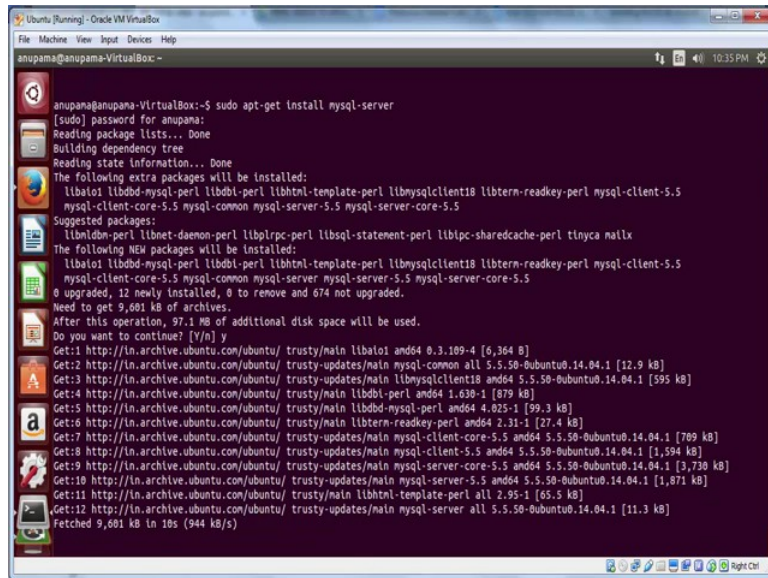


**Privacy and Security in Online Social Networks**  
**Department of Computer Science and Engineering**  
**Indian Institute of Technology, Madras**

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

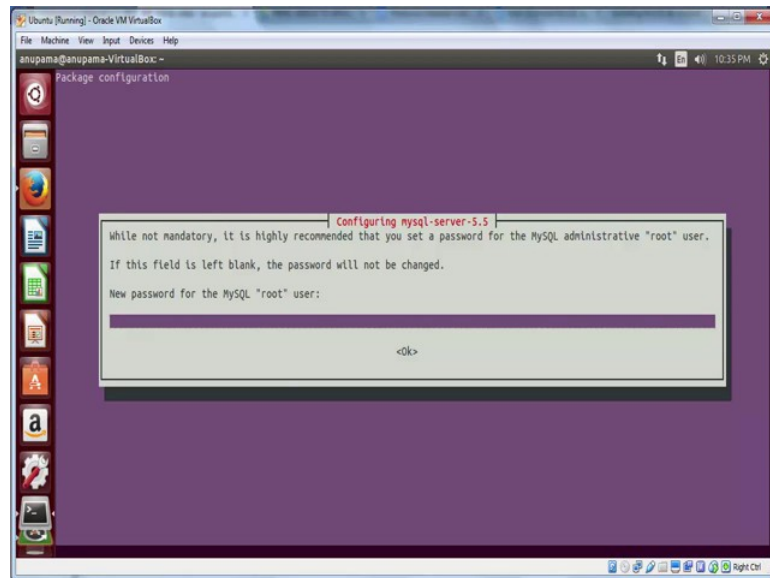
(Refer Slide Time: 00:14)



```
anupama@anupama-VirtualBox:~$ sudo apt-get install mysql-server
[sudo] password for anupama:
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following extra packages will be installed:
  libatoi libbdb-mysql-perl libdbi-perl libhtml-template-perl libmysqlclient18 libterm-readkey-perl mysql-client-5.5
  mysql-client-core-5.5 mysql-common mysql-server-5.5 mysql-server-core-5.5
Suggested packages:
  libnids-perl libnet-daemon-perl libssl-perl libssl-dev libssl-doc libterm-readkey-perl libterm-readkey-perl libterm-readkey-perl
  libterm-readkey-perl libterm-readkey-perl libterm-readkey-perl libterm-readkey-perl libterm-readkey-perl libterm-readkey-perl
The following NEW packages will be installed:
  libatoi libbdb-mysql-perl libdbi-perl libhtml-template-perl libmysqlclient18 libterm-readkey-perl mysql-client-5.5
  mysql-client-core-5.5 mysql-common mysql-server-5.5 mysql-server-core-5.5
0 upgraded, 12 newly installed, 0 to remove and 674 not upgraded.
Need to get 9,601 kB of archives.
After this operation, 97.1 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://ln.archive.ubuntu.com/ubuntu/ trusty/main libatoi amd64 0.3.109-4 [6,364 B]
Get:2 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-common all 5.5.50-0ubuntu0.14.04.1 [12.9 kB]
Get:3 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main libmysqlclient18 amd64 5.5.50-0ubuntu0.14.04.1 [595 kB]
Get:4 http://ln.archive.ubuntu.com/ubuntu/ trusty/main libdbi-perl amd64 1.630-1 [879 kB]
Get:5 http://ln.archive.ubuntu.com/ubuntu/ trusty/main libbdb-mysql-perl amd64 4.025-1 [99.3 kB]
Get:6 http://ln.archive.ubuntu.com/ubuntu/ trusty/main libterm-readkey-perl amd64 2.31-1 [27.4 kB]
Get:7 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-client-core-5.5 amd64 5.5.50-0ubuntu0.14.04.1 [709 kB]
Get:8 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-client-5.5 amd64 5.5.50-0ubuntu0.14.04.1 [1,594 kB]
Get:9 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-server-core-5.5 amd64 5.5.50-0ubuntu0.14.04.1 [3,730 kB]
Get:10 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-server-5.5 amd64 5.5.50-0ubuntu0.14.04.1 [1,871 kB]
Get:11 http://ln.archive.ubuntu.com/ubuntu/ trusty/main libhtml-template-perl all 2.95-1 [65.5 kB]
Get:12 http://ln.archive.ubuntu.com/ubuntu/ trusty-updates/main mysql-server all 5.5.50-0ubuntu0.14.04.1 [11.3 kB]
Fetched 9,601 kB in 10s (944 kB/s)
```

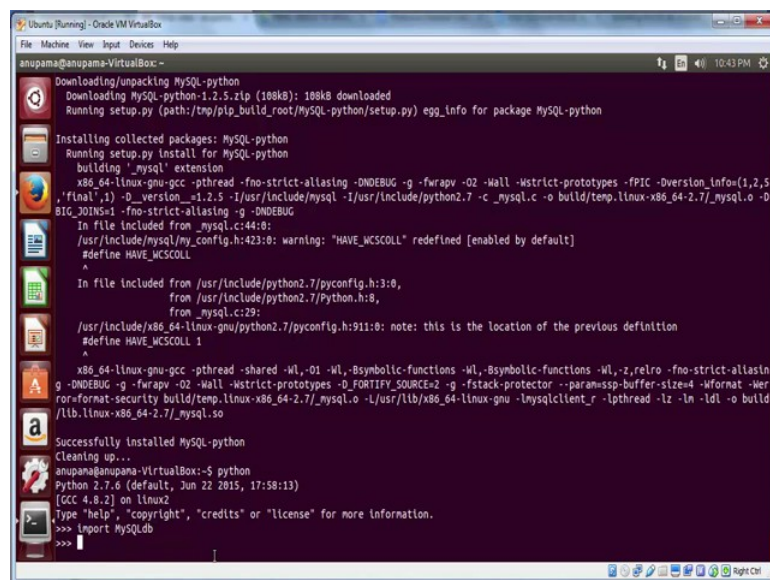
In this tutorial, we will learn how to use data bases to store the data collected from various **OSNs**. We will start off with MySQL. MySQL is a database management system which uses a relational data base and SQL. As first step we will install MySQL. To do that, go to your terminal and type `sudo apt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



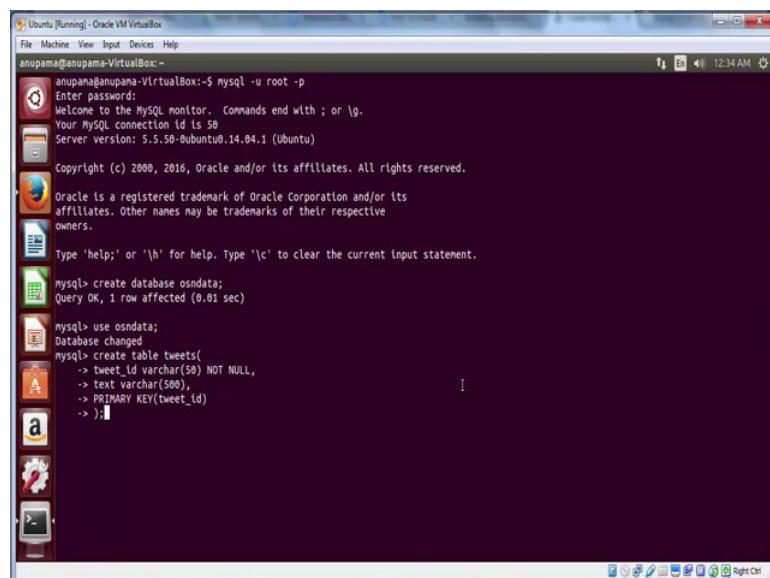
In the next step, we will see how to configure MySQL. To do that, type `sudo mysql` secure installation. Enter current password for root. Now, do not change the root password because you have already set it. Therefore, press `n`. You can remove anonymous accounts by pressing `n` and also press `n` for 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 --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 -u root -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 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-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 `MySQLdb`. If it has been successfully installed, there will be no error.

(Refer Slide Time: 04:14)



```
anupama@anupama-VirtualBox:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 50
Server version: 5.5.50-0ubuntu0.14.04.1 (Ubuntu)

Copyright (c) 2000, 2016, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

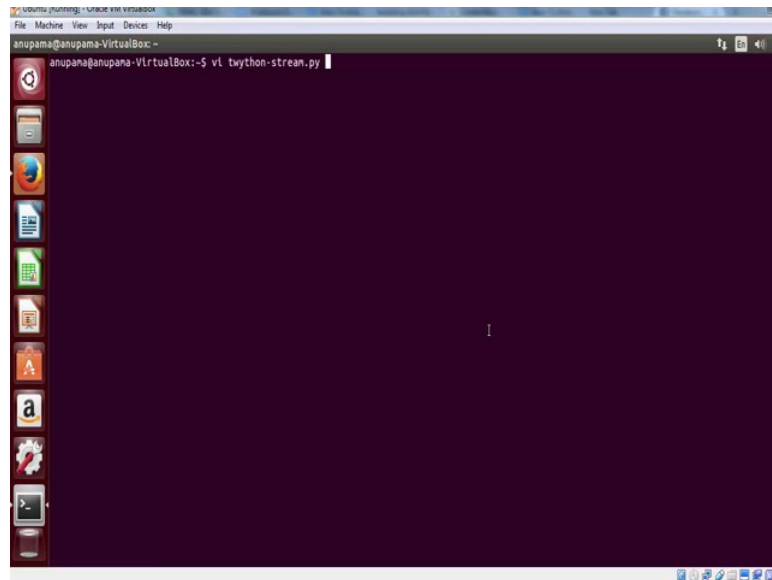
mysql> create database osndata;
Query OK, 1 row affected (0.01 sec)

mysql> use osndata;
Database changed
mysql> create table tweets(
-> tweet_id varchar(50) NOT NULL,
-> text varchar(500),
-> PRIMARY KEY(tweet_id)
-> );
```

Now, we will see how to use MySQL. Log in to your MySQL account by typing `MySQL -u root -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`

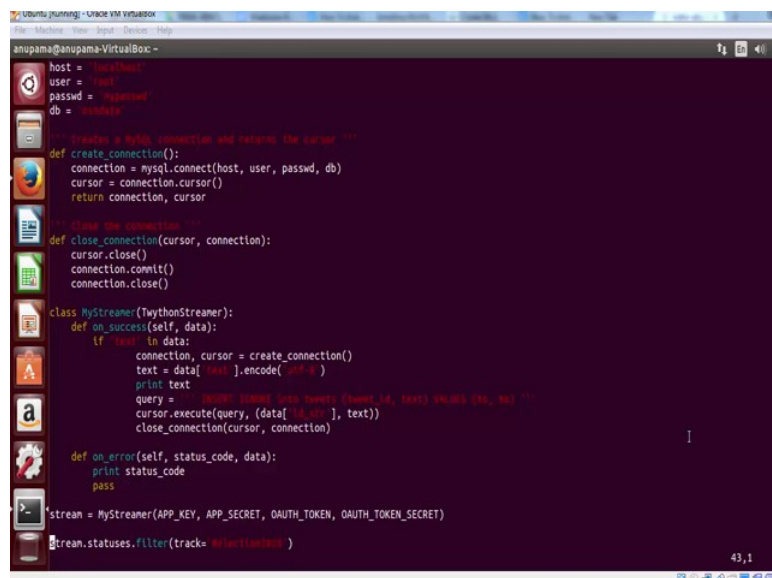
data. Now, let us create a table to store tweets collected by a **Twitter** API. We will first need to switch to the newly created database by typing use osn data and now we will create a table using the following command. Note that here, we are setting the primary key as the tweet id. Now, we have a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

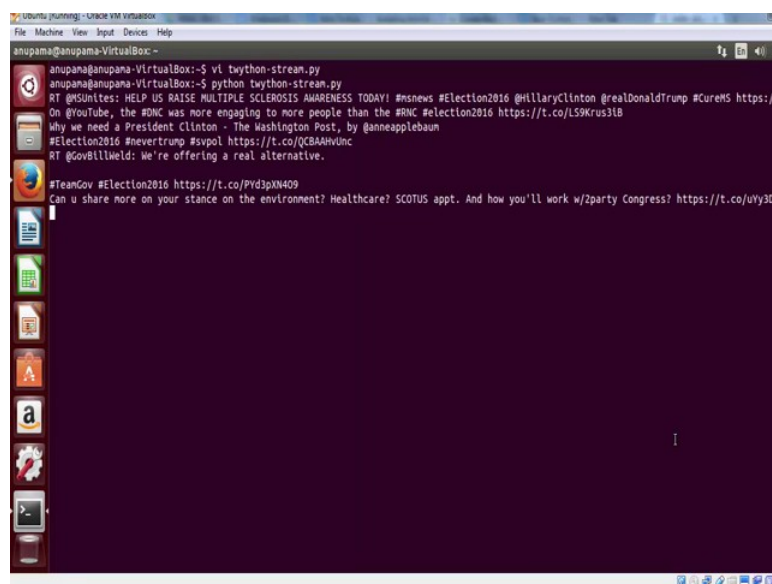


In case you had to monitor a set of keywords for a large amount of time, storing the **output** in a database can be helpful, so that it is indexed and can be easily **queried** later. Now, let us look at the changes which we have made to the existing script to store data into a MySQL table. We start 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.

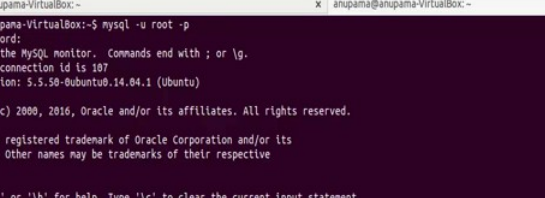
Now, let us take this file and try running it.

(Refer Slide Time: 07:28)



```
anupama@anupama-VirtualBox:~$ vi twython-stream.py
anupama@anupama-VirtualBox:~$ python twython-stream.py
RT @MSUnites: HELP US RAISE MULTIPLE SCLEROSIS AWARENESS TODAY! #MSnews #Election2016 @HillaryClinton @realDonaldTrump #CureMS https://
On YouTube, the #DNC was more engaging to more people than the #RNC #Election2016 https://t.co/L59Krus3tB
Why we need a President Clinton - The Washington Post, by @anneapplebaum
#Election2016 #NeverTrump #supol https://t.co/QCBAAHvMnc
RT @GovBillWeld: We're offering a real alternative.
#TeamGov #Election2016 https://t.co/PYd3pXN409
Can u share more on your stance on the environment? Healthcare? SCOTUS appt. And how you'll work w/2party Congress? https://t.co/uYy30o
```

(Refer Slide Time: 07:52)



The screenshot shows a VirtualBox window titled "anupama@anupama-VirtualBox -". The terminal window inside shows the following output:

```
anupama@anupama-VirtualBox:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 187
Server version: 5.5.50-0ubuntu0.14.04.1 (Ubuntu)

Copyright (c) 2000, 2016, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> use osndata;
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_osndata |
+-----+
| tweets            |
+-----+
1 row in set (0.01 sec)

mysql> select * from tweets
```

You will notice that in a while the tweets **will** 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)

```

kali@kali:~$ python3 scraper.py
File Machine View Input Devices Help

anupama@anupama-VirtualBox: ~
x anupama@anupama-VirtualBox: ~

anupama@anupama-VirtualBox: ~
| 75986712877143296 | RT jorgnet: â€¦people who donâ€¦ trust â€¦the mainstreamâ€¦ media are more likely to swallow disinformation.
https://t.co/30DrKghBP
#Election2016
| 759867128187892992 | #JustNotTrump 0V''0V# #hillaryclinton #gay #instagram #cat #funny #election2016 https://t.co/2KE3dF80MT
| 759867147673743368 | RT How the #MuslimBrotherhood controls #HillaryClinton & the #Democrats
Give widest distribution!!
#Election2016
https://t.co/RmfnLcStfif
| 759867167932157952 | Best campaign promise yet! https://t.co/pfIIeuuWvD
| 75986717462286878 | RT @GovGaryJohnson: People are hungry to vote FOR someone as opposed to voting against someone else.
#teagovov #election2016 https://t.co/sâ€¦
| 759867185032568648 | RT @PSYCHICTWINS: Mark Cuban endorses Hillary Clinton, calls Donald Trump 'batshit crazy' @CNNPolitics https://
pezfzB #Election2016
-----+
12 rows in set (0.01 sec)

mysql> select count(*) from tweets;
+-----+
| count(*) |
+-----+
| 12 |
+-----+
1 row in set (0.00 sec)

mysql>

```

You will notice that there is a bunch of text and at the bottom it says 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.