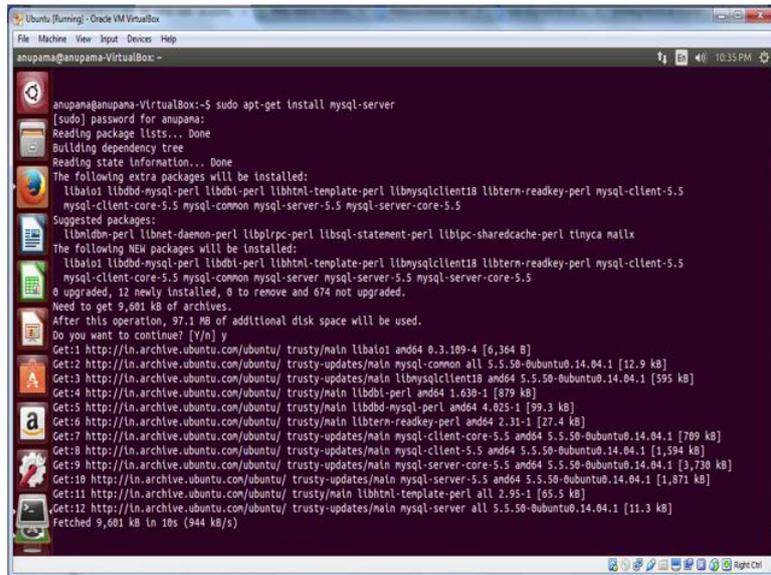**Privacy and Security in Online Social Networks**
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

**Lecture – 13**
Tutorial 3 Part 2 MySQL

(Refer Slide Time: 00:14)

In this tutorial, we will learn how to use data bases to store the data collected from various OSNs. We will start off with MySQL. MySQL is a database management system which uses a relational data base and SQL. As first step we will install MySQL. To do that, go to your terminal and type sudo apt get install MySQL hyphen server press enter and enter your sudo password.

(Refer Slide Time: 01:00)
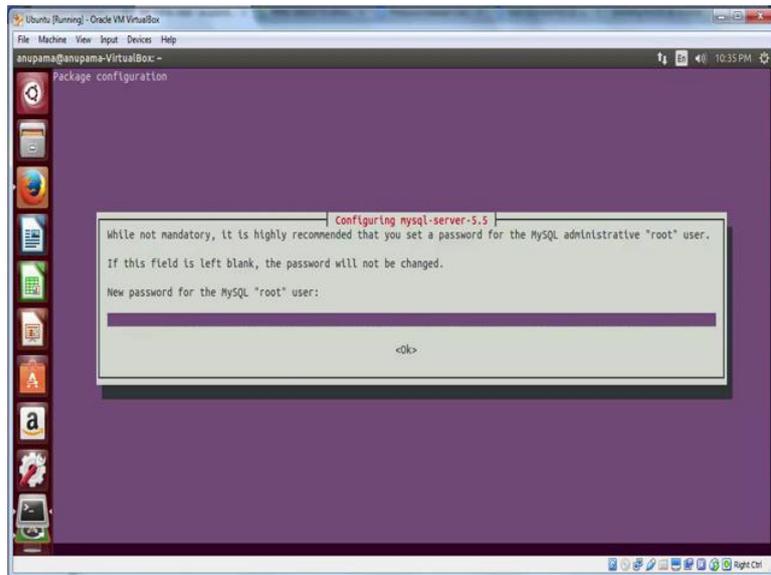
Now, you will be prompted for a password. Make sure that you would take a note of it because you will need it later.

(Refer Slide Time: 01:08)
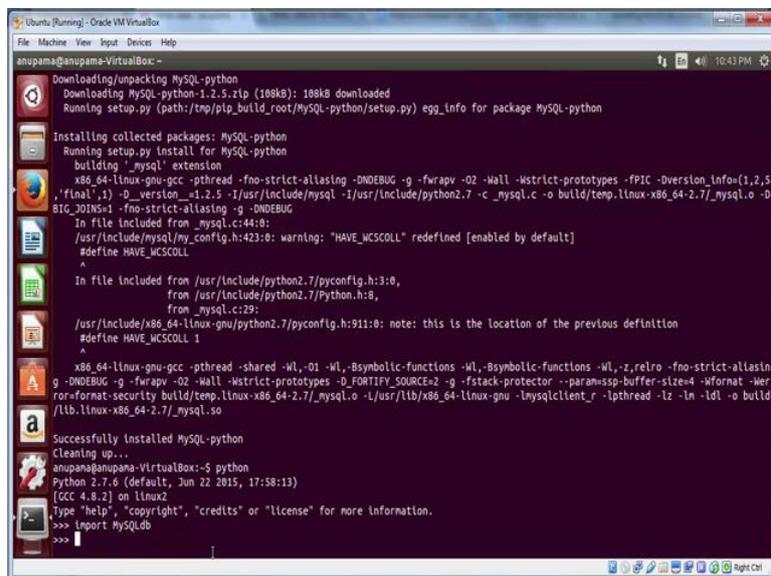
In the next step, we will see how to configure MySQL. To do that, type sudo mysql secure installation. Enter current password for root. Now, do not change the root password because you have already set it. Therefore, press n. You can remove anonymous accounts by pressing n and also press n for disallowing root login remotely. If you want you can remove the test data base or leave it as it is. We will choose to

delete it. Let us reload the privilege tables now and press yes. This finishes off the installation of MySQL.
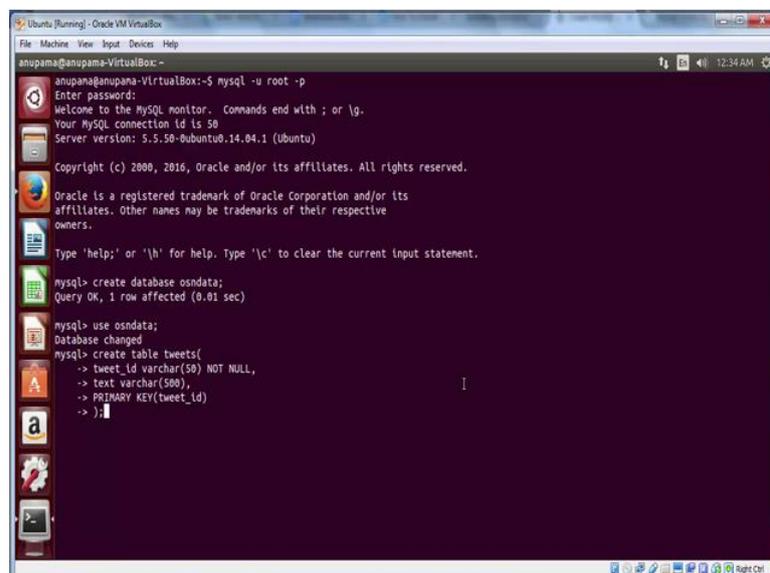
Now, to check whether MySQL has indeed been installed, let us check the MySQL version by typing the command MySQL minus minus version. It says MySQL version 5.5.5 has been installed. You can test whether MySQL has indeed been installed or not by trying to log into MySQL. To do that, type MySQL minus u root minus p and press enter. Enter the password which you created for your MySQL admin account and you are able to login.

Now, to connect python to MySQL, we will need MySQL dd module. To do that, type the following command, note that we are not typing, apt get install, but we are using the command apt get build dep which is used to build dependencies for a following package. Continue. Next we will install the python MySQL library by typing sudo pip install MySQL hyphen python and we are done with the MySQL installation along with integration with python. Now, to make sure that python module and MySQL are indeed installed, let us try going to the python CLI. Type python and inside the console, import MySQL db. If it has been successfully installed, there will be no error.

(Refer Slide Time: 04:14)

Now, we will see how to use MySQL. Log in to your MySQL account by typing MySQL minus u root minus p, enter your password and you will be inside the MySQL prompt. Let us create a database to store data using the command, create database osn

data. Now, let us create a table to store tweets collected by a Twitter API. We will first need to switch to the newly created database by typing use osn data and now we will create a table using the following command. Note that here, we are setting the primary key as the tweet id. Now, we have a data base and a table ready to store the data.

(Refer Slide Time: 05:25)
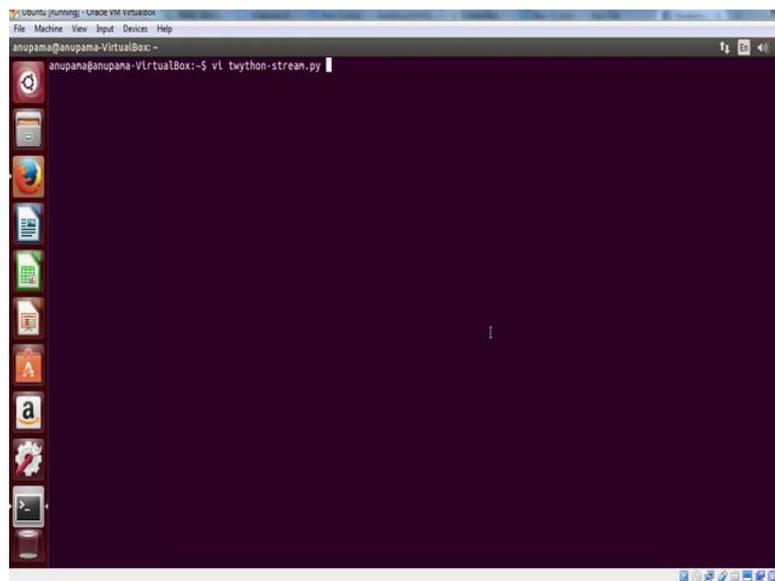
Let us now look at how we can store previously fetched data from twitter into a MySQL database. Let us go back to the script we wrote for collecting real time data from twitter.
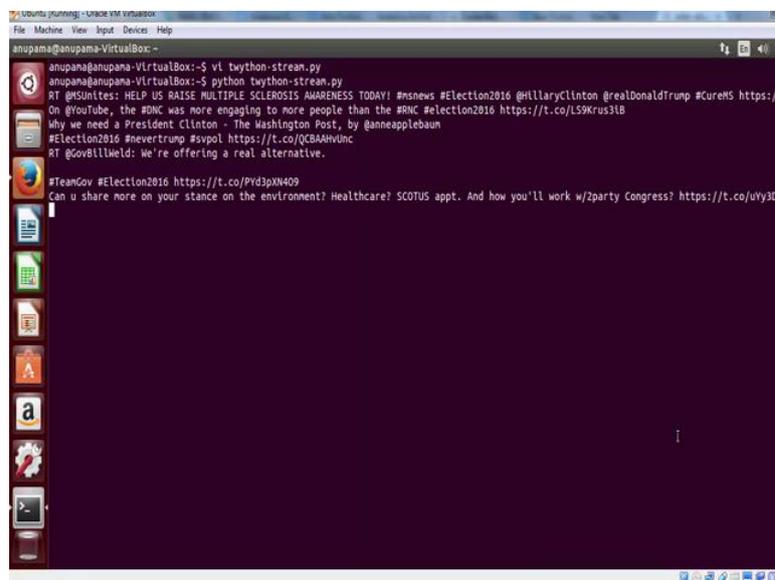
(Refer Slide Time: 05:42)
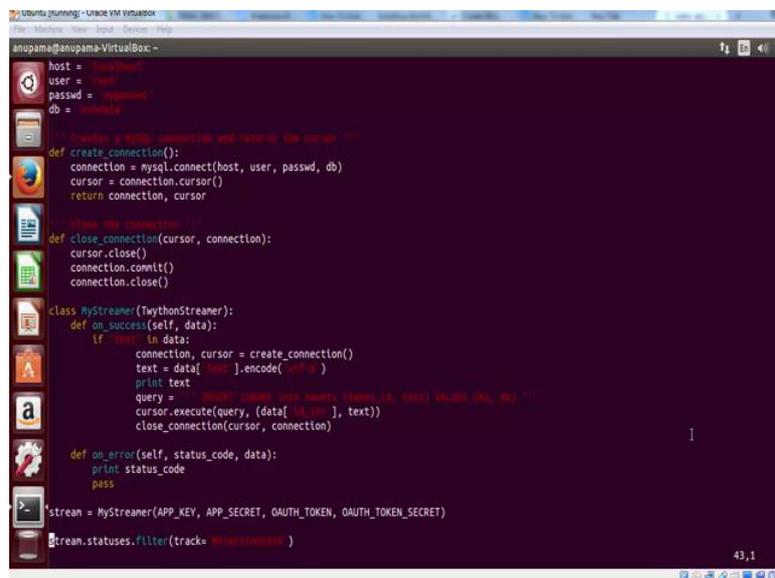
In case you had to monitor a set of keywords for a large amount of time, storing the output in a database can be helpful, so that it is indexed and can be easily queried later. Now, let us look at the changes which we have made to the existing script to store data into a MySQL table. We start off with importing MySQL db, then we have specified the credentials to access MySQL. Note that in this case, the host will be local host since we are accessing MySQL on our local machine.

Db is the name of the database where we will store the data. Next, we have two functions, create connection and close connection which are used to open and close connection to MySQL. In the previous tutorial, we were tracking a keyword hash election 2016 and simply printing the tweet text in the terminal, this time additionally we will store this into the data base. Now, remember we have already created a table named tweets which had two fields tweet id and text. We only need to execute the insert query to set the value of tweet id and tweet text field. Note that the insert ignore command ignores the insertion of a new entry with previously existing primary key.
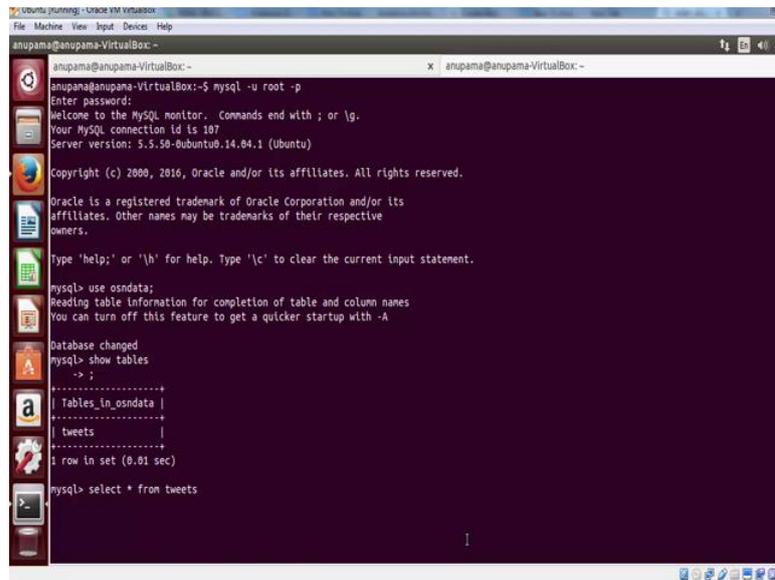
Now, let us take this file and try running it.

(Refer Slide Time: 07:28)

(Refer Slide Time: 07:52)

You will notice that in a while the tweets <mark>will</mark> start coming. Now, open a new tab in the terminal and let us go to MySQL. Once you are in the MySQL prompt, switch to the data base OSN data. Now, if you run the command, show tables, you will be able to see the list of tables in this particular database. We have the tweets table which has the tweet id and the tweet text from the streaming API running in parallel. Let us try to access the data in the tweet table, type the command, select star from tweets.

(Refer Slide Time: 08:50)

You will notice that there is a bunch of text and at the bottom <mark>it says</mark> 12 rows in set, let

us verify the number of tweets present in the table so far. To do that, go to the previous command, but this time add count in front of the star and enclose it in the brackets. This command will count the number of rows in the table. Now, we have 25 tweets in this particular table. The select query can also be integrated into python script for a more effective data collection.