800+ Q&As | Logout | Contact

Java-Success.com

Prepare to fast-track, choose & go places with 800+ Java & Big Data Q&As with lots of code & diagrams.

search here ...

Go

300+ Java FAQs ▼ 300+ Big Data FAQs ▼ Home

Membership • Your Career ▼

Home > bigdata-success.com > Tutorials - Big Data > TUT - Starting Spark & Scala >

06: Setting up Spark-shell on Mac & getting started

06: Setting up Sparkshell on Mac & getting started



Posted on November 8, 2018

This tutorial outlines the basic steps to get started with Spark on Mac OS.

1. Install Xcode

Xcode can be installed via Apple appstore. Xcode is Apple's Integrated Development Environment (IDE). Xcode is a large suite of software development tools and libraries from Apple.

2. Install the Apple command line tools

300+ Java **Interview FAQs**

300+ Java FAQs



16+ Java Key Areas Q&As



150+ Java Architect FAQs



80+ Java Code Quality Q&As



150+ Java Coding 0&As



300+ Big Data **Interview FAQs**

300+ Big Data FAOs (



Tutorials - Big Data



TUT - M Starting Big Data

TUT - Starting Spark Scala

Once Xcode is installed, install the command line tools via "Xcode menu" -> "preferences" -> "command lines tools", and click the install button. This may take a while to install. Once installed you can verify with a Terminal window

```
1
2 $ xcode-select -h
3
```

The Xcode Command Line Tools are part of XCode. The Xcode Command Line Tools include a GCC compiler, and many common Unix-based tools require the GCC compiler

3. Install homebrew

Homebrew is a package manager for OS. On a Terminal window type

```
1 | $ ruby -e "$(curl -fsSL https://raw.github.com/mxc] 3
```

Verify if brew is installed properly by typing the following on a Terminal window:

```
1 | $ brew doctor 3
```

4. Install Java & Scala

```
1
2 $ brew cask install java
3
```

TUT - Starting with Python

TUT - Kafka

TUT - Pig

TUT - Apache Storm

TUT - Spark Scala on Zeppelin

TUT - Cloudera

TUT - Cloudera on Docker

TUT - File Formats

TUT - Spark on Docker

TUT - Flume

TUT - Hadoop (HDFS)

TUT - HBase (NoSQL)

TUT - Hive (SQL)

TUT - Hadoop & Spark

TUT - MapReduce

TUT - Spark and Scala

TUT - Spark & Java

TUT - PySpark on Databricks

TUT - Zookeeper

800+ Java Interview Q&As

300+ Core Java Q&As



300+ Enterprise Java Q&As



150+ Java Frameworks Q&As



120+ Companion Tech Q&As



Tutorials -Enterprise Java



```
1 | 2 | $ brew install scala 3 |
```

5. Install Apache Spark

```
1
2 $ brew install apache-spark
3
```

You can check where a package has been installed with:

```
1 | $ brew info apache-spark 3
```

6. Spark shell to run Spark code interactively

```
1
2
  $ spark-shell
3
  Welcome to
2
3
4
5
6
7
  Using Scala version 2.11.8 (Java HotSpot(TM) 64-B
8
  Type in expressions to have them evaluated.
10 Type :help for more information.
11
12 | scala>
13
```

7. Spark example in Scala

Use ":paste" to paste the following code. After pasting the code press "ctrl-D" to run the Spark code.

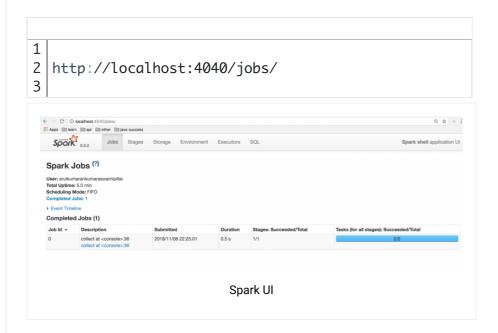


```
1
2
   scala> :paste
3
   // Entering paste mode (ctrl-D to finish)
4
5
   import org.apache.spark.sql.SparkSession;
6
   import org.apache.spark.sql.Dataset;
   val spark = SparkSession.builder.appName("Simple /
8
9
                                    .config("spark.mas
10
                                    .getOrCreate()
11
12 | val list = List(1,2,3,4,5)
13 | val rdd = spark.sparkContext.parallelize(list, 2)
14
15 | val multipleBy2Rdd = rdd.map(_ * 2)
16 | println(multipleBy2Rdd.collect().toList)
17
```

Use "ctrl+c" to quit the Spark REPL.

8. Spark UI

When the spark-shell is running, you can bring up the Spark UI in a browser.



The .bash_profile

This assumes that Hadoop has been installed as per one of the other tutorials so that you can do



\$(hadoop classpath). Also, when you install it with "brew install" as opposed downloading the *.tar.gz and unzipping it.

```
1 //....
2 export SPARK_HOME=/usr/local/Cellar/apache-spark/2
3 export SPARK_DIST_CLASSPATH=$(hadoop classpath)
4 export PATH=$SPARK_HOME/bin:${PATH}
5 //...
6
```

Activate the change:

```
1 | $ source ~/.bash_profile
```

Starting the Spark master

```
1 sudo spark-class org.apache.spark.deploy.master.Mas
2
```

Note down the **host:port** where the Spark master is listening via "http://localhost:8080".

Spark worker joins the Spark master

In a separate shell window start the worker node, specify the master host:location to join.

```
1 sudo spark-class org.apache.spark.deploy.worker.Wor
2
```

You can submit a job as shown below. If you don't specify the URL of the master, it defaults to "local[*]".





```
2 | $ spark-submit --master spark://192.168.0.18:7077
```

05: Setting up & getting started with Spark local mode with Sbt & Scala

07: Getting started with Spark-submit >>

Disclaimer

The contents in this Java-Success are copyrighted and from EmpoweringTech pty ltd. The EmpoweringTech pty ltd has the right to correct or enhance the current content without any prior notice. These are general advice only, and one needs to take his/her own circumstances into consideration. The EmpoweringTech pty ltd will not be held liable for any damages caused or alleged to be caused either directly or indirectly by these materials and resources. Any trademarked names or labels used in this blog remain the property of their respective trademark owners. Links to external sites do not imply endorsement of the linked-to sites. Privacy Policy

© 2022 java-success.com

Top

