

Assignment 12.1:

Problem Statement:

Follow the below link document steps to download and import Acadgild Spark VM in the Oracle Virtual Box.

<https://drive.google.com/file/d/0ByJLBTmJojizQ2hLc1RLX3pST0U/view>

Steps:

Apache Spark was downloaded and installed. Attached below are the screen shots:

```
[acadgild@localhost ~]$ ls
Assignment3.2-0.0.1-SNAPSHOT.jar      hadoop
Assignment3.3-0.0.1-SNAPSHOT.jar      hdfs:
assignment11.txt                      hive
company.java                         hive-site.xml
customer.java                       mapreduce-0.0.1-SNAPSHOT.jar
derby.log                           max-temp.txt
Desktop                             metastore_db
Documents                           Music
Downloads                           mysql-connector-java-5.1.44.zip
eclipse                             Pictures
eclipse-jee-neon-M3-linux-gtk-x86_64.tar.gz  pig
employee.java                       pig_1510851976774.log
[acadgild@localhost ~]$ ls spark-2.2.0-bin-hadoop2.7
bin  conf  data  derby.log  examples  jars  LICENSE  licenses  metastore_db  NOTICE  python  R  README.md  RELEASE  sbin  yarn
[acadgild@localhost ~]$ ls spark-2.2.0-bin-hadoop2.7/bin/
beeline      employee.java  metastore_db  run-example  spark-class.cmd  spark-shell  spark-submit
beeline.cmd  find-spark-home  pyspark      run-example.cmd  sparkR           spark-shell2.cmd  spark-submit2.cmd
company.java  load-spark-env.cmd  pyspark2.cmd  spark-class    sparkR2.cmd      spark-shell.cmd  spark-submit.cmd
derby.log    load-spark-env.sh  pyspark.cmd   spark-class2.cmd  sparkR.cmd       spark-sql

[acadgild@localhost ~]$ cat ~/.bash_profile
# ~/.bash_profile

# Get the aliases and functions
if [ -f ~/.bashrc ]; then
    . ~/.bashrc
fi

# User specific environment and startup programs

PATH=$PATH:$HOME/bin:/home/acadgild/spark-2.2.0-bin-hadoop2.7/bin

export PATH
[acadgild@localhost ~]$ spark-shell
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
17/11/26 20:11:48 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/11/26 20:11:50 WARN Utils: Your hostname, localhost.localdomain resolves to a loopback address: 127.0.0.1; using 192.168.43.85 instead (on interface eth5)
17/11/26 20:11:50 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
17/11/26 20:12:43 WARN ObjectStore: Failed to get database global_temp, returning NoSuchObjectException
Spark context Web UI available at http://192.168.43.85:4040
Spark context available as 'sc' (master = local[*], app id = local-1511707321679).
Spark session available as 'spark'.
Welcome to

  ____
 /  __ \
/   /  \
/_____/

 version 2.2.0

Using Scala version 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_65)
Type in expressions to have them evaluated.
Type :help for more information.

scala> █
```