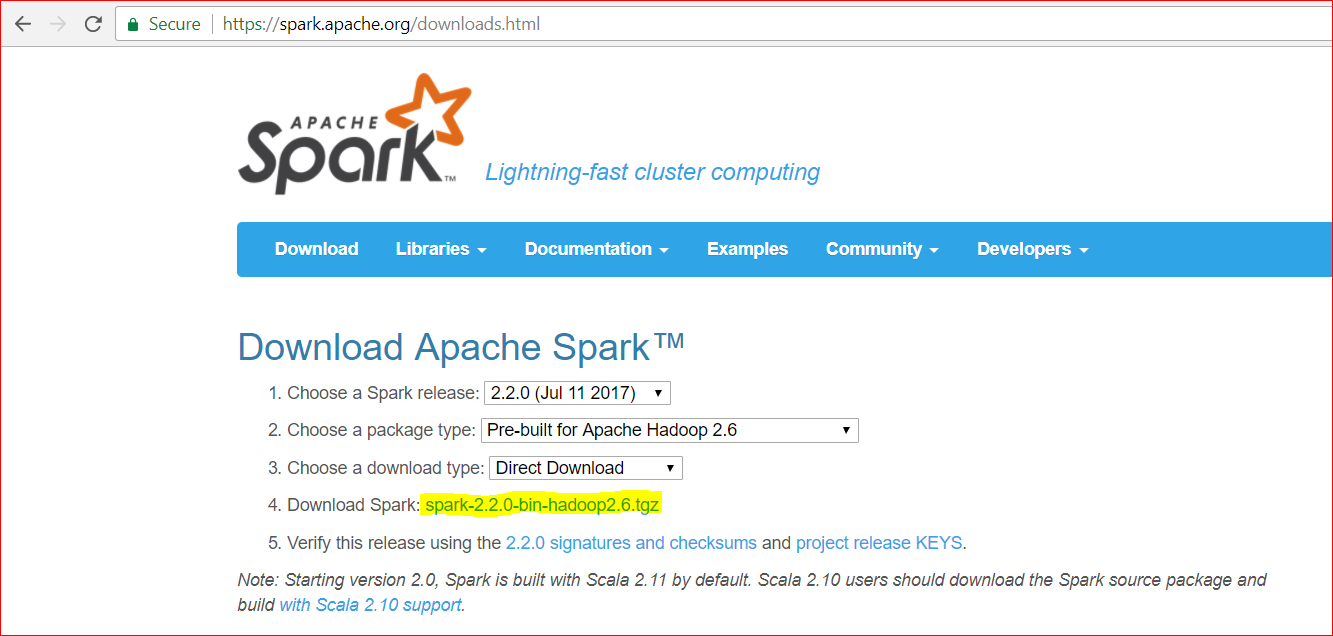
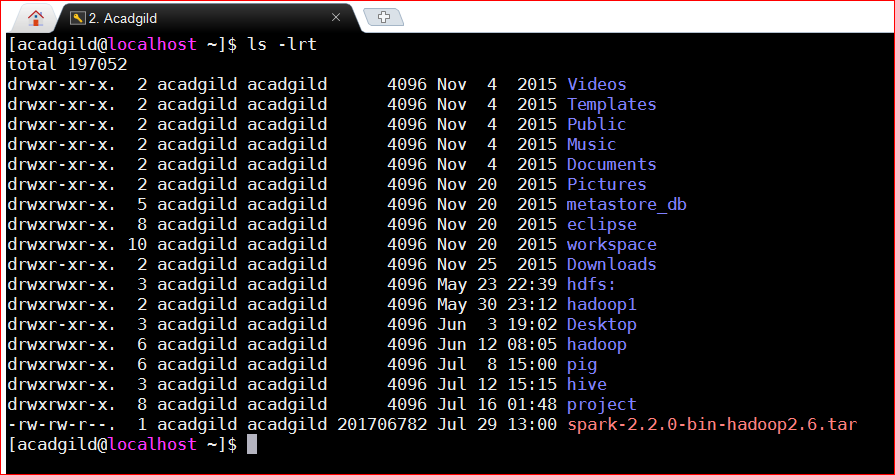
**Assignment 12.3: Spark Installation**

**Step 1:** Go to below URL and select appropriate version of spark and download clicking on the highlighted link.

<https://spark.apache.org/downloads.html>

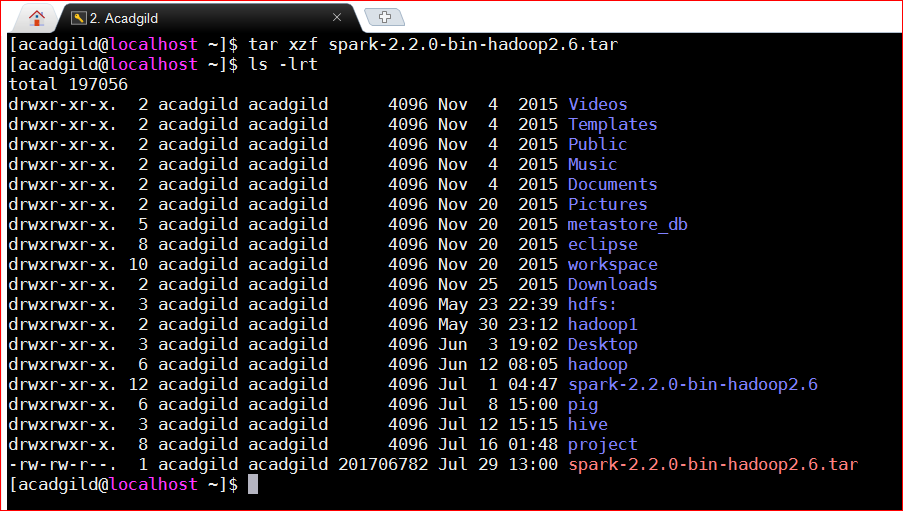


**Step 2:** Move downloaded file to Acadgild sandbox.

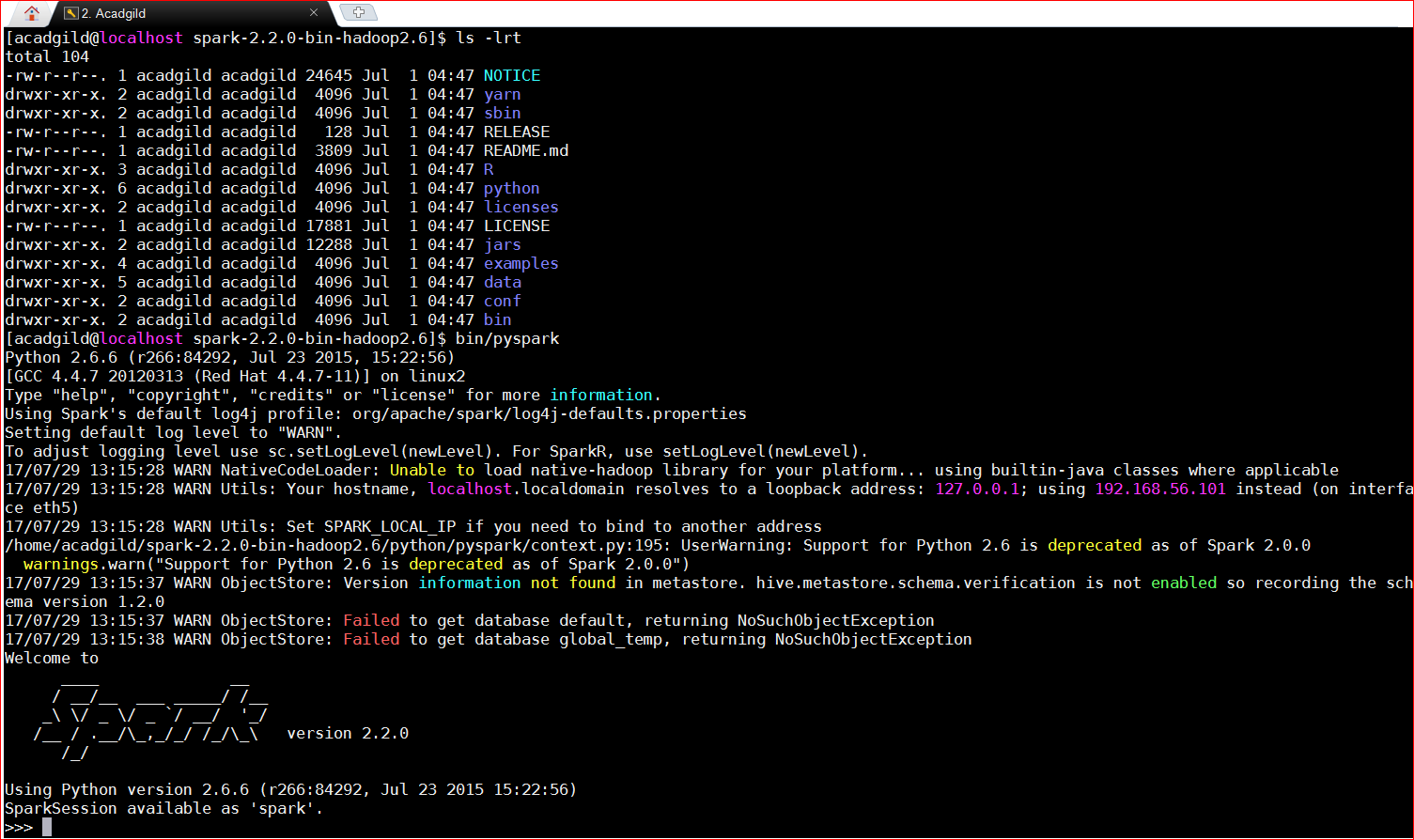


**Step 3:** Unzip/untar the file using below command.

**[acadgild@localhost ~]$ tar xzf spark-2.2.0-bin-hadoop2.6.tar**



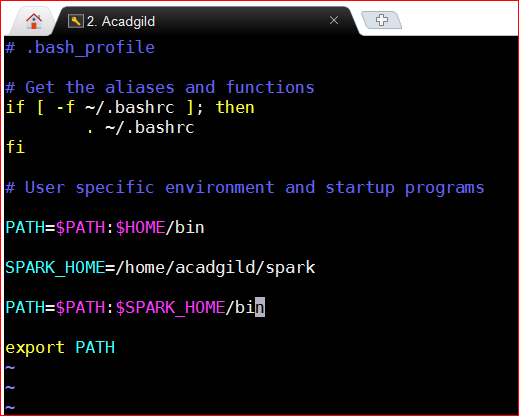
**Step 4:** Verify the content of directory **spark-2.2.0-bin-hadoop2.6** and try launching spark with python.



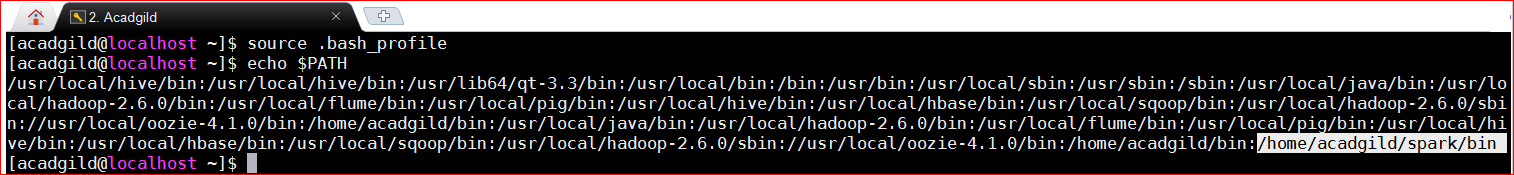
**Step 5:** close the pyspark (cntrl+d) and rename the directory **spark-2.2.0-bin-hadoop2.6** to **spark** for the simplicity.

**[acadgild@localhost ~]$ mv spark-2.2.0-bin-hadoop2.6 spark**

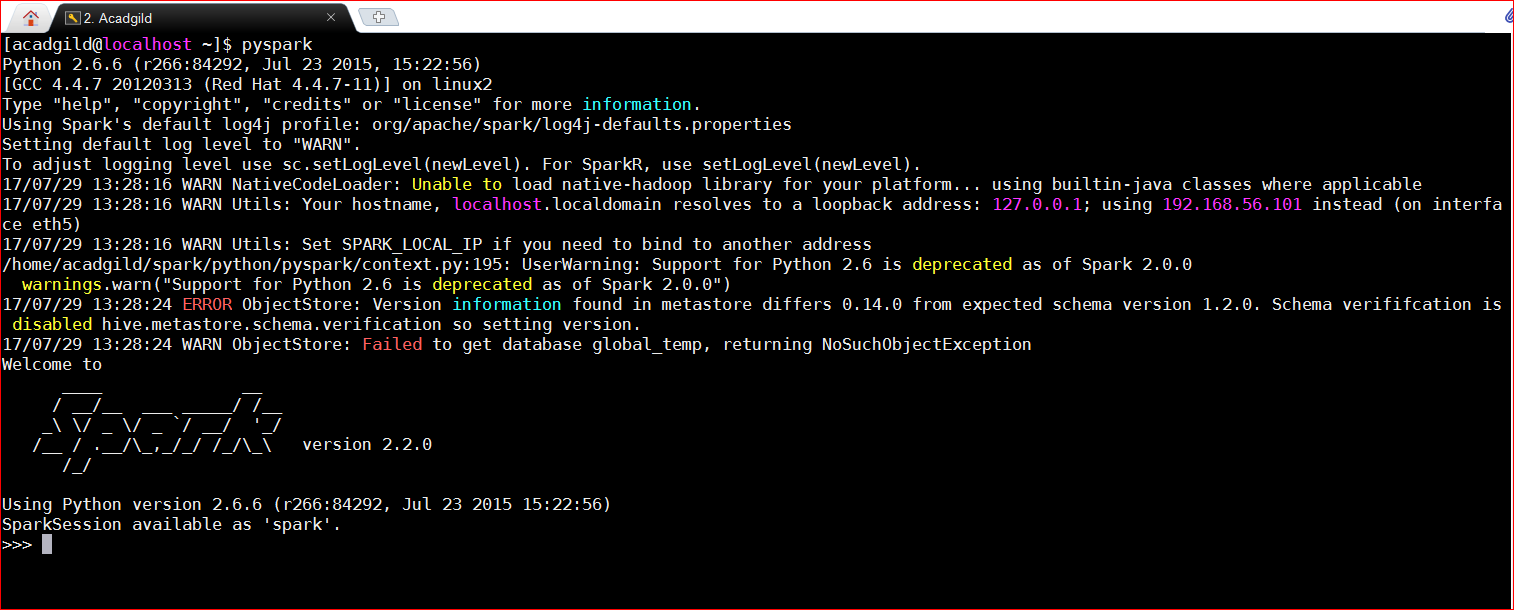
**Step 6:** update .bash\_profile file with the path of spark bin directory so that pyspark can be launched from any directory. Make sure that you **source .bash\_profile** file.



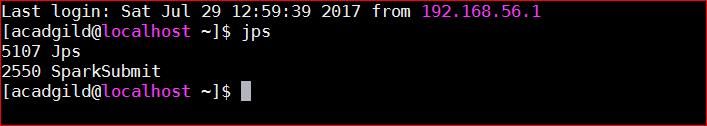
**Step 7:** Verify the $PATH.



**Step 8:** launch spark using pyspark from anywhere.



**Step 9:** launch another acadgild sandbox session and run jps command to see spark daemon running.



**Spark Installed successfully.**