How to compile and run your Scala/Spark Program

(Steps 1 through 4 are same for scala or spark code.)

1. Create a folder under which all your spark or scala applications will be kept. Let the path to the folder be

/home/student/studentName/SparkProjects

2. Let the app name be WordCount. Create the directories as below.

/home/student/studentName/SparkProjects/WordCount/src/main/scala/

3. Place the file containing the scala code in the "wordcount" folder.

Let us assume that the file is "wordcount.scala". Hence

ls /home/student/studentName/SparkProjects/WordCount/src/main/scala/wordcount.scala should find the file.

4. In /home/student/studentName/SparkProjects/WordCount, create the sbt dependencies file if required (our vanilla scala apps will probably not need it while our spark apps will probably need it): WordCount.sbt

(Steps 5 and 6 for spark code)

5. In folder /home/student/studentName/SparkProjects/WordCount, create a java package of your code by running:

sbt package

6. Execute your code using spark-submit as follows:

(note if your program expects arguments, specify them after the jar file)

spark-submit --master local[\*] target/scala-2.11/wordcount\_2.11-1.0.jar

(Steps 7 and 8 if your code is scala code)

7. In folder /home/student/studentName/SparkProjects/WordCount, run the compile command as:

sbt compile

8. run the code as:

sbt run

Note: if the scala app requires parameters, then they have to be passed as follows:

sbt “run param1 param2”

Note: if the param is a string, then it needs to be passed as below:

sbt “run \”This string is a parameter or argument required by the app.\” “