**ASSIGNMENT - 35.2**

**Problem Statement:**

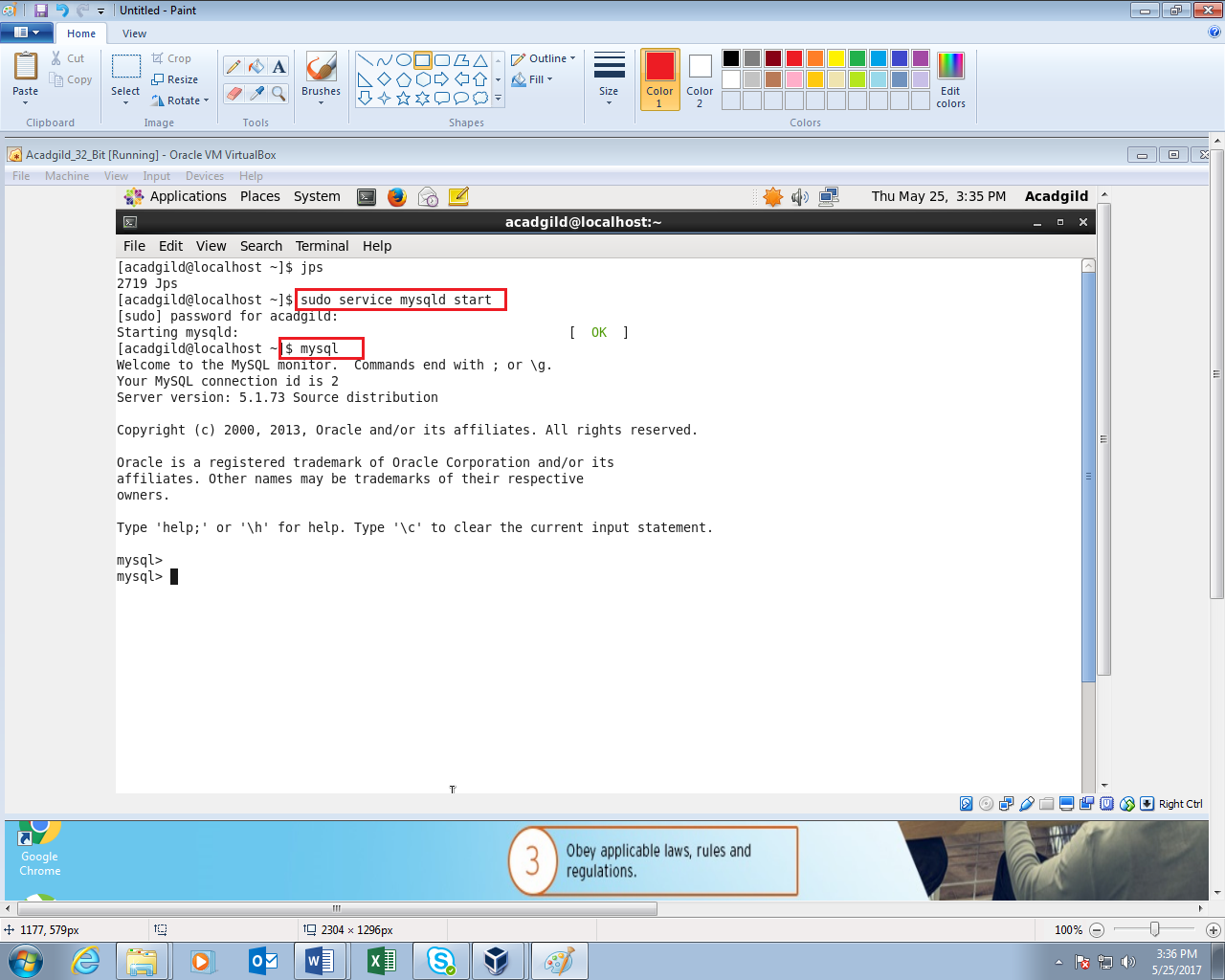
* Explain and perform Sqoop Incremental Import operation to load data from Mysql to HDFS using Sqoop.
* Explain the procedures performed, Share the screenshots of commands and results for the same.

**Solution:**

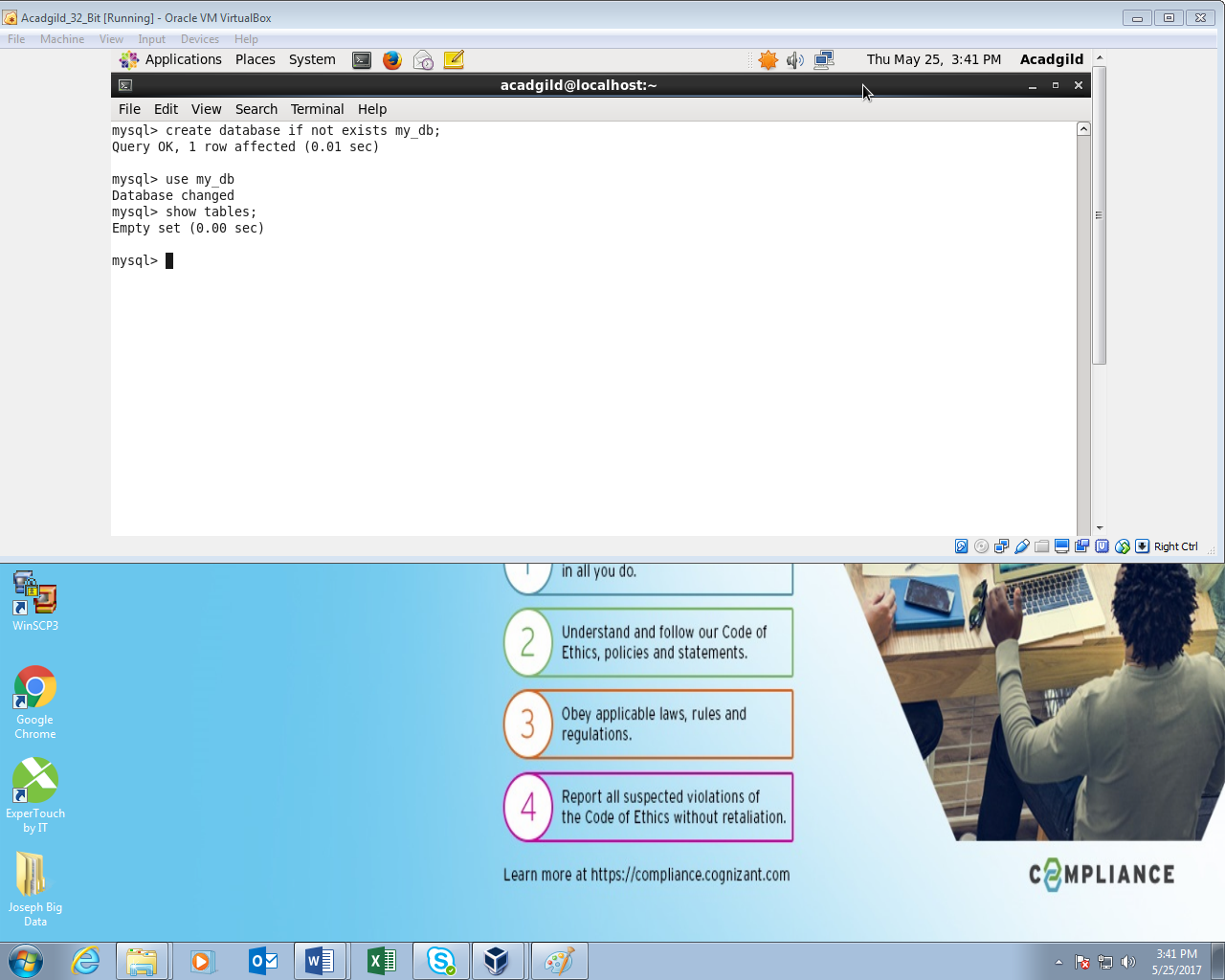
* **Incremental import operation using sqoop:**

**Creating table ‘emp’ in MYSQL:**

First we have to login to your MySQL shell.

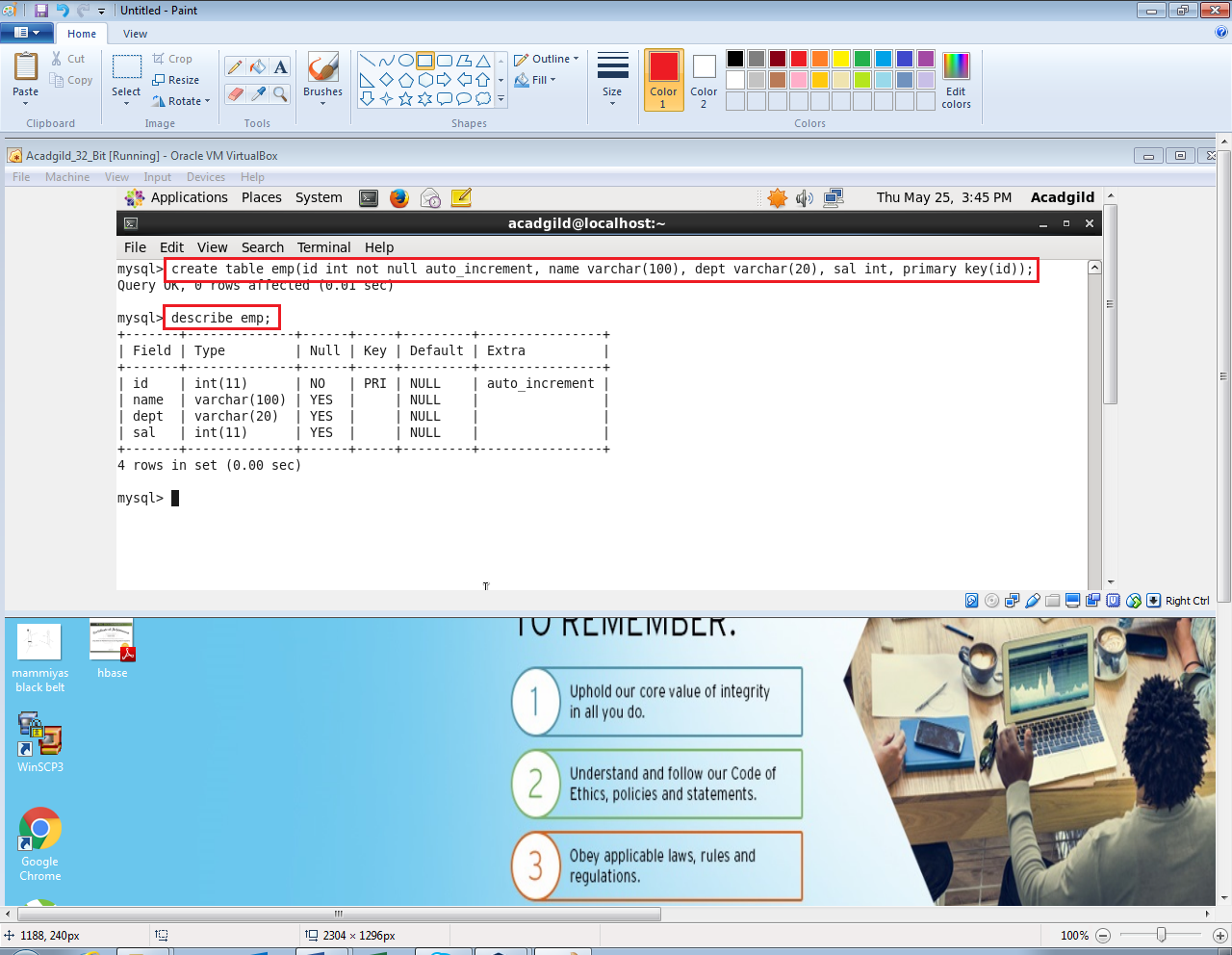


Creating a database by using the **‘create’** command and to work in the created database, we use **‘use’** command.

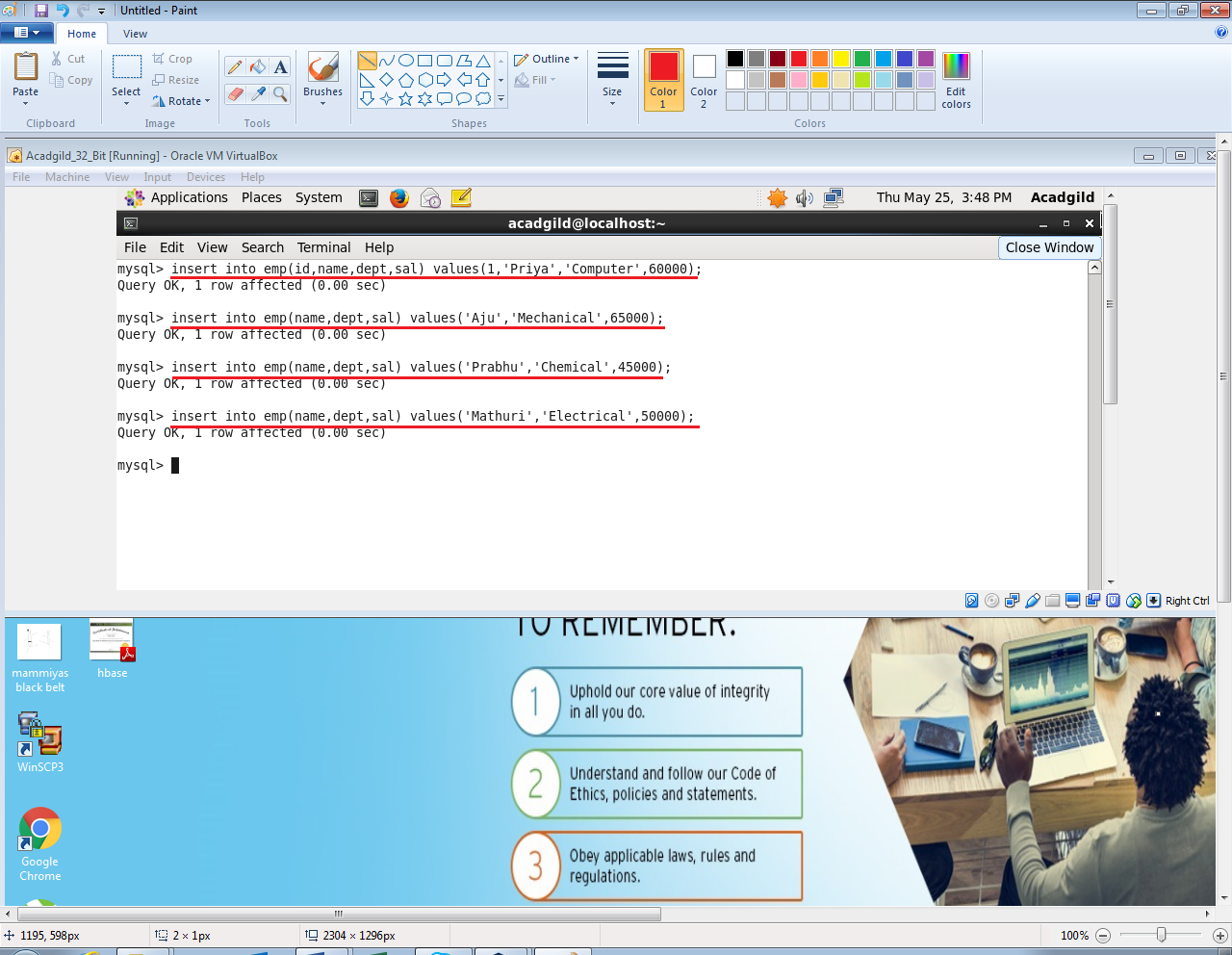


Creating a table **emp** by using the **‘create’** command:

After a table has been created with name **emp** and with the columns id, name, dept, sal. The scheme of this table can be checked using the **‘describe’** command:

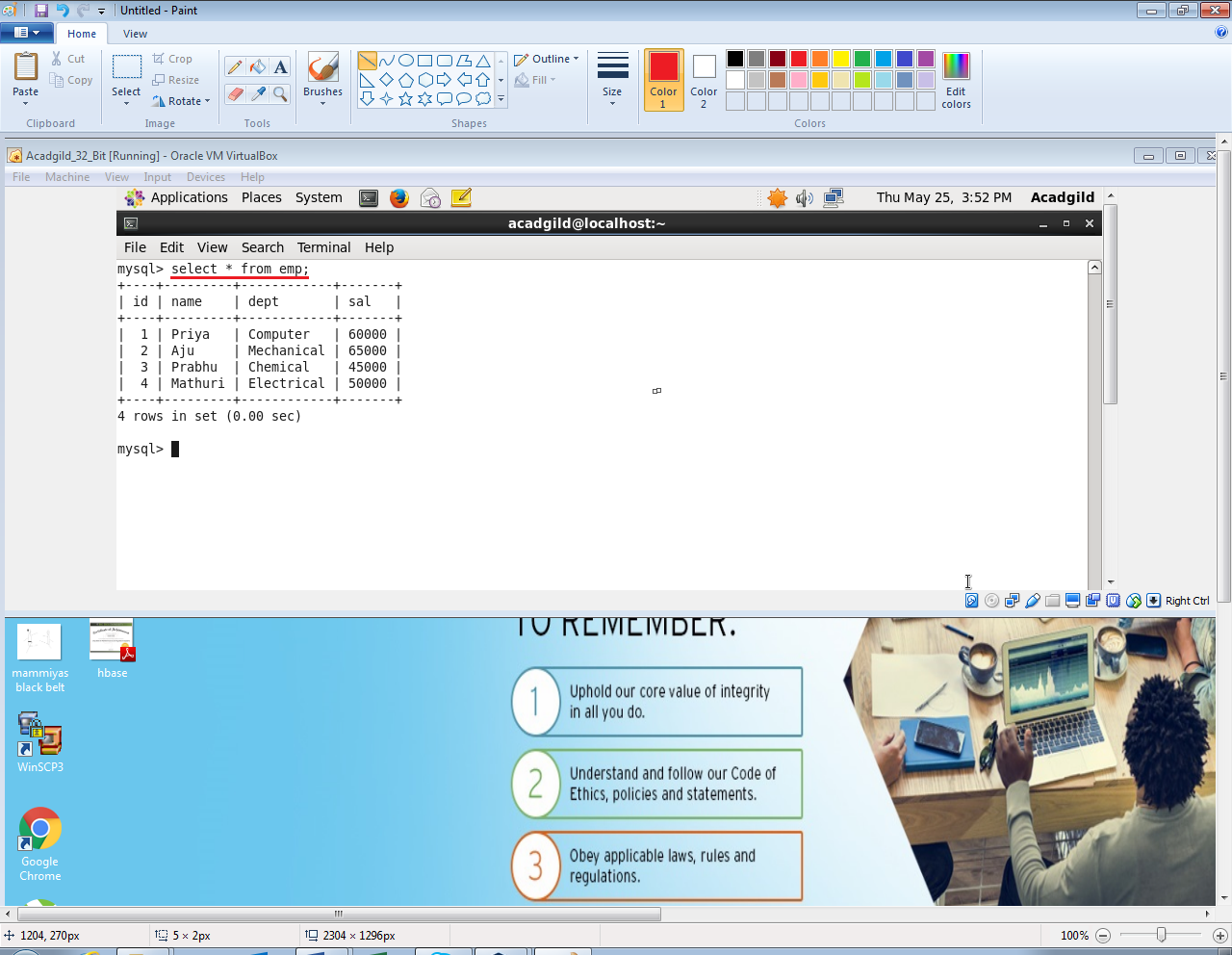


Inserting some sample data into the created table by using **‘insert’** command:



We can check the inserted using this command:

select \* from employee;



We have successfully created a table in MySQL, and we will now import the same into HDFS by using Sqoop.

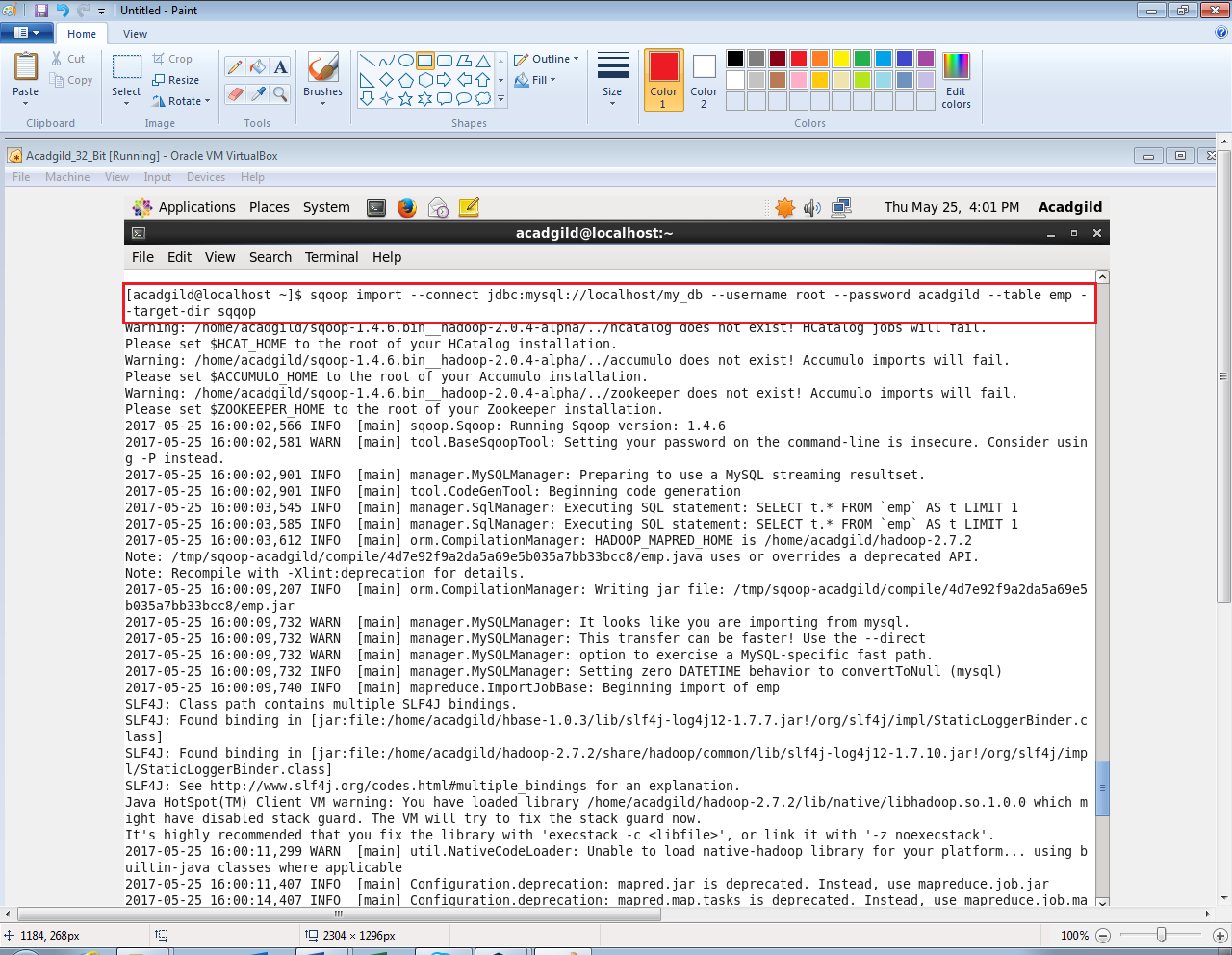
**Importing a Table from RDBMS to HDFS:**

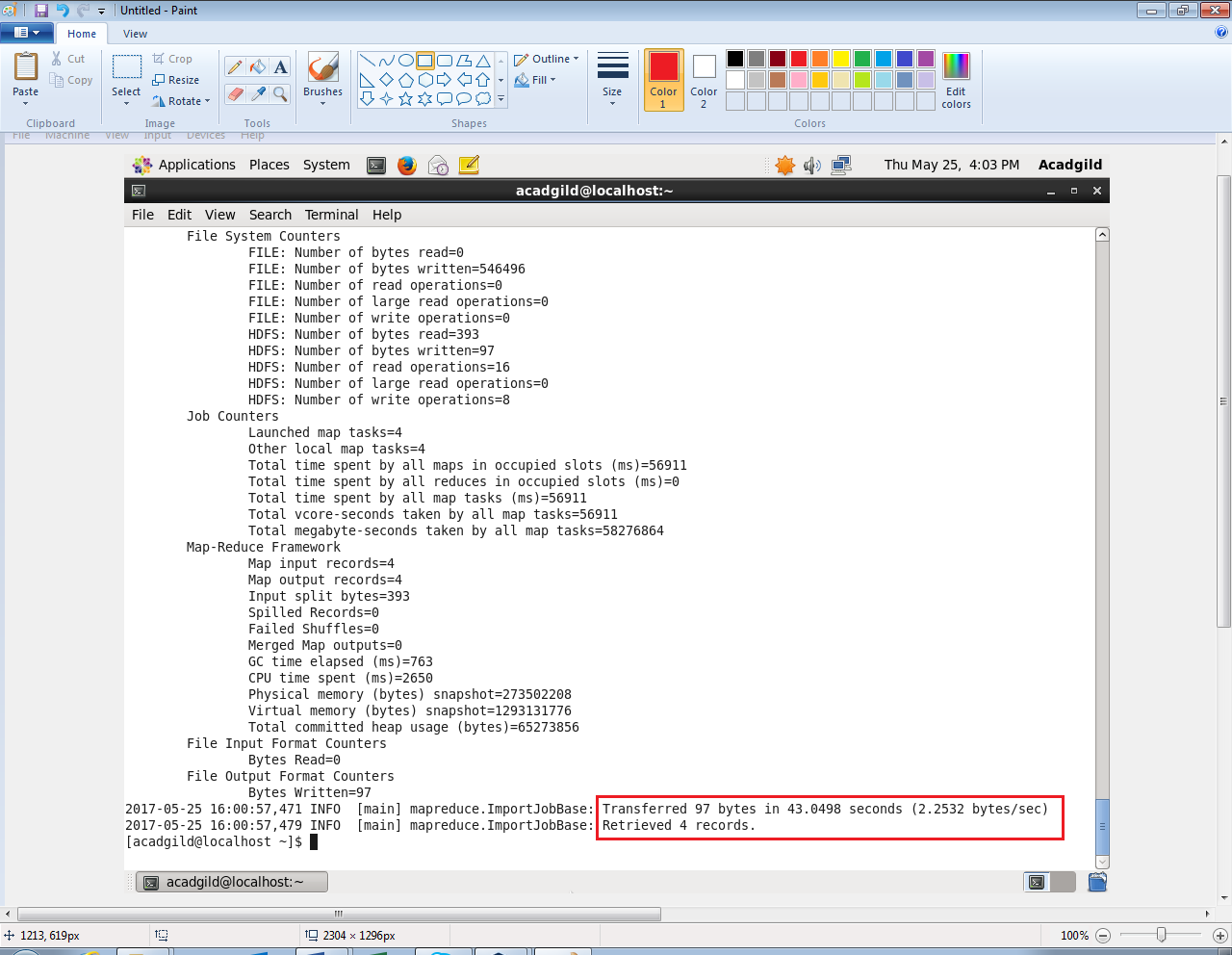
Here we are using a sample MySQL data and importing it into HDFS using Sqoop. Let us see how to create a table in MYSQL.

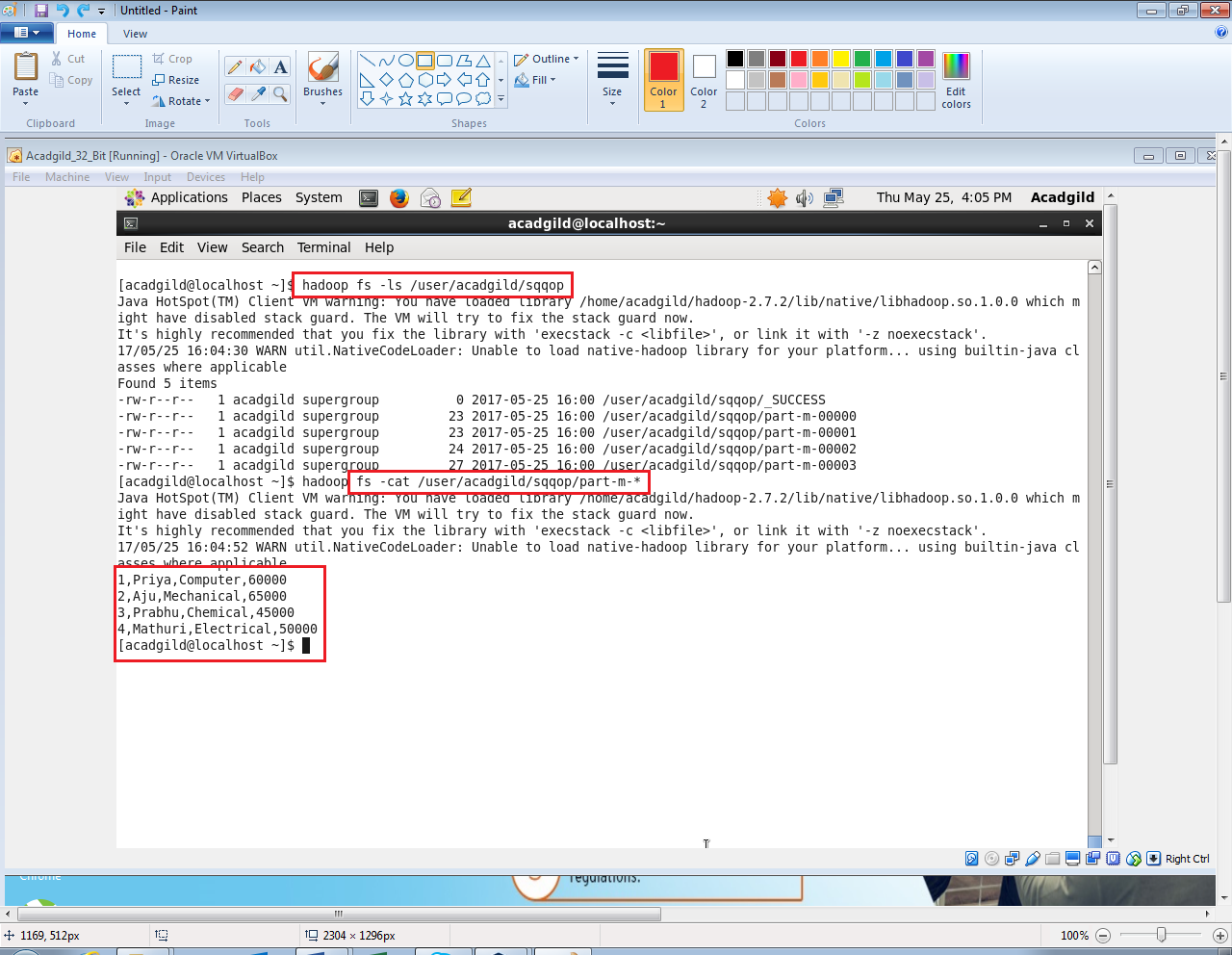
The following command can be used to import the table into HDFS.

Here we are connecting to MySQL through JDBC connectors and using the database Acadgild. Here it is necessary to specify the MySQL‘s username and password and the table name.

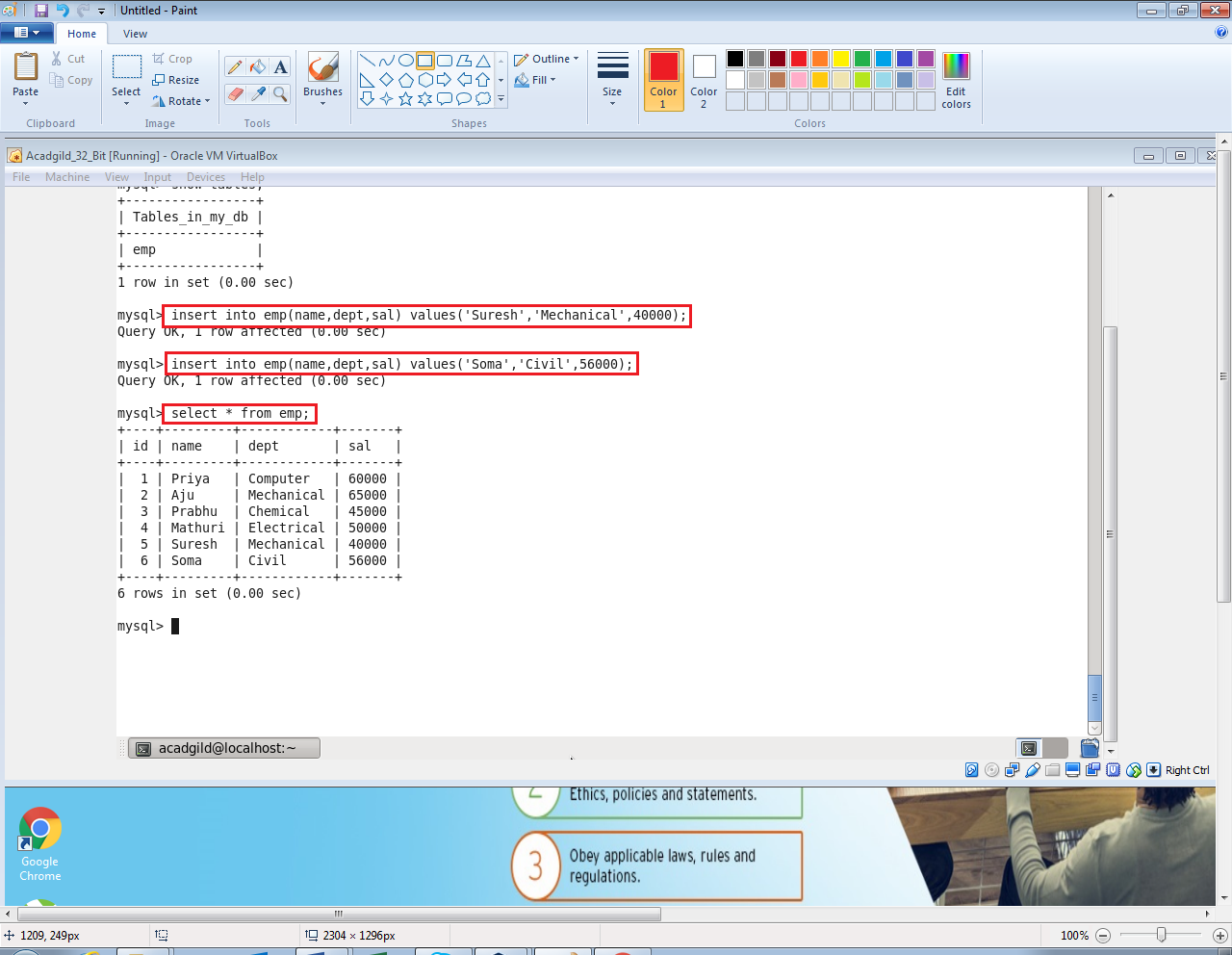
The command will be as shown below:







**Manually inserting few extra values in mysql table:**

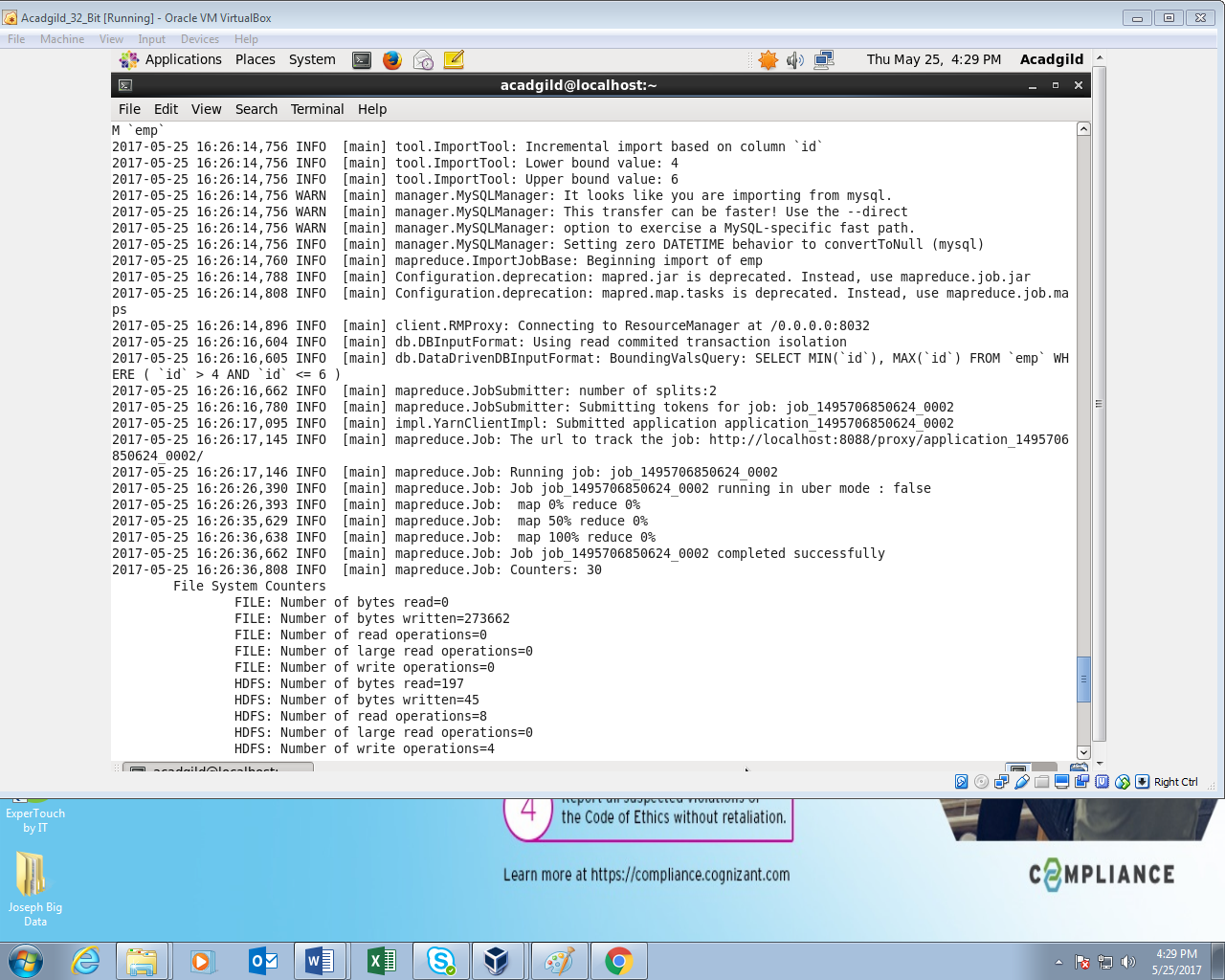
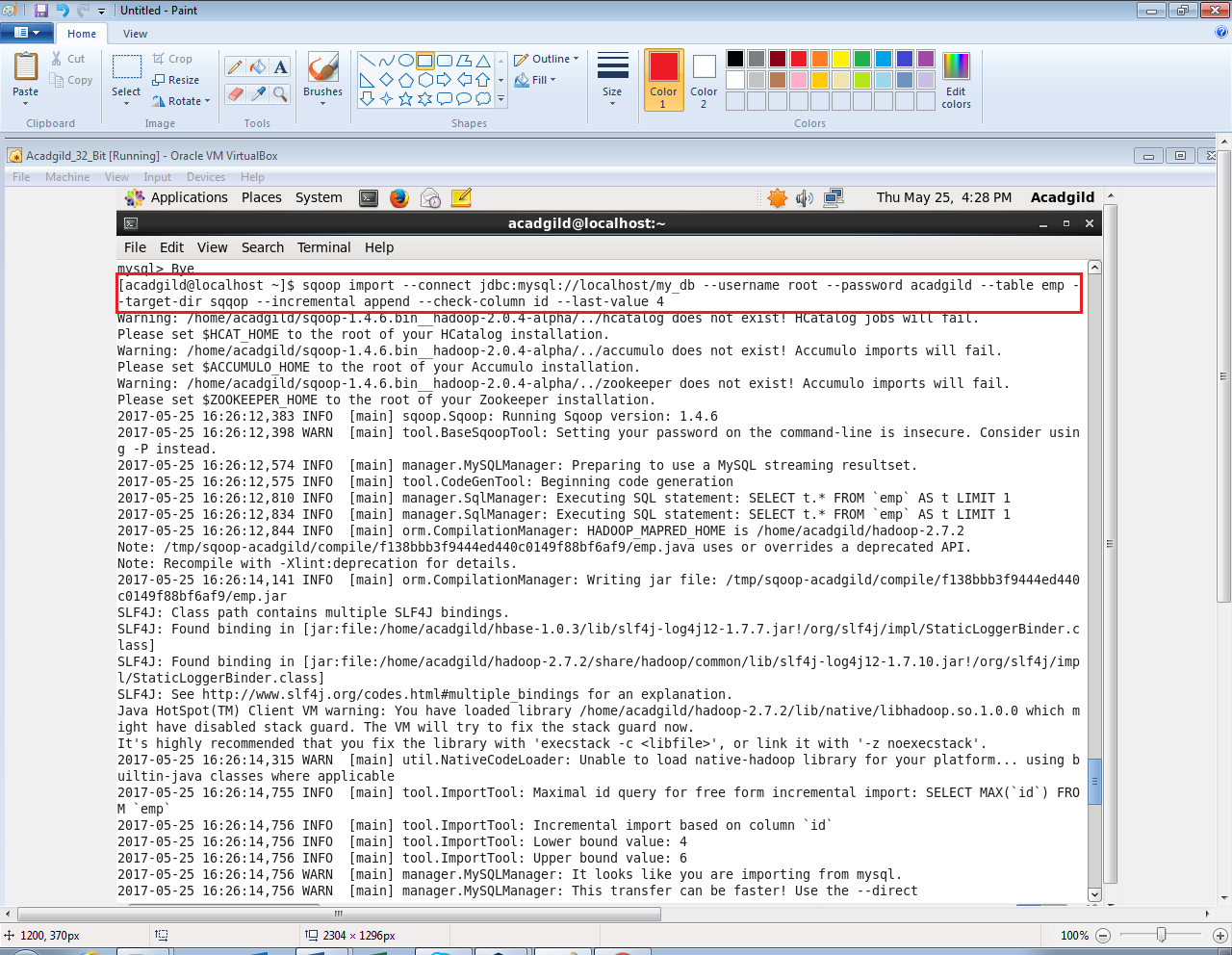
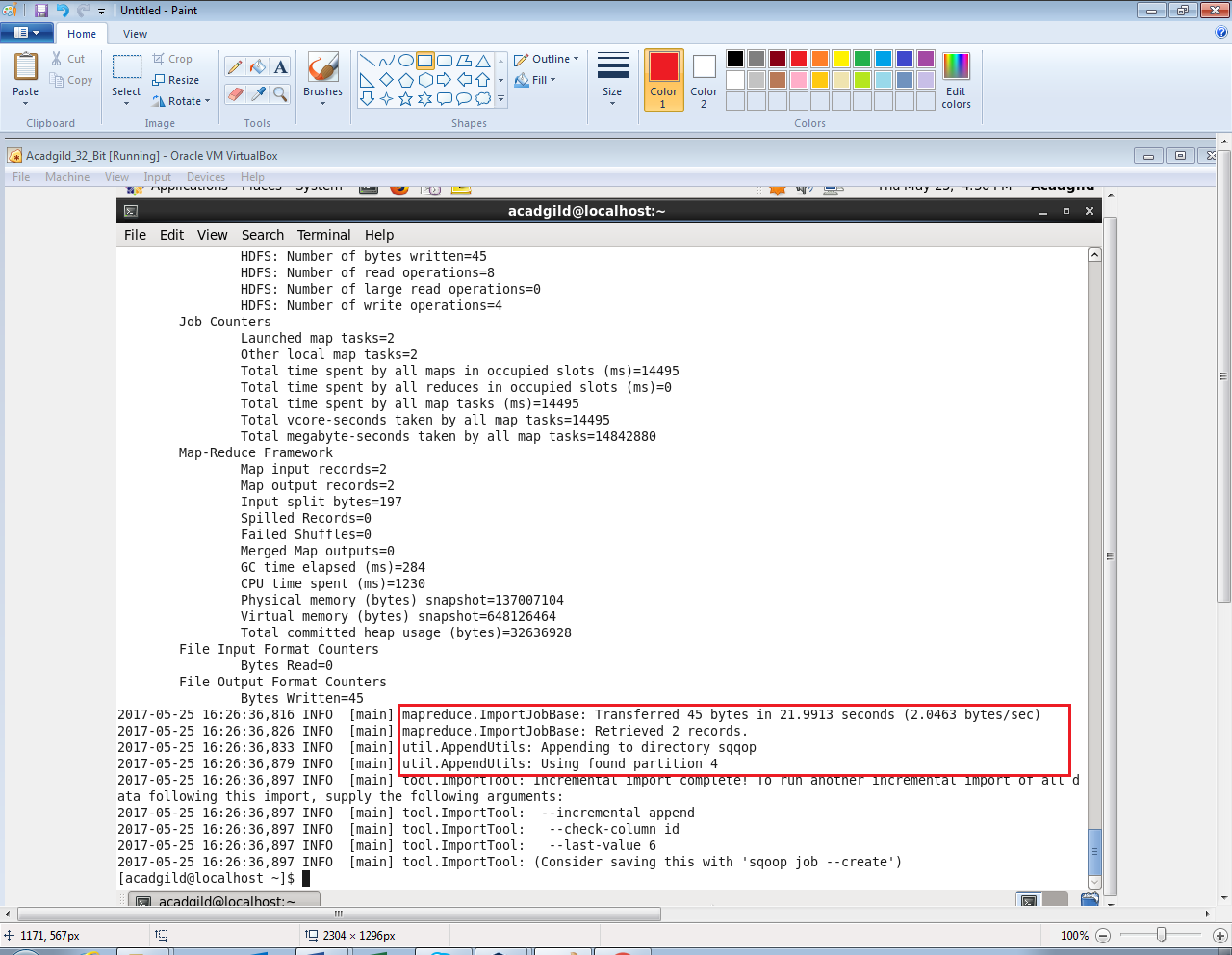
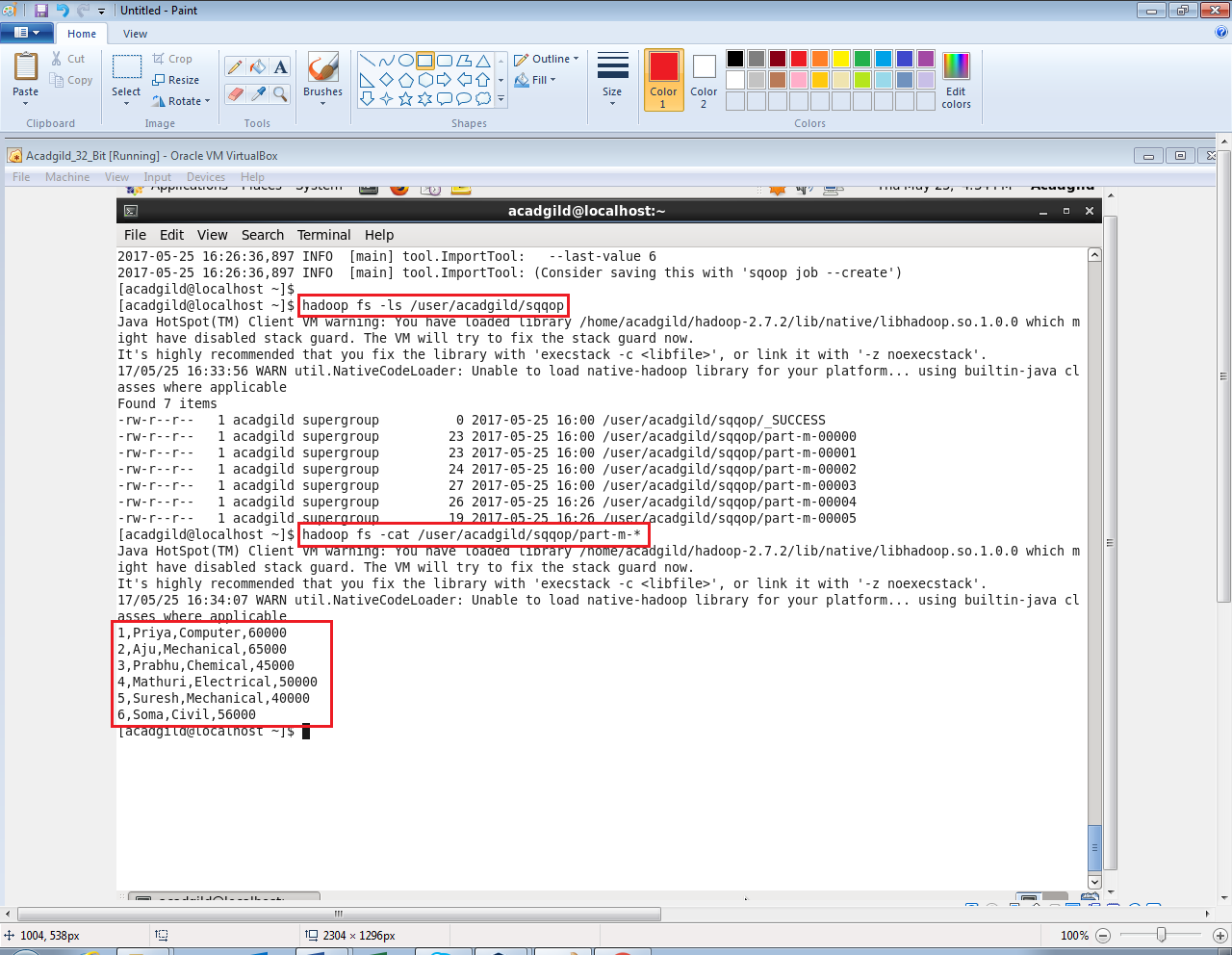


**The following syntax is used for the incremental option in Sqoop import command.**

**--incremental <mode>**

**--check-column <column name>**

**--last value <last check column value>**

As we can see in above image, 2 more records have been retrieved and the incremental import is now complete.

Along with message for next incremental import, you need to give last value as 6.

We can check and confirm the new data inside the new data inside HDFS. This is how incremental import is done every time for any number of new rows.