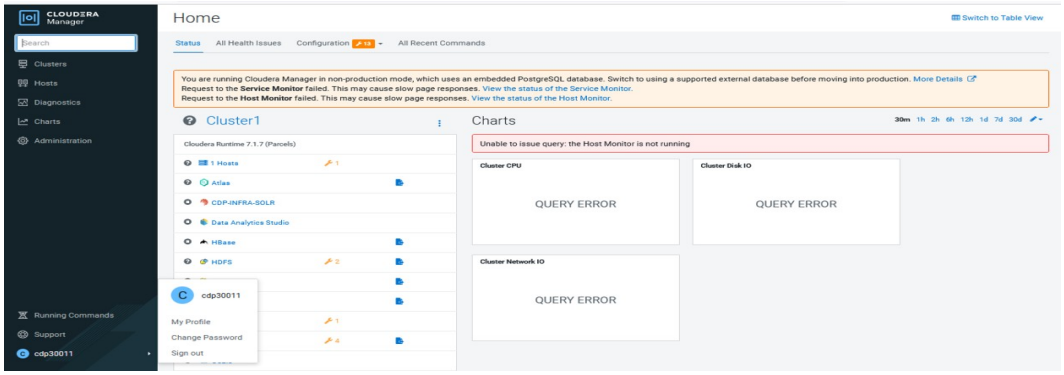
Prepared by : Amrit Chhetri, RCS(Siliguri)

CDP Developer Participant Environment Setup

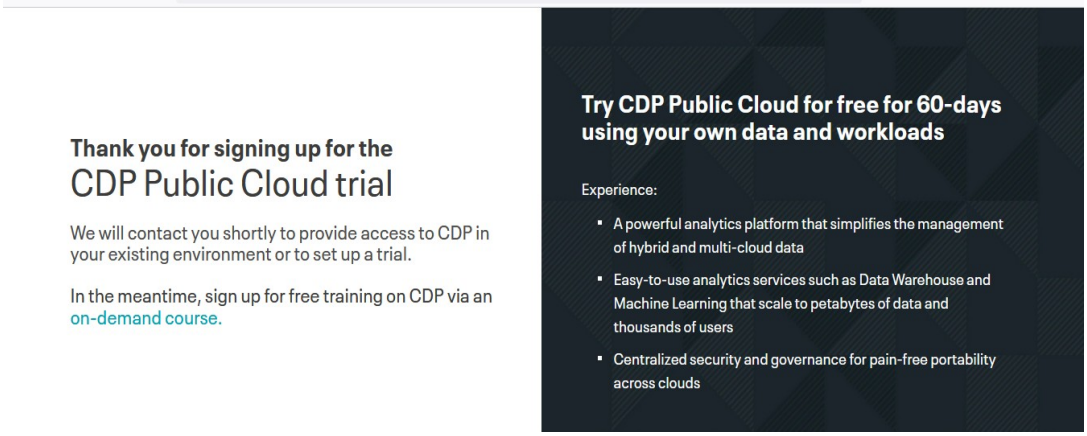
1. **Software Stacks for CDP 3001 Certification Training:**
    1. Cloudera Platform : CDP Private Cloud on GCP
    2. SDKs : Java, Scala, Python and Spark
    3. IDES : Jupyter Notebook, PyCharm Edu and Intellij Edu
    4. Remote Terminal : WinSCP and Smart Putty
    5. Utility : Notepad++, 7Zip
    6. DevSecOps and MLOps : Jenkins, Git, GitHub and Splunk

    For download links and version, refer "***CDP Developer Workstation for CDP 3001 Certification 1.0.docx***"

2. **CDP Cloud Systems:**
    1. **CDP Private Cloud:** Installation on Cloud/Data Centre, private installation. CDP Private Cloud CDP Private Cloud Address: http://35.223.46.173:7180 . It is installed on Ubuntu 20.04 LTS in GCP System with all Services.

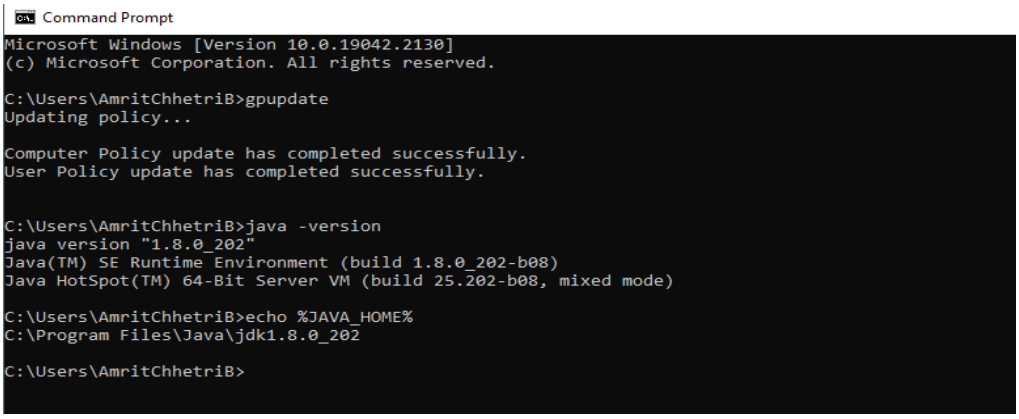


    2. **CDP Public Cloud:** No Installation, Cloudera SASS. CDP Public Cloud Data Services: URL: https://www.cloudera.com/campaign/try-cdp-public-cloud.html



3. **Software Stack for CDP-3001-Installations Steps:**
    1. **Java 1.8 and Java 11 Installation:**
        i. Download installers of JDK 1.8 and JDK 11
        ii. Run setup with default settings and click on finish
        iii. Post installation, configure JAVA_HOME in Environment Properties

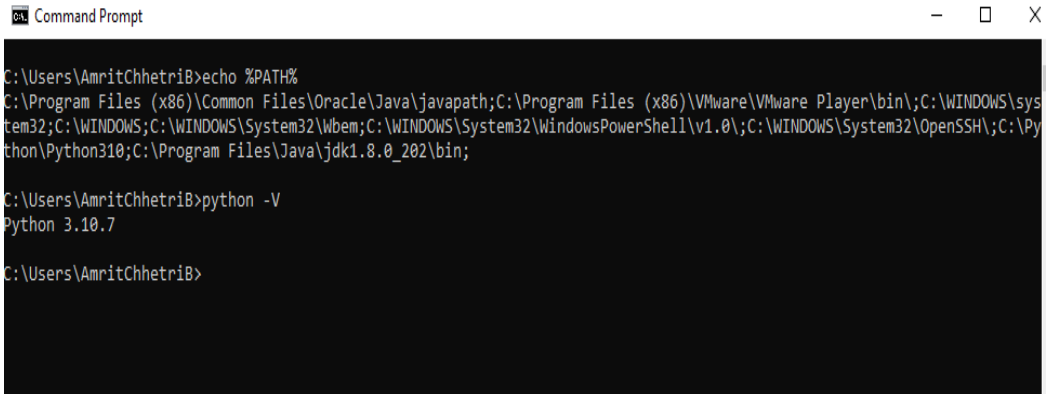
iv. Open Windows Command Prompt and type gpupdate, java -version and echo %JAVA_HOME%



Follow similar steps for Java 11 for Jenkins and features comparisons. Site https://www.oracle.com/in/java/technologies/javase/jdk11-archive-downloads.html

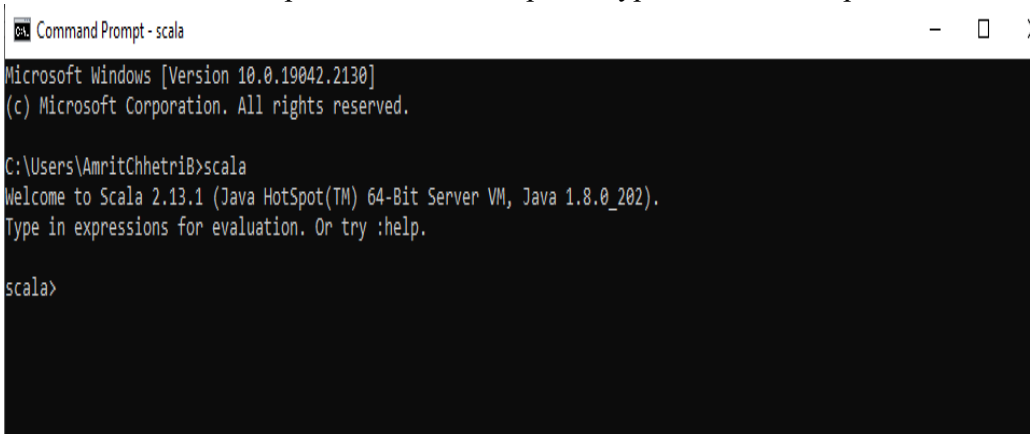
2. Python 3.6 and Python 3.10:

i. Get installer of python 3.10 and change installation directory to C:\Python\Python310
ii. Post Installation, add python home folder C:\Python\Python310 in path
iii. Type echo %PATH%, python -V

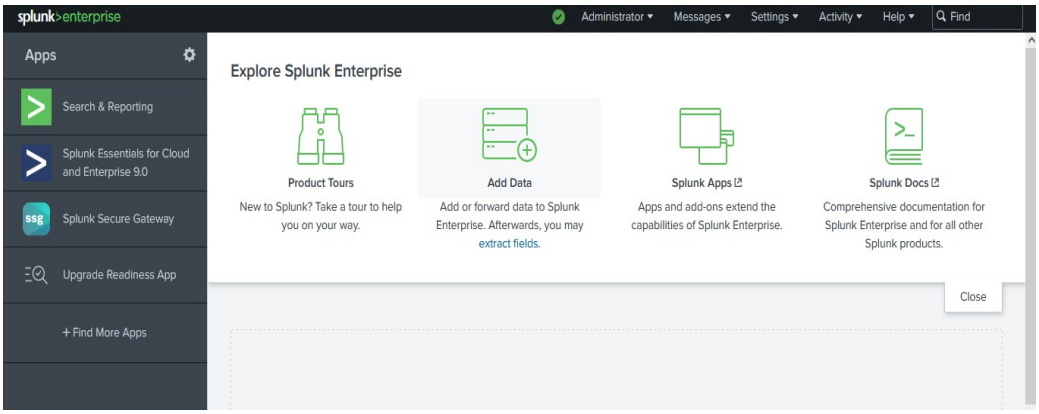


3. Scala Installation:

i. Get installer of Scala and install https://downloads.lightbend.com/scala/2.13.1/scala-2.13.1.msi
ii. Post Installation, open Windows Prompt and type scala on Prompt

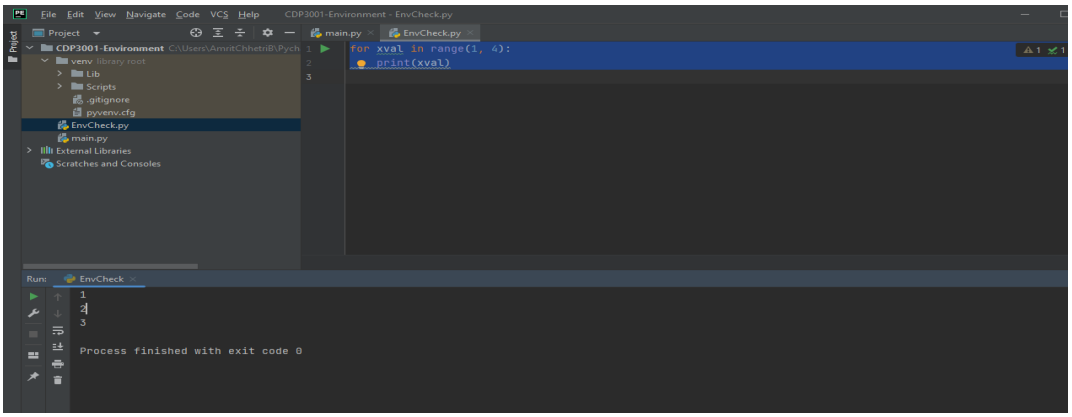


4. **Splunk Installation:**
    i. Get installer of Splunk and install with default settings, whenever promted create your user credentials - username/password: amritchhetrib/Qwerty123456
    ii. Post Installation, access by typing http://127.0.0.1:8000

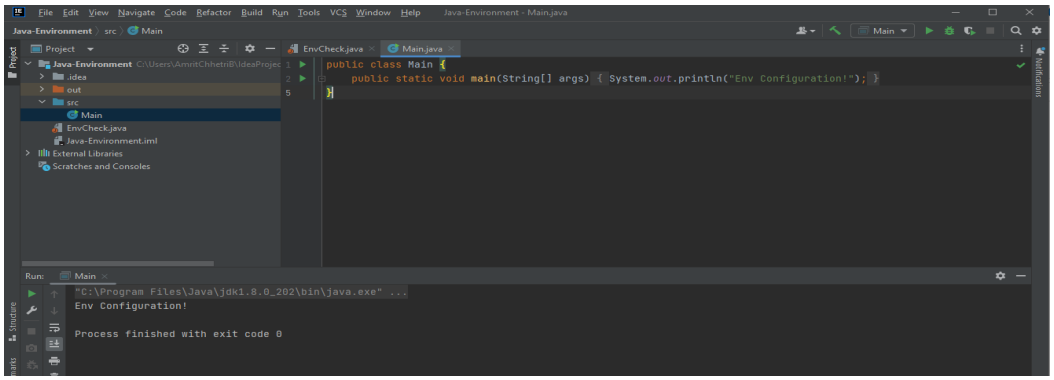


5. **PyCharm Installation:**
    i. Download/Get installer of PyCharm and install with default settings
    ii. **Post** Installation, open and click on New ->Project->Python-> create file EnvCheck.py and write

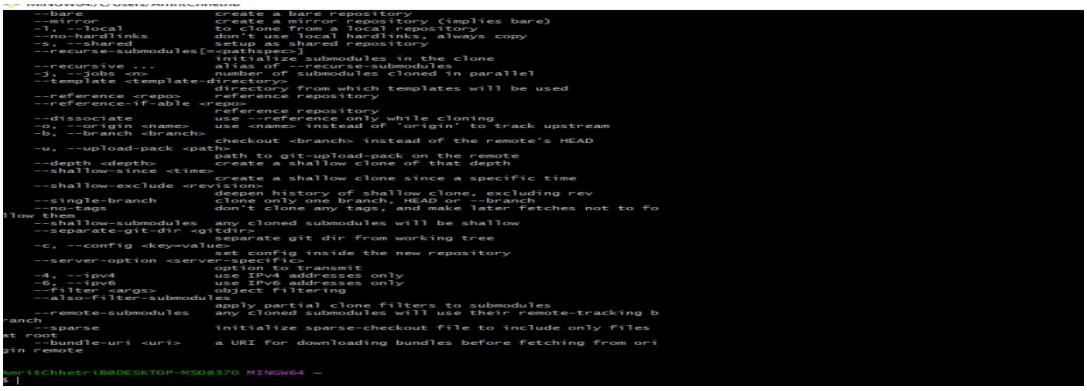
```
for xval in range(1, 4):
    print(xval)
```



6. **Intellij Edu Installation:**
    i. Download/Get installer of Intellij and install with default settings
    ii. Post Installation, Post Installation, open and click on New -Project-Java Project

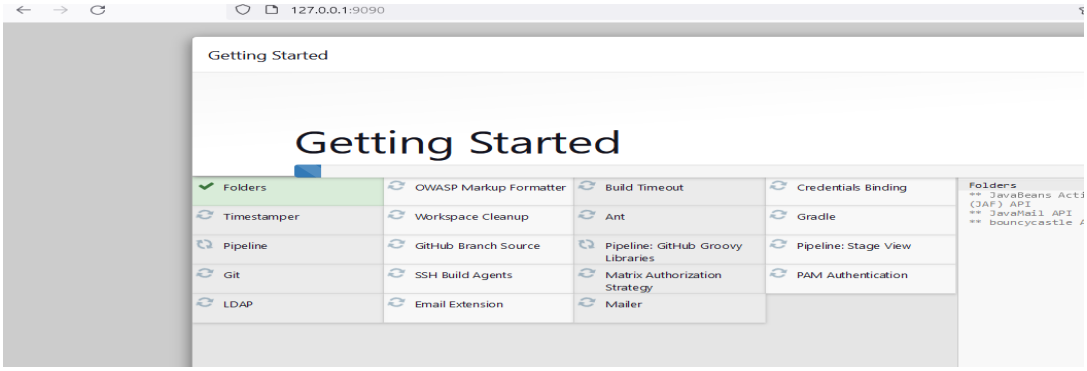


7. **Git Installation:**
    i. Get installer of Git and install with default settings
    ii. Post Installation, type git clone

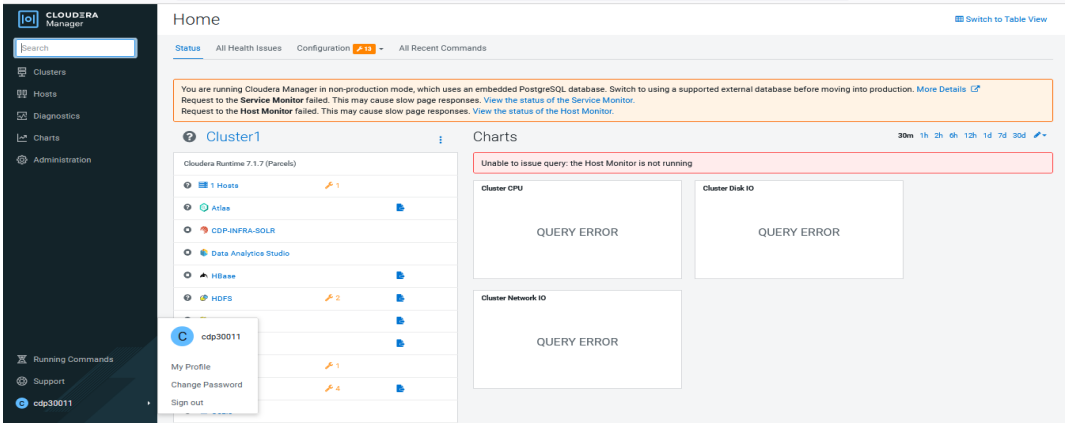


Create a folder CDP-3001-Certification in C or D and Examples
Open Windows Prompt, go to Examples and types
git -c http.sslVerify=false clone https://github.com/spark-examples/spark-scala-xamples.git

8. **Jenkins Installation:**
   i. Get installer of Jenkins and install with default settings
   ii. Post Installation, access typing htt://ip:9090 , as given during installation
   iii. Login with password in C:\ProgramData\Jenkins\.jenkins\secrets\initialAdminPassword



4. **Software Stack for CDP-3001-** Check **Steps:**
   1. **CDP Private Cloud(setup for Training):**
      Web Address: http://35.223.46.173:7180
      Username: cdp30011
      Password : Qwerty123456$
      Access Type: Private Cloud



   2. CDP Public Cloud (SAAS) – SAAS Access :

   3. HDFS Access: with SSH
      IP Address: 35.223.46.173 Port: 22
      Username: cdp30011
      Password : Qwerty123456$

Create directory to load file:
Leave Safe Mode: sudo -u hdfs hdfs dfsadmin -safemode leave
sudo -u hdfs hadoop fs -mkdir /cdproot
sudo -u hdfs hadoop fs -mkdir /cdproot/spark
sudo -u hdfs hadoop fs -mkdir /cdproot/nifi
sudo -u hdfs hadoop fs -mkdir /cdproot/spark
sudo -u hdfs hadoop fs -ls /cdproot
sudo -u hdfs hadoop fs -ls /cdproot
sudo -u hdfs hadoop fs -put Data.txt /cdproot/data
sudo -u hdfs hadoop fs -cat /cdproot/data/Data.txt
Deleting files: sudo -u hdfs hadoop fs -rmdir /cdproot/spark