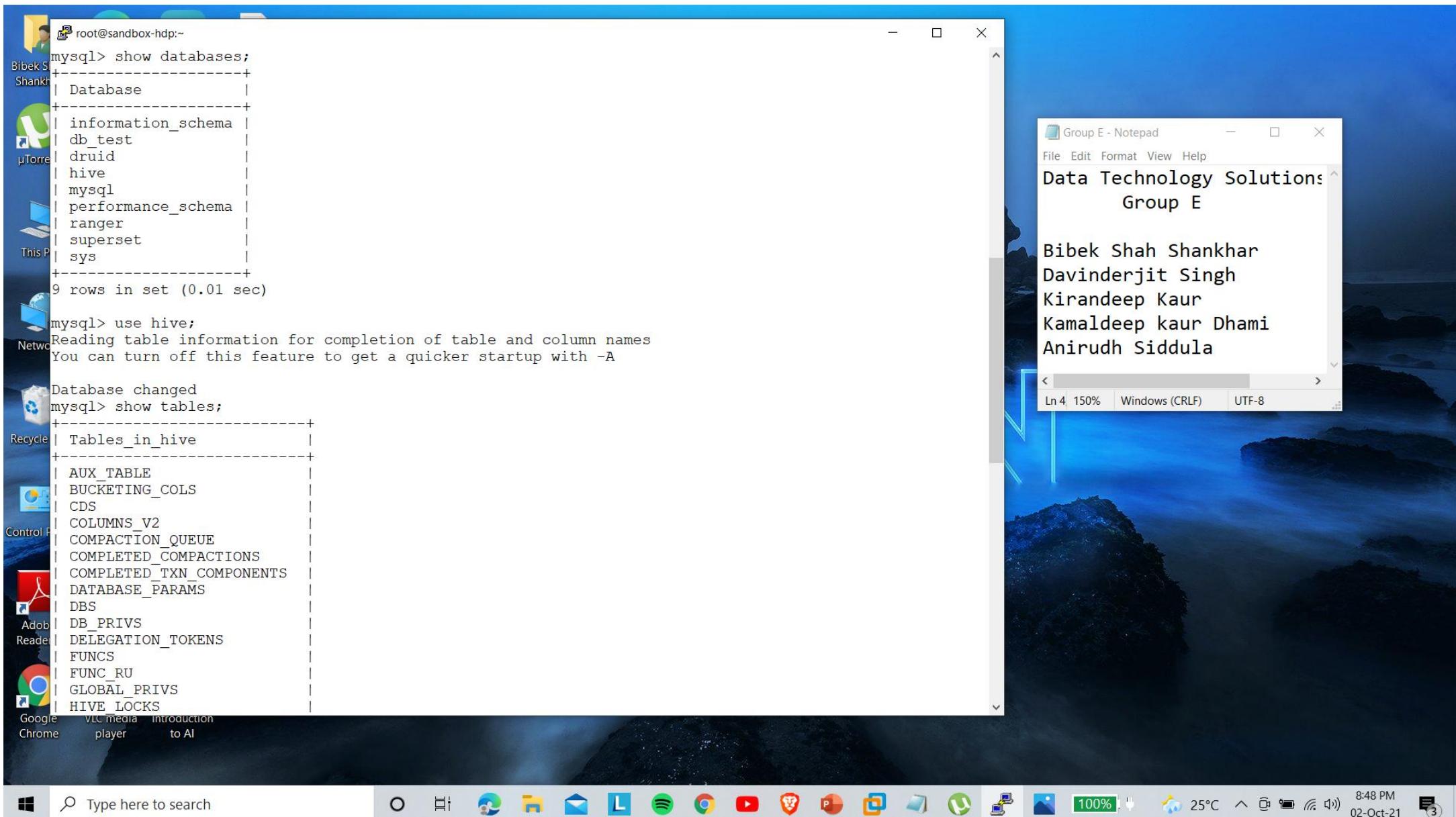
Kirandeep Kaur

Kamaldeep kaur Dhami

Anirudh Siddula

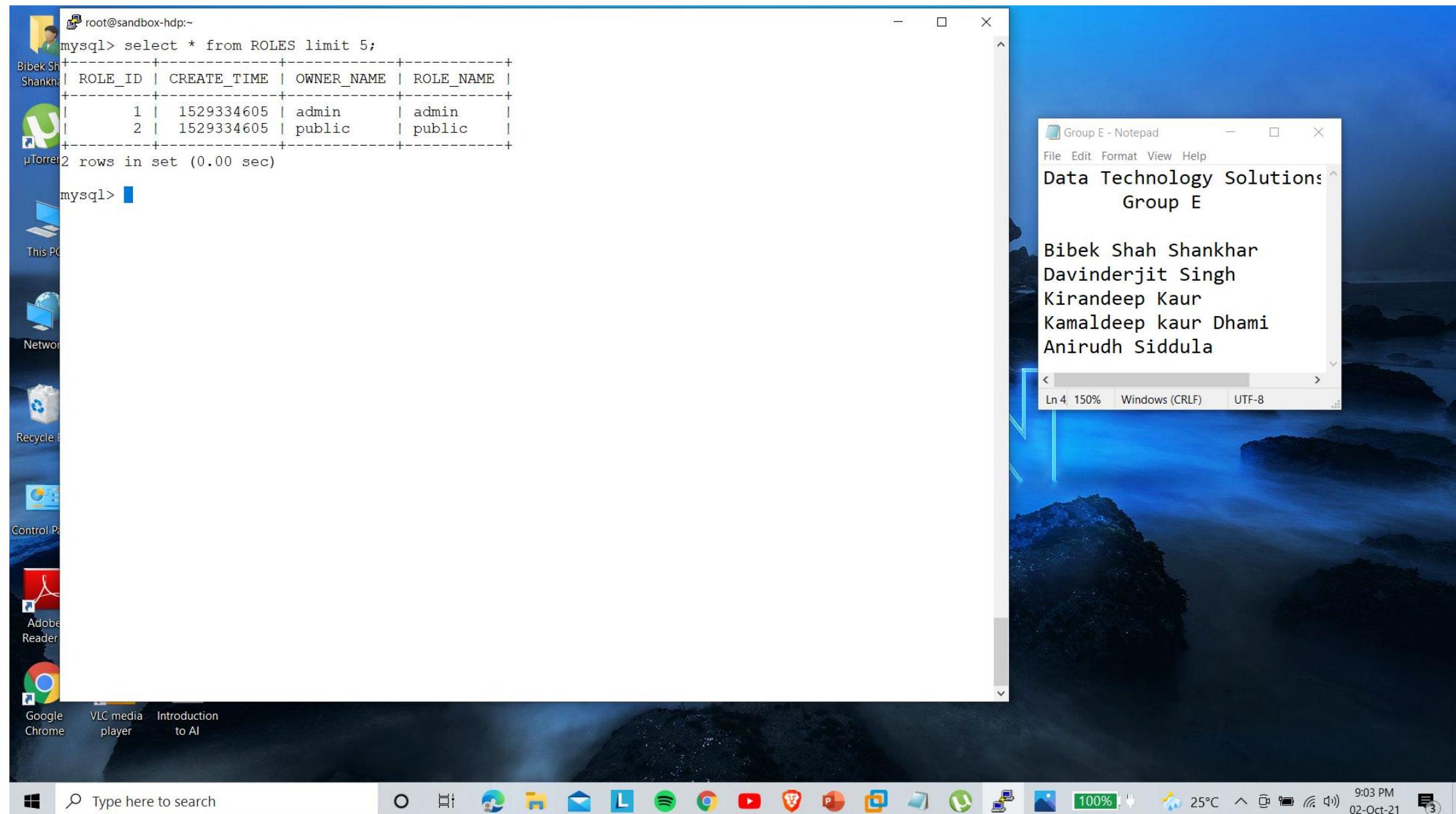
Ln 4 150% Windows (CRLF) UTF-8

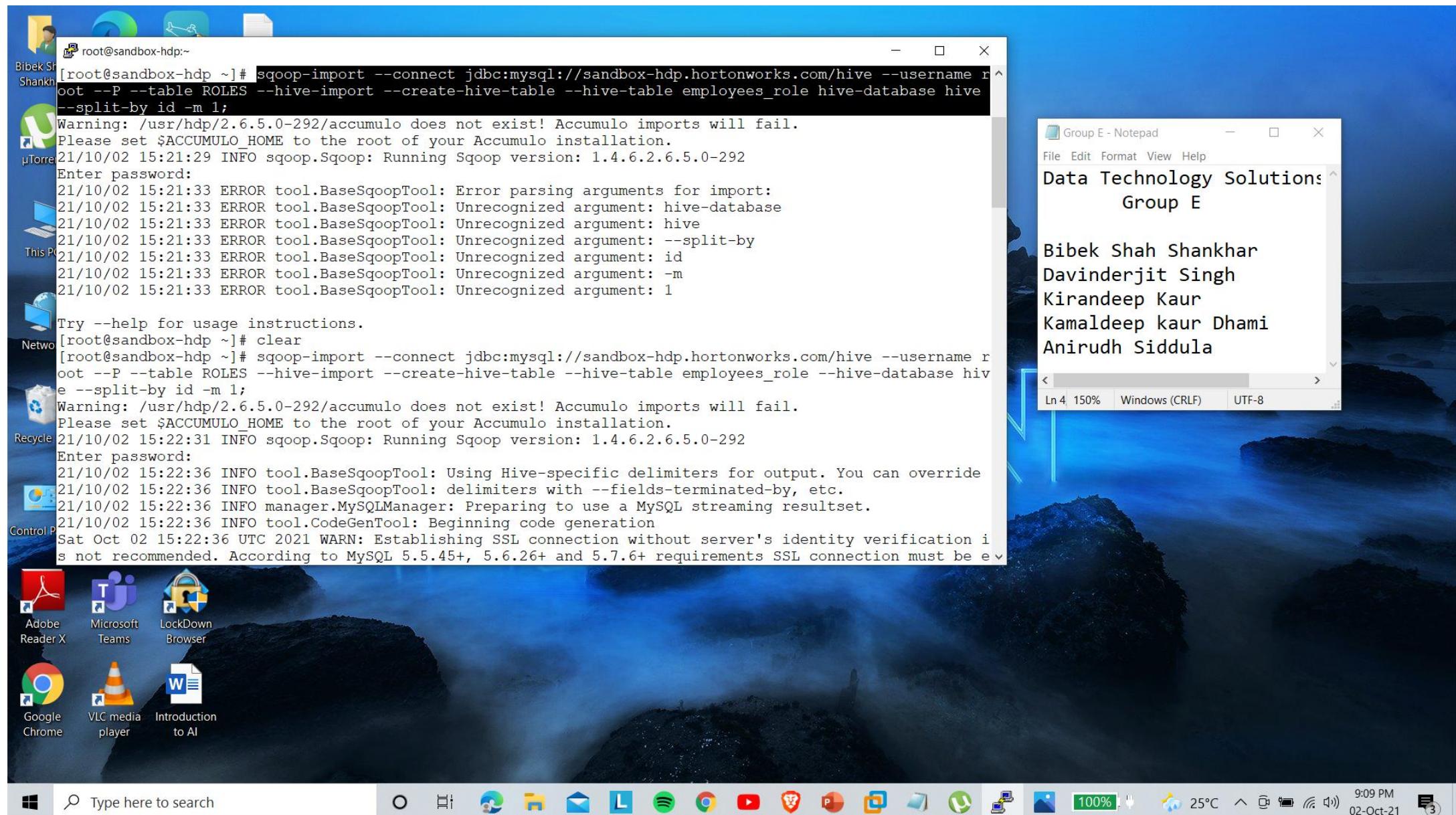
```
root@sandbox-hdp:~  
Bibek Shah  
Shankhar  
μTorre  
This PC  
Network  
Recycle  
Control Panel  
Adobe Reader  
Google Chrome  
VLC Media player  
Introduction to AI  
mysql> show databases;  
+-----+  
| Database |  
+-----+  
| information_schema |  
| db_test |  
| druid |  
| hive |  
| mysql |  
| performance_schema |  
| ranger |  
| superset |  
| sys |  
+-----+  
9 rows in set (0.01 sec)  
  
mysql> use hive;  
Reading table information for completion of table and column names  
You can turn off this feature to get a quicker startup with -A  
  
Database changed  
mysql> show tables;  
+-----+  
| Tables_in_hive |  
+-----+  
| AUX_TABLE |  
| BUCKETING_COLS |  
| CDS |  
| COLUMNS_V2 |  
| COMPACTION_QUEUE |  
| COMPLETED_COMPACTIONS |  
| COMPLETED_TXN_COMPONENTS |  
| DATABASE_PARAMS |  
| DBS |  
| DB_PRIVS |  
| DELEGATION_TOKENS |  
| FUNCS |  
| FUNC_RU |  
| GLOBAL_PRIVS |  
| HIVE_LOCKS |  
+-----+  
Ln 4 150% Windows (CRLF) UTF-8  
8:48 PM 02-Oct-21 3  
Type here to search
```

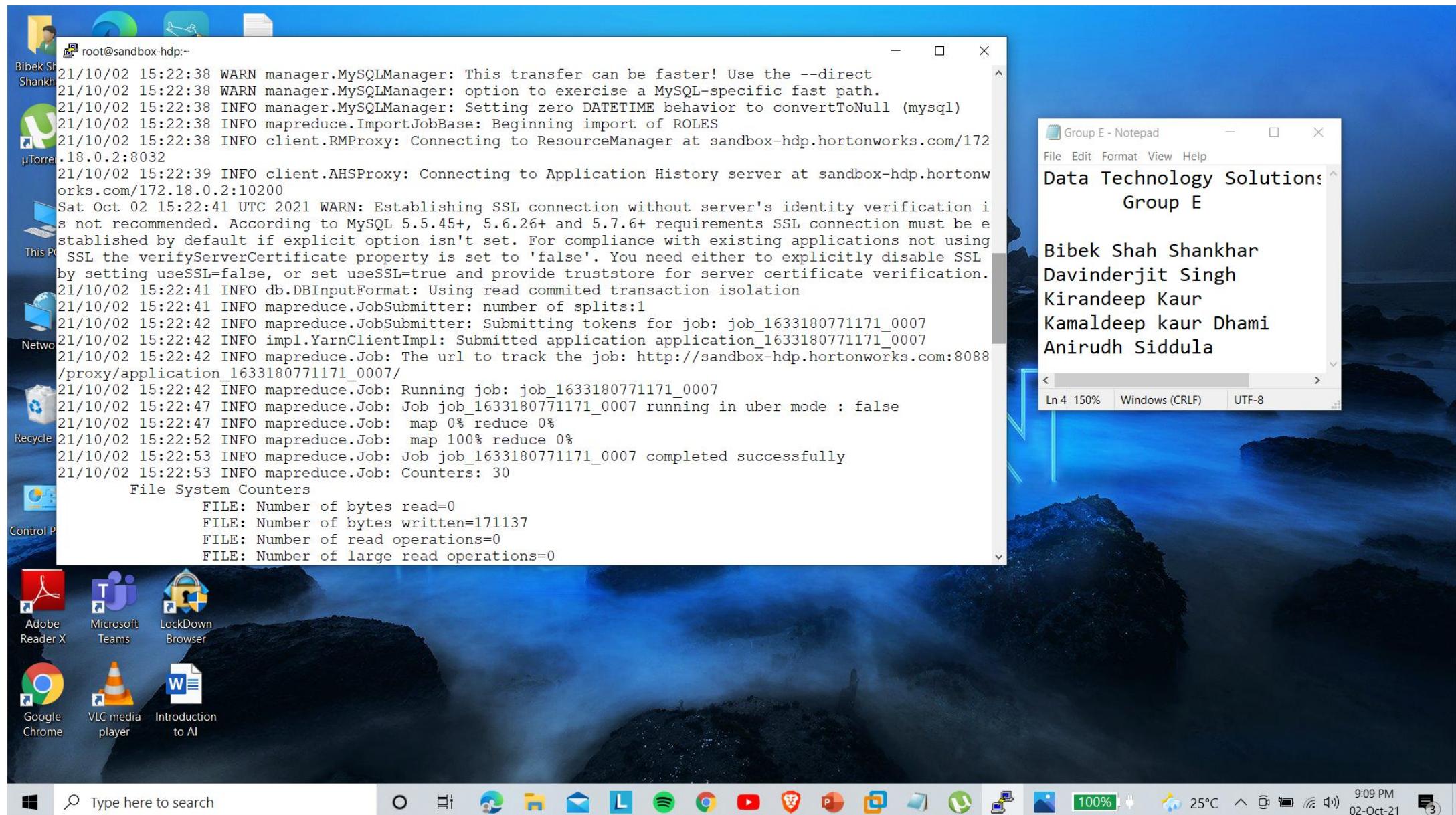


The image shows a Windows desktop environment. On the left, a terminal window titled 'root@sandbox-hdp:~' displays MySQL command-line interface output. It lists databases like 'information\_schema', 'db\_test', 'druid', 'hive', 'mysql', 'performance\_schema', 'ranger', 'superset', and 'sys'. It also shows the result of the 'use hive;' command and the output of 'show tables;'. On the right, a Notepad window titled 'Group E - Notepad' contains a document with the title 'Data Technology Solutions' and 'Group E'. The document lists five names: Bibek Shah, Shankhar, Davinderjit Singh, Kirandeep Kaur, Kamaldeep kaur Dhami, and Anirudh Siddula. The desktop background is a dark blue landscape image. The taskbar at the bottom includes icons for various applications like VLC, Spotify, and Google Chrome, along with system status icons.

```
root@sandbox-hdp:~| 109 | NEXT_TXN_ID  
| NOTIFICATION_LOG  
| NOTIFICATION_SEQUENCE  
| NUCLEUS_TABLES  
| PARTITIONS  
| PARTITION_EVENTS  
| PARTITION_KEYS  
| PARTITION_KEY_VALS  
| PARTITION_PARAMS  
| PART_COL_PRIVS  
| PART_COL_STATS  
| PART_PRIVS  
| ROLES  
| ROLE_MAP  
| SDS  
| SD_PARAMS  
| SEQUENCE_TABLE  
| SERDES  
| SERDE_PARAMS  
| SKEWED_COL_NAMES  
| SKEWED_COL_VALUE_LOC_MAP  
| SKEWED_STRING_LIST  
| SKEWED_STRING_LIST_VALUES  
| SKEWED_VALUES  
| SORT_COLS  
| TABLE_PARAMS  
| TAB_COL_STATS  
| TBLS  
| TBL_COL_PRIVS  
| TBL_PRIVS  
| TXNS  
| TXN_COMPONENTS  
| TYPES  
| TYPE_FIELDS  
| VERSION  
| WRITE_SET  
+-----+  
57 rows in set (0.00 sec)  
  
mysql>   
  
File Edit Format View Help  
Data Technology Solutions  
Group E  
Bibek Shah Shankhar  
Davinderjit Singh  
Kirandeep Kaur  
Kamaldeep kaur Dhami  
Anirudh Siddula  
Ln 4 150% Windows (CRLF) UTF-8  
  
Slide 114 of 117 Notes Comments 125% 8:48 PM 25°C 100% 02-Oct-21 3  
Type here to search O L S C G M L 100% 25°C 02-Oct-21 3
```







```
Bibek Shah  
Shankhar  
μTorren  
This PC  
Network  
Recycle  
Control Panel  
Adobe Reader X  
Microsoft Teams  
LockDown Browser  
Google Chrome  
VLC media player  
Introduction to AI  
root@sandbox-hdp:~  
HDFS: Number of read operations=4  
HDFS: Number of large read operations=0  
HDFS: Number of write operations=2  
Job Counters  
Launched map tasks=1  
Other local map tasks=1  
Total time spent by all maps in occupied slots (ms)=2439  
Total time spent by all reduces in occupied slots (ms)=0  
Total time spent by all map tasks (ms)=2439  
Total vcore-milliseconds taken by all map tasks=2439  
Total megabyte-milliseconds taken by all map tasks=609750  
Map-Reduce Framework  
Map input records=2  
Map output records=2  
Input split bytes=87  
Spilled Records=0  
Failed Shuffles=0  
Merged Map outputs=0  
GC time elapsed (ms)=61  
CPU time spent (ms)=1580  
Physical memory (bytes) snapshot=175566848  
Virtual memory (bytes) snapshot=1957707776  
Total committed heap usage (bytes)=45088768  
File Input Format Counters  
Bytes Read=0  
File Output Format Counters  
Bytes Written=52  
21/10/02 15:22:53 INFO mapreduce.ImportJobBase: Transferred 52 bytes in 15.0351 seconds (3.4586 bytes/sec)  
21/10/02 15:22:53 INFO mapreduce.ImportJobBase: Retrieved 2 records.
```

```
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 4 150% Windows (CRLF) UTF-8
```

