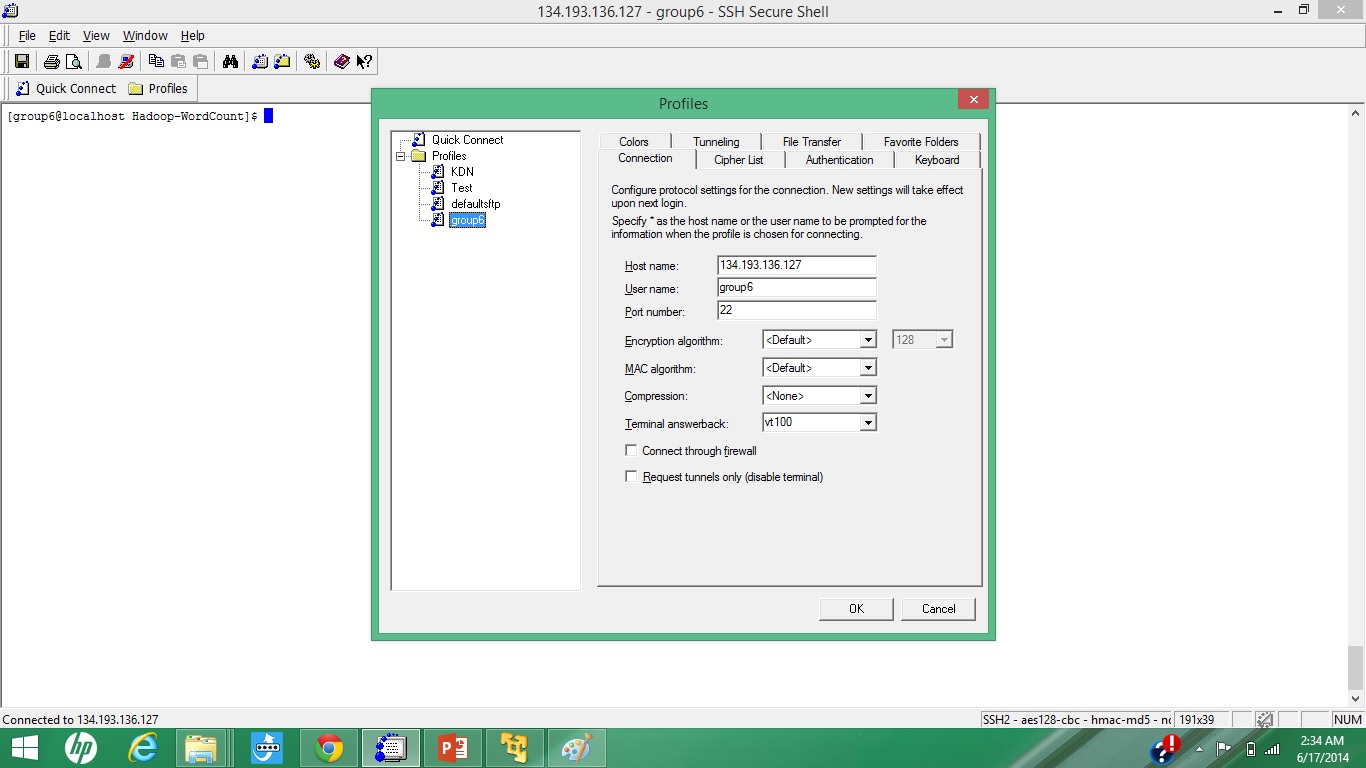
|  |
| --- |
| CS590BD Big Data Analytics and Applications |
| Lab 1 Assignment |
| Goutham Marikanti  16171450 |

TASK 2

# Subtask1:

# How to access UMKC Cloudera Servers?

* Download SSH/Putty from the internet, install on your machine.
* Connect to UMKC VPN using SSO credentials and enter the details of Host, Username, Port as captured below and save it as a Profile to get connected each time using this profile.

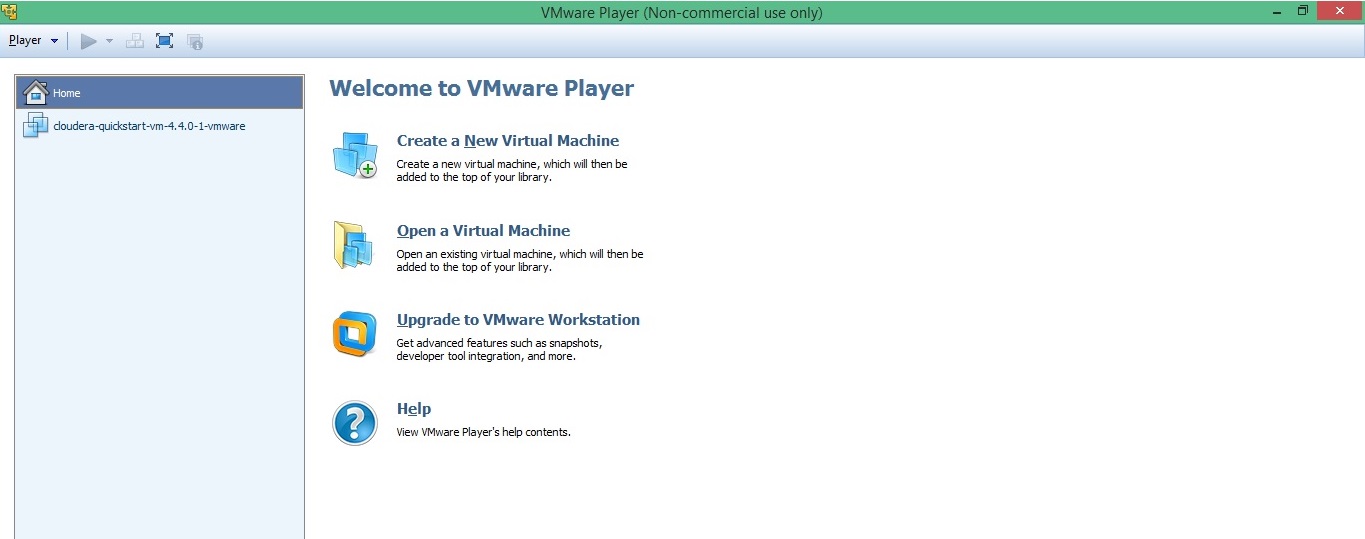


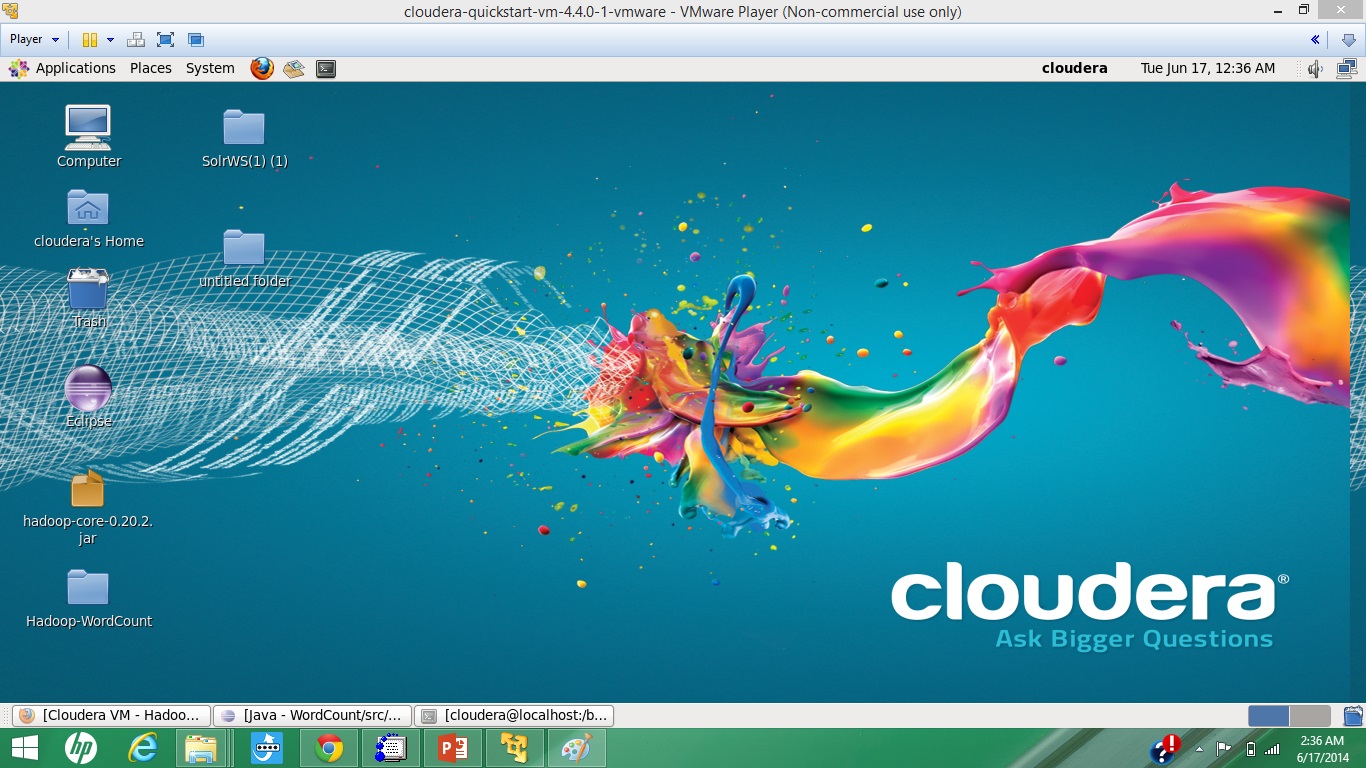
* Once the system accepts the credentials, you are connected to UMKC Cloudera Server.

# Subtask 2:

# How to install your own Cloudera Server?

* Download VMWare Player, Cloudera Image from the internet.
* Once VMWare player isinstalled, load Cloudera Image onto it.



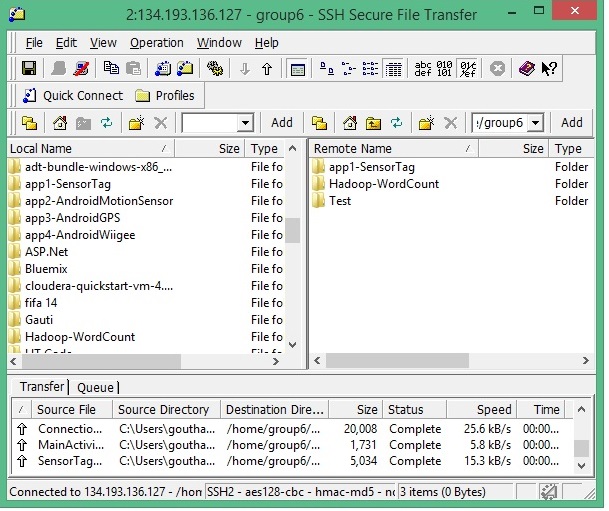


* Cloudera Server can now be run on your local machine.

# Subtask 3:

# How to transfer files to Cloudera?

* Download and install WinScp from the internet for Putty or open SSH Secure connection.
* Connect to host and then just drag and your files from your local system to Cloudera Server.



* Files are thus transferred onto Cloudera from your local machine.

# Subtask 4:

# How to run a program “Word Count” on Cloudera?

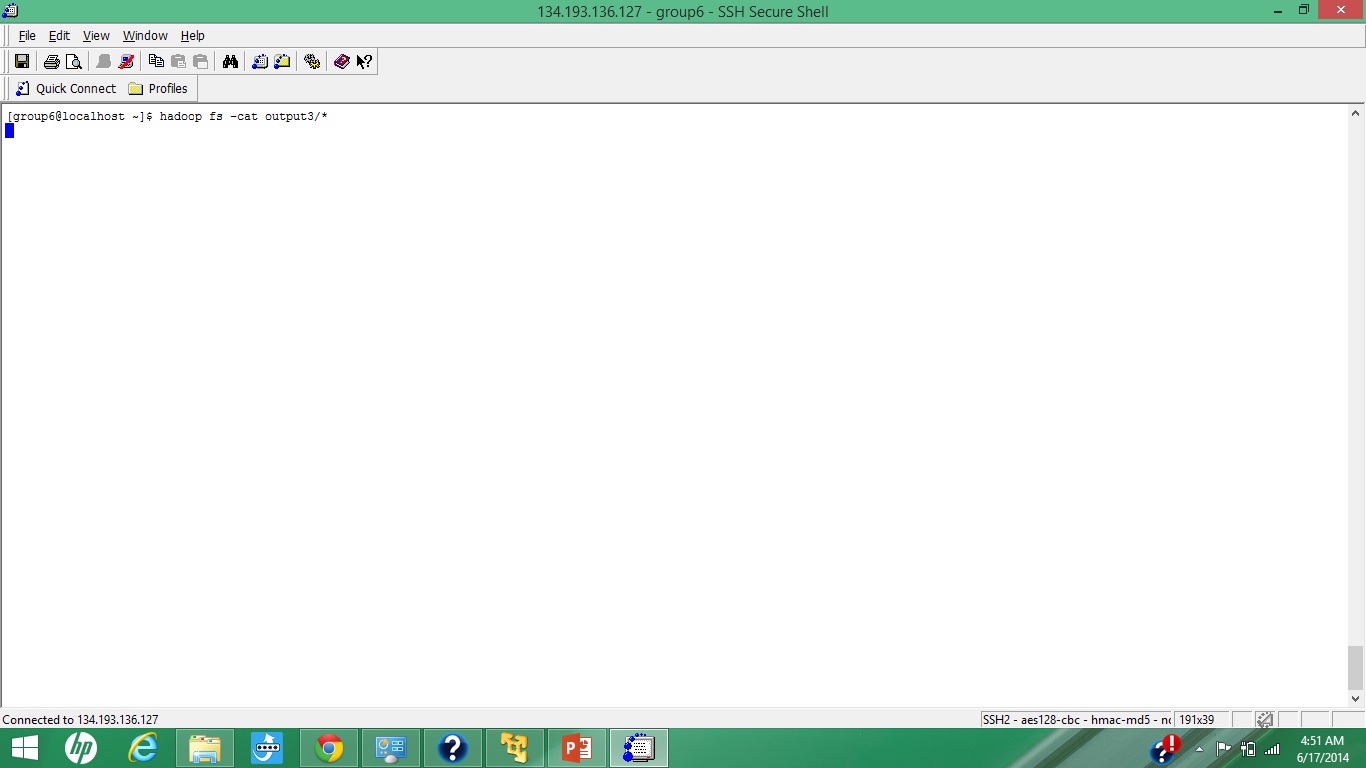
* Download the Source code for “Word Count” from the given link.
* Copy the Source Code from local machine to Cloudera Server and run the following commands:

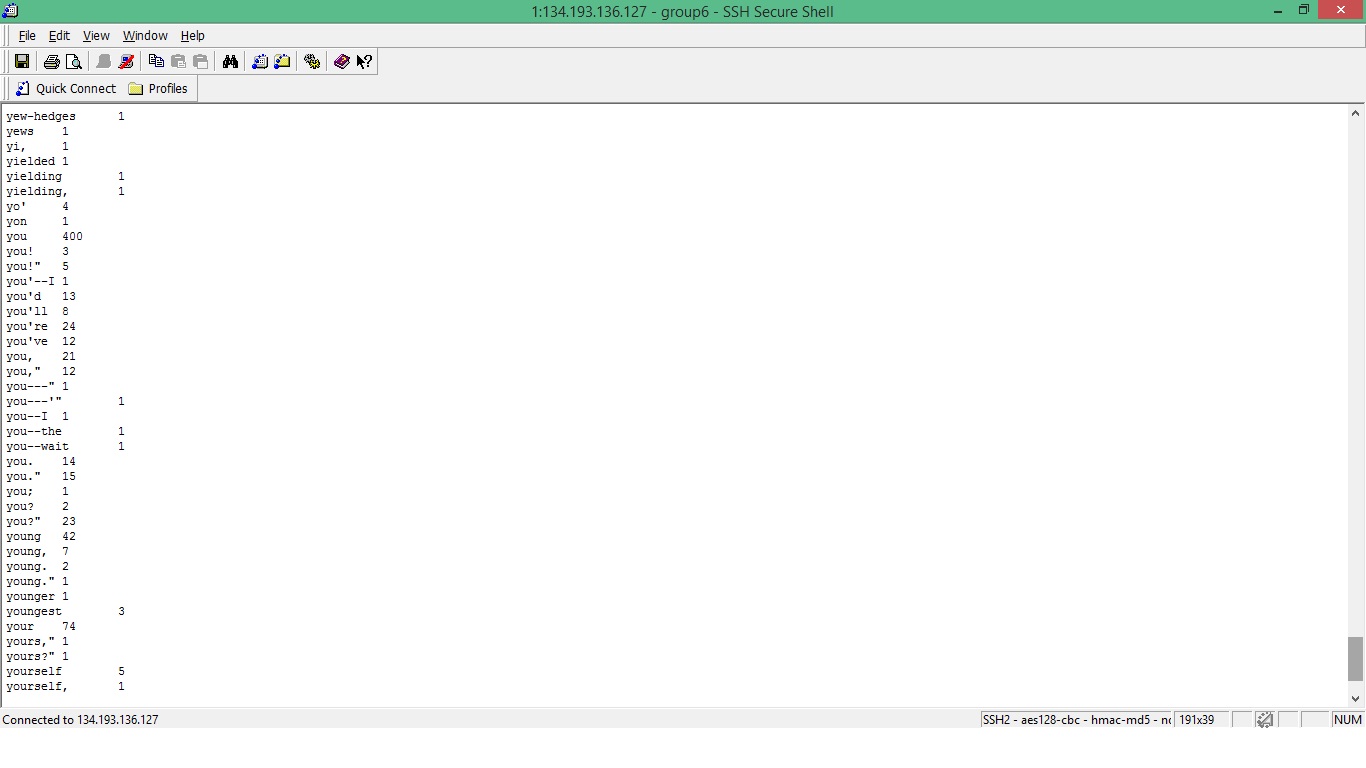
*hadoop fs –put input input*

*hadoop jar wordcount.jar WordCount input output3*

**

*hadoop fs –cat output3/\**





* The Word Count program is run and the output is displayed by the last command used above, the occurrence of each word with the corresponding no of times is displayed.

# Subtask 5:

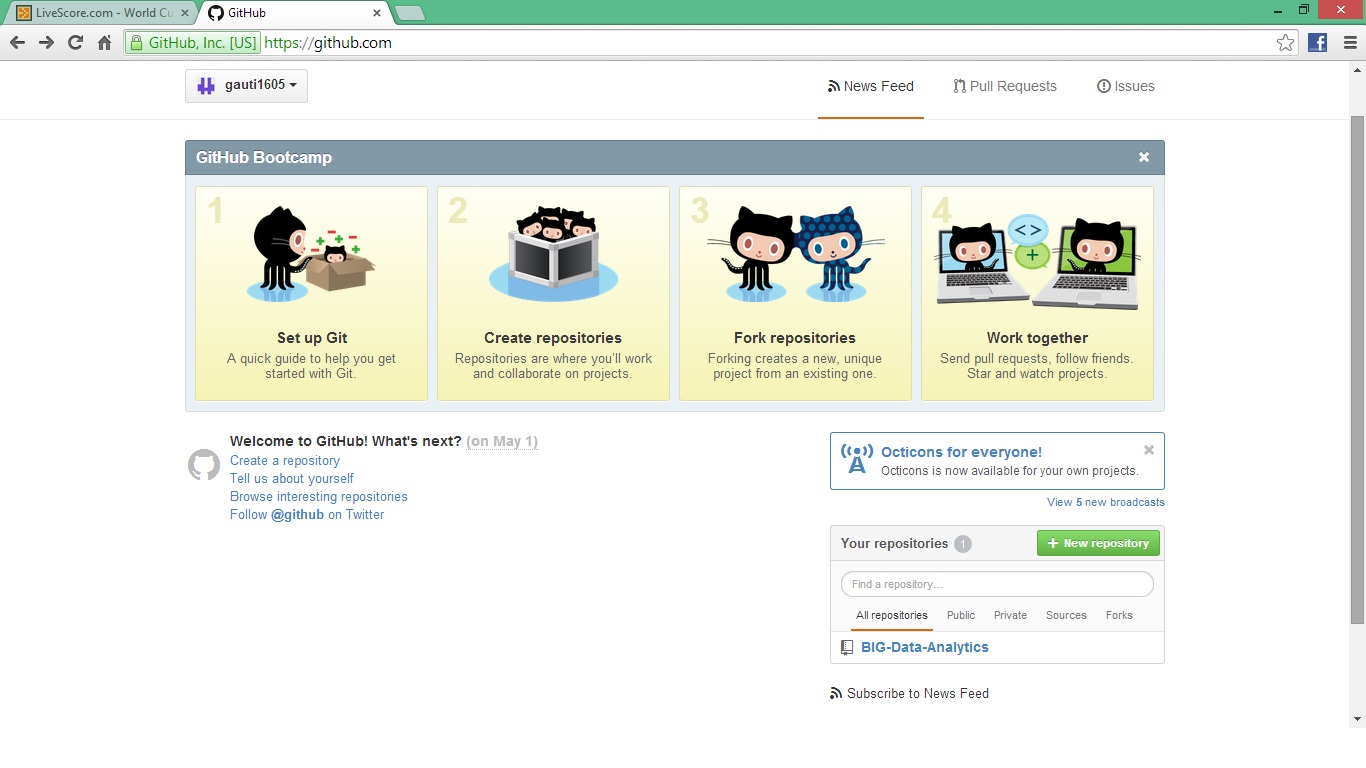
# How to run a program “Word Count” on Cloudera?

TASK 2

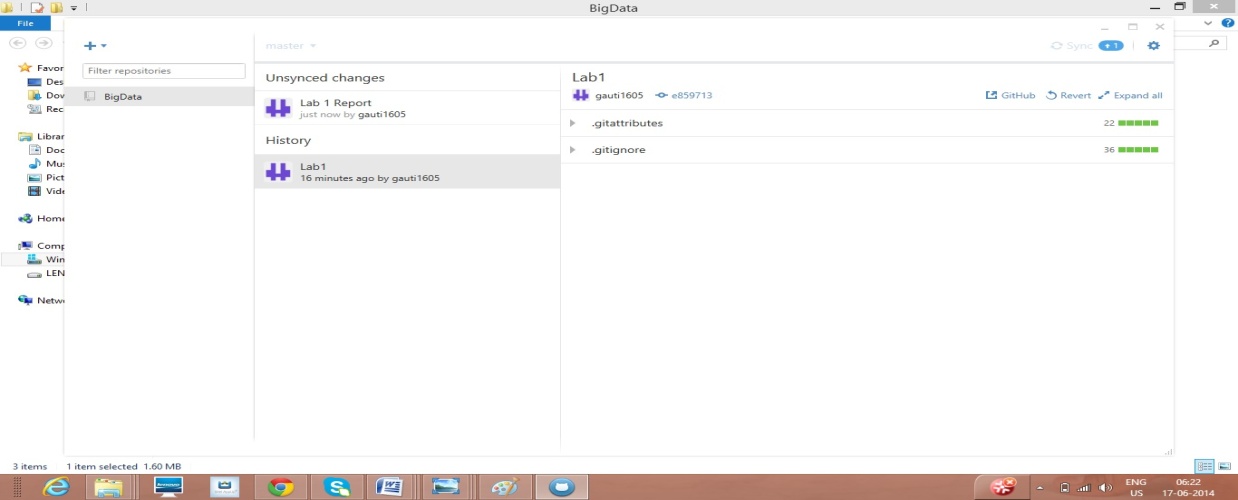
# Subtask1:

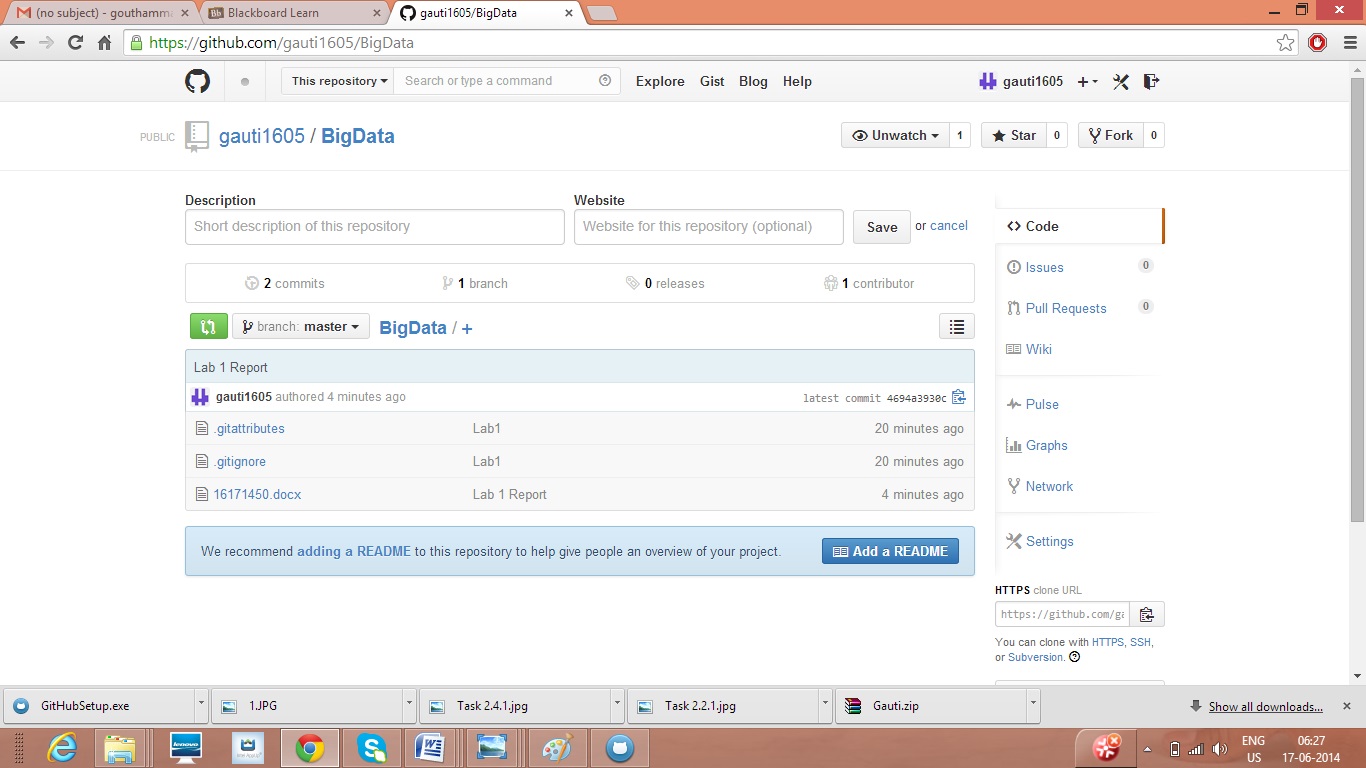
# How to create an account and deploy files to GitHub?

* Create an account with GitHub.



* Download and GitHub for Windows and transfer files into GitHub folder.
* Commit the Uncommitted changes, and then synchronise by clicking the ‘Publish Repository’ and ‘Sync’ buttons and the changes will be reflected in your online account.

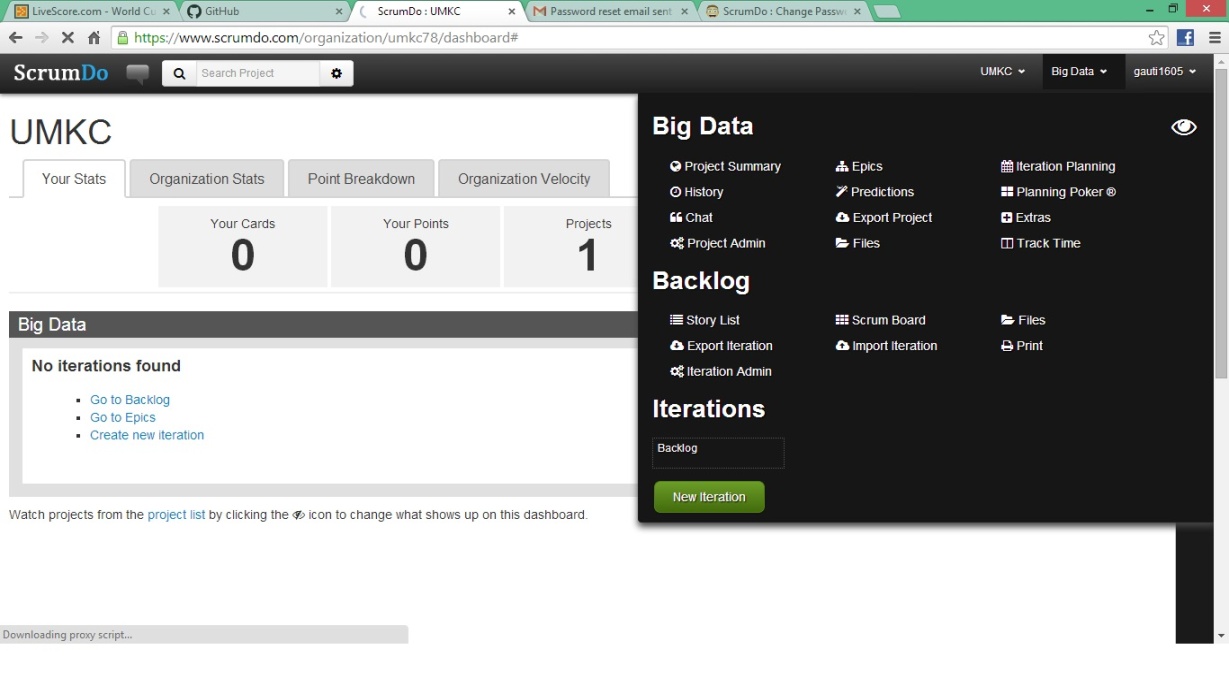




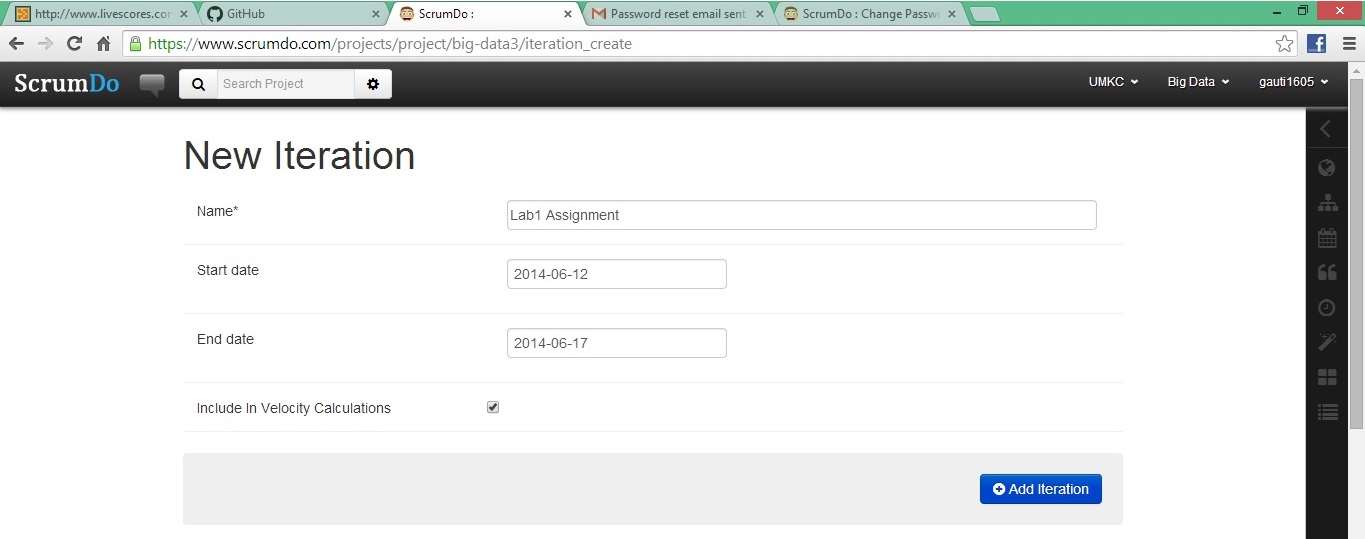
# Subtask2:

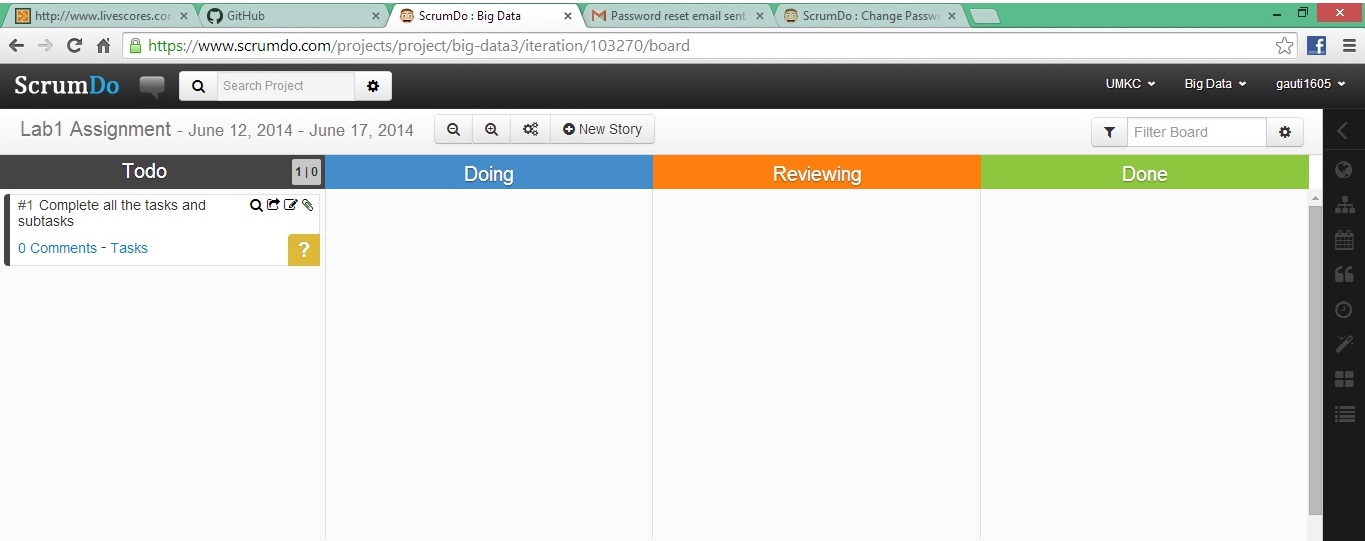
# How to create an account and design projects with Scrumdo?

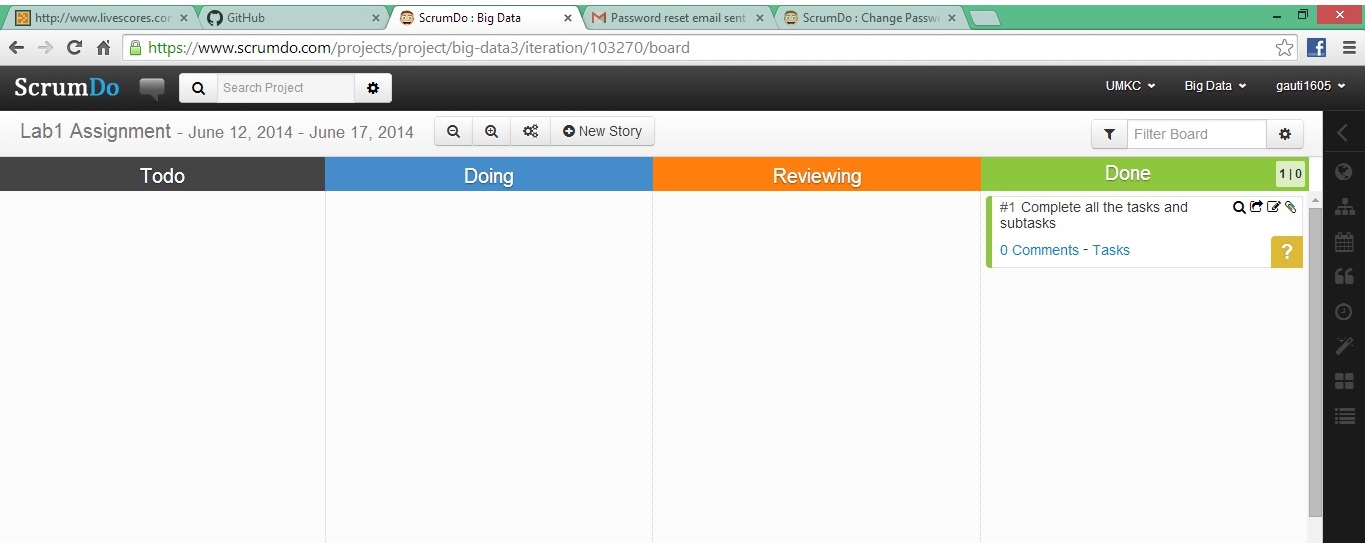
* Create an account with Scrumdo.
* Login to your account and create a new iteration.



* You can now create a new Iteration and see the status of your Iteration as captured below.







* Thus, projects can be designed using Scrumdo.