Setting UP Spark 2.0 environment on intellij community edition version 2016.2.2

23/08/2016

Laser Nahoom-Kabakov

Contents:

Download and install latest intellij IDEA community edition

Click the link below for the download:

For windows:

https://www.jetbrains.com/idea/download/download-thanks.html?platform=windows&code=IIC

for linux:

https://www.jetbrains.com/idea/download/#section=linux

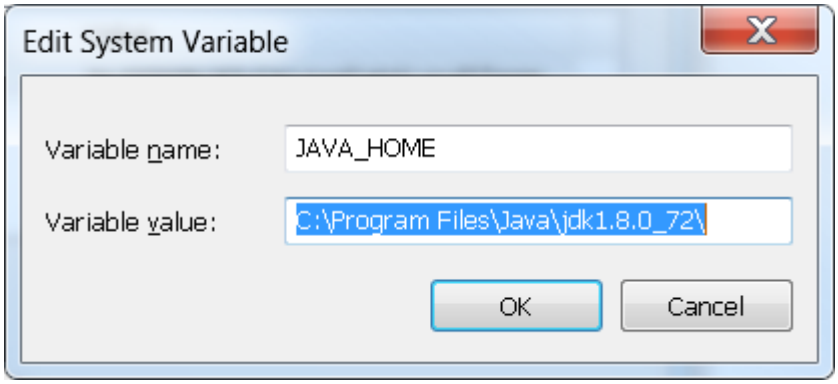
Download and install the latest Oracle/Sun JDK

Click the link below for the download:

http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html

Set up your JAVA_HOME variable

Make sure to set up your JAVA_HOME to: “C:\Program Files\Java\<YOUR_JVM_LOCATION>”, pictured below:



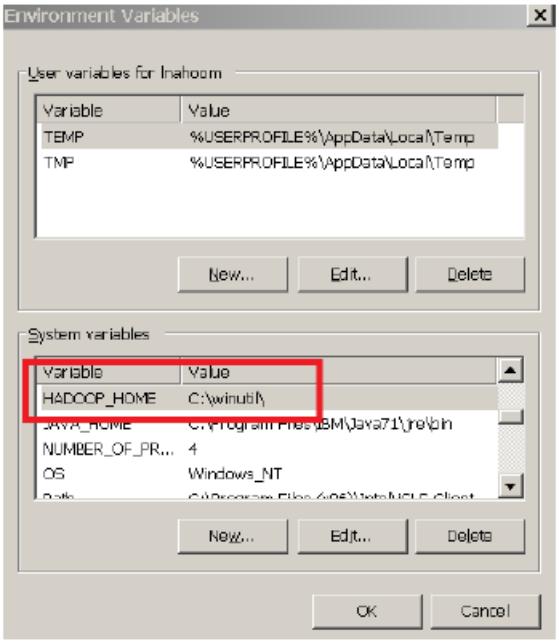
Download and configure winutil (windows only)

Download winutil.exe from here:

http://public-repo-1.hortonworks.com/hdp-win-alpha/winutils.exe

Make the following path on your file system and add the winutils.exe there
`C:\winutil\bin\winutils.exe`

Create a new Environment variable named HADOOP_HOME and point it to C:\winutil\



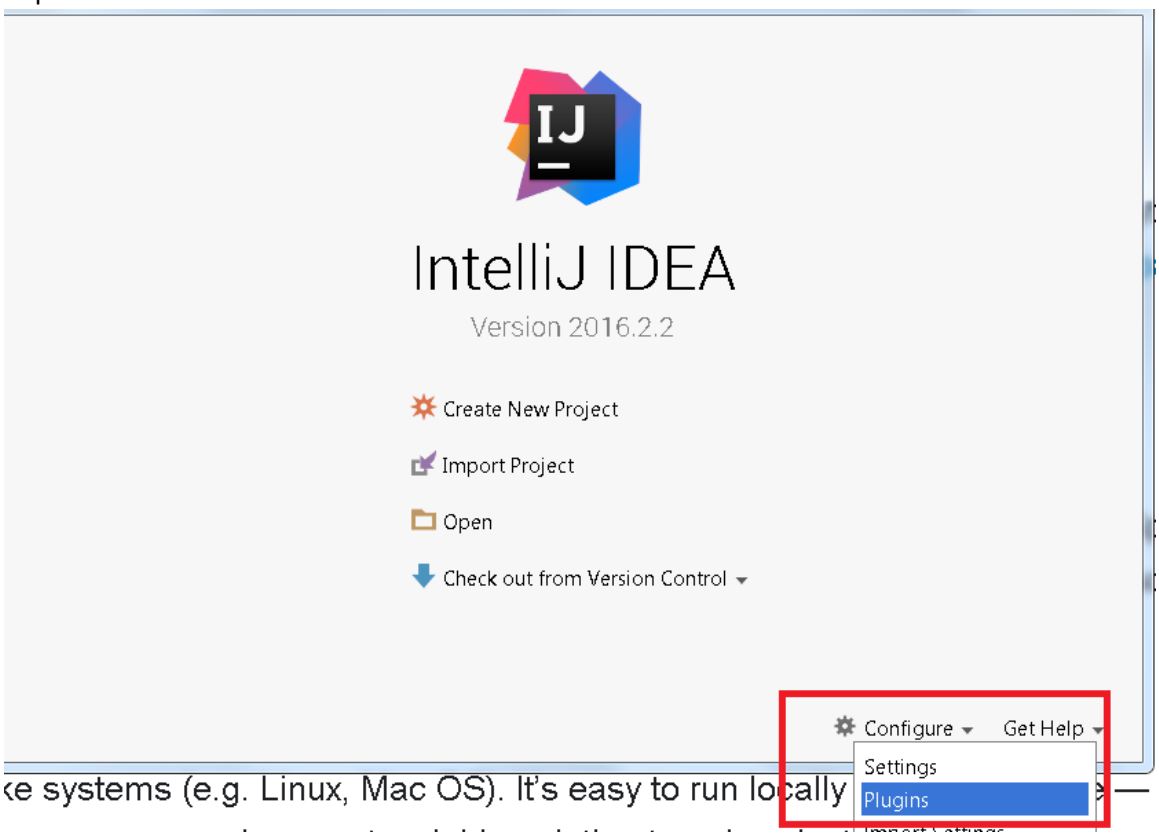
Start intellij IDEA

Skip all configurations at startup and don't install anything. Until you get to the "create new project screen", pictured below:

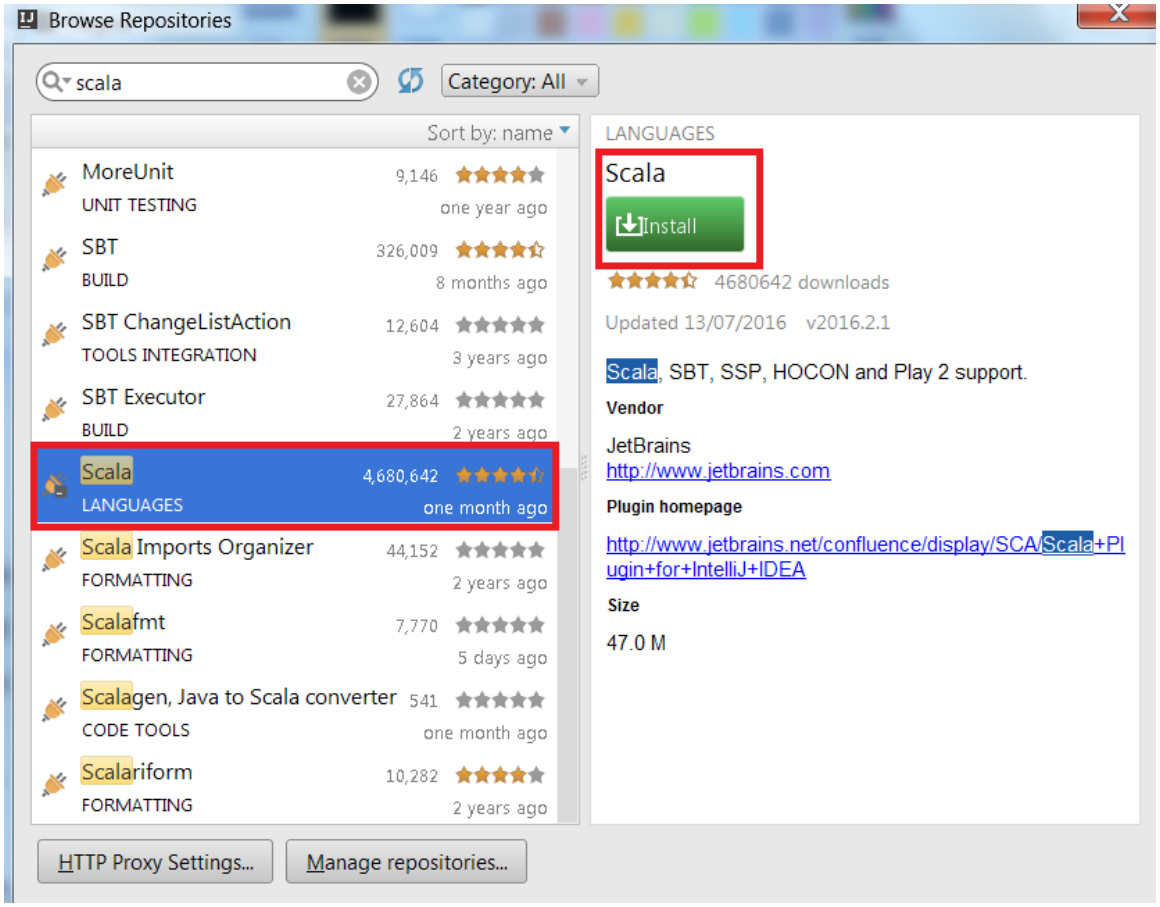

IntelliJ IDEA
Version 2016.2.2
Create New Project
Import Project
Open
Check out from Version Control

Install Scala and SBT plugins for Intellij

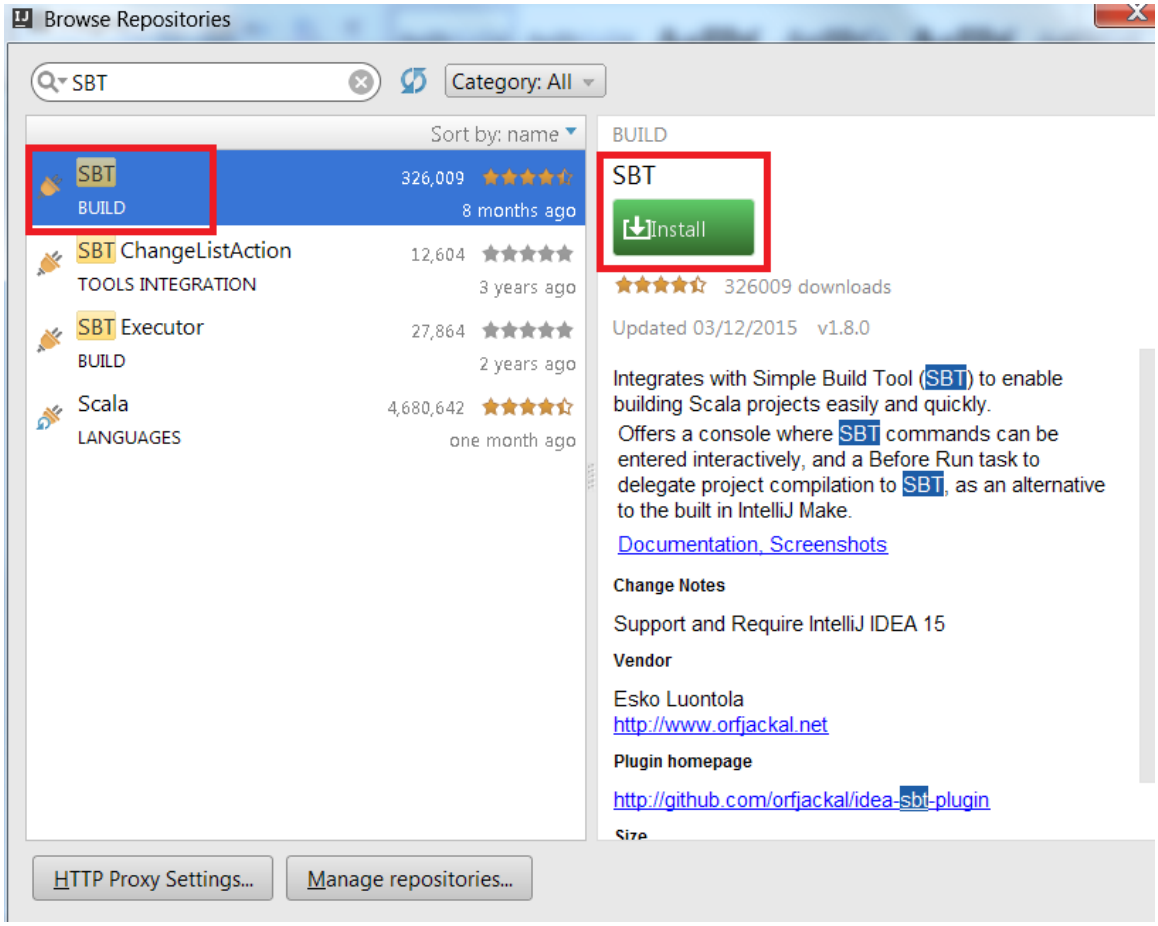
1. On this page go to the configure→plugins→ Browse Repositories



2. Install the Scala plugin listed below:



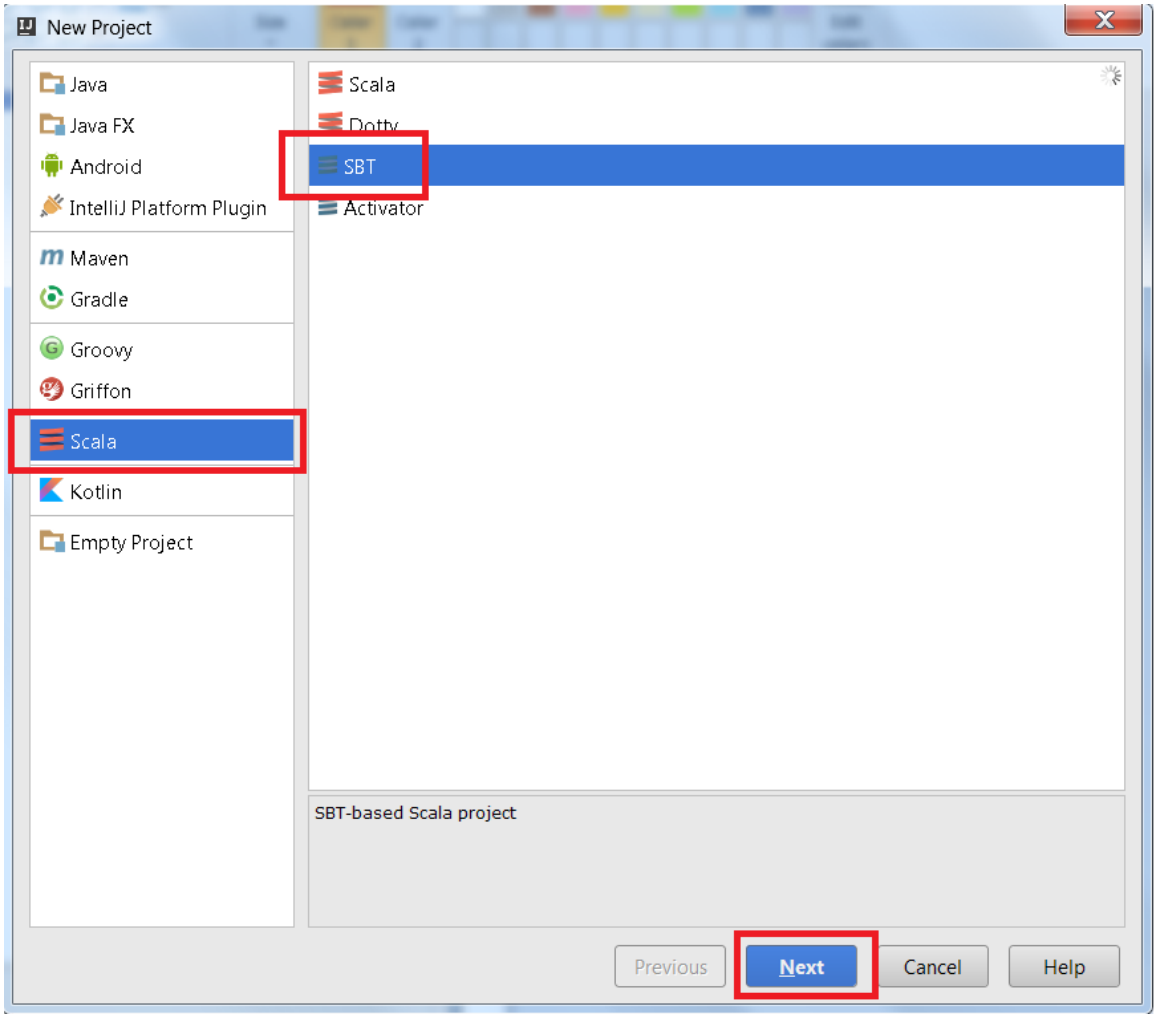
3. Install the SBT plugin listed below



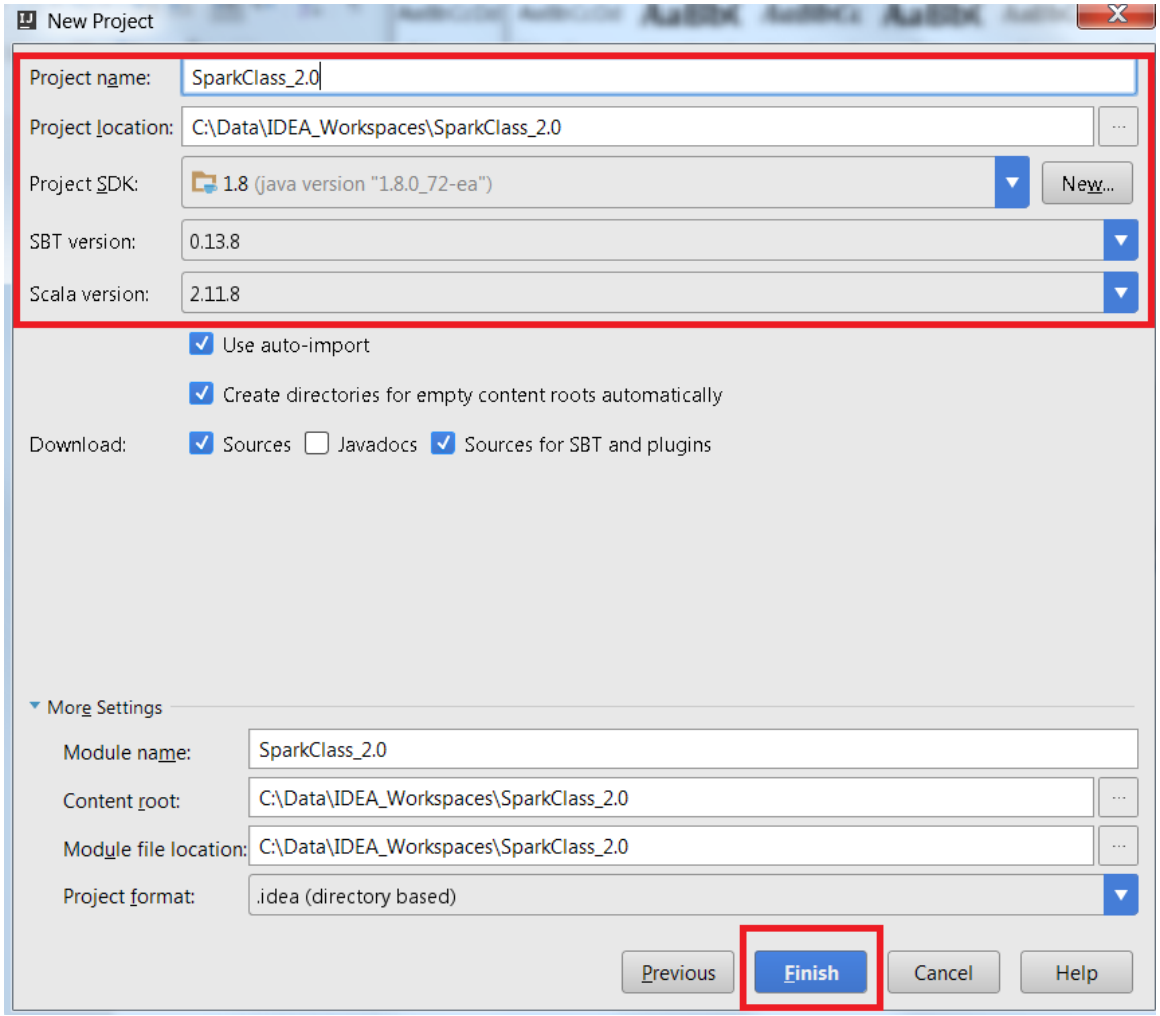
4. Restart intellij IDEA as requested

Create and configure a new Scala and SBT project

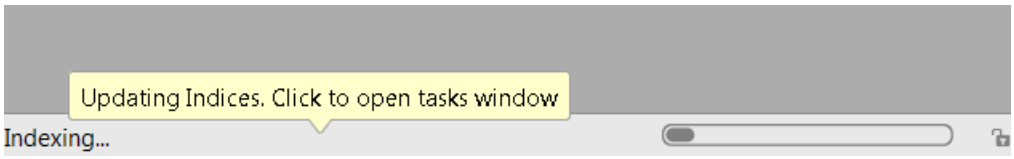
1. Create a new Scala SBT project



2. Create the project with the following configuration settings

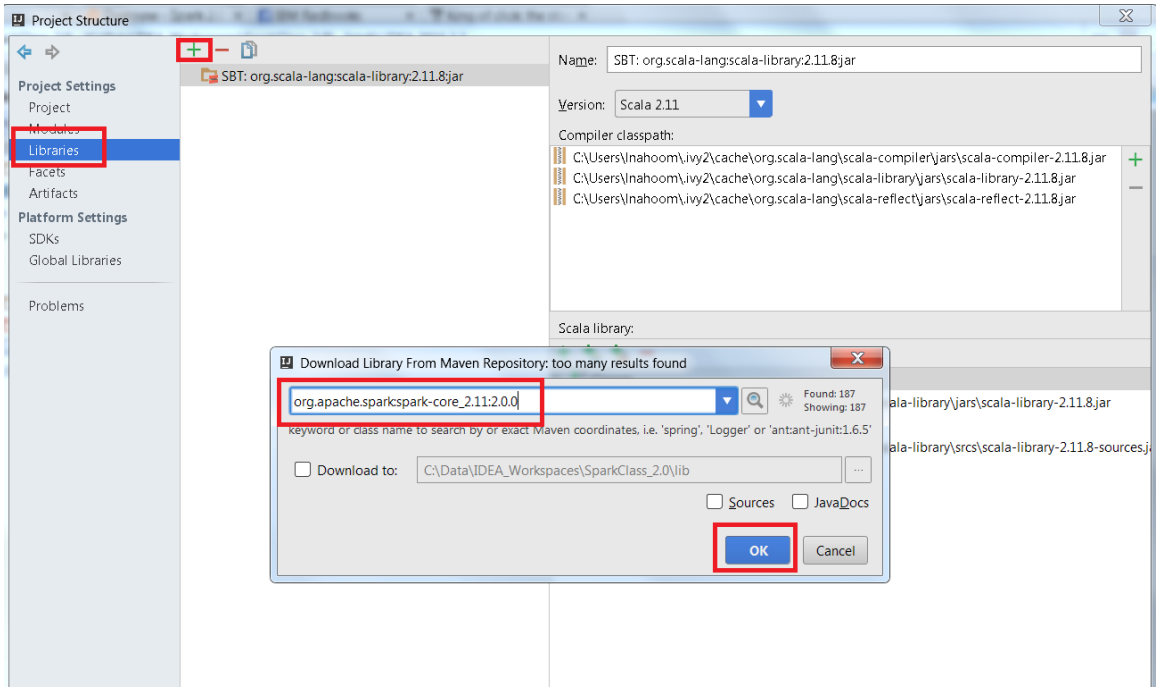


3. Wait for the indexing completion!!



Configure libraries for the newly created project

1. Open the module setting and go to the Libraries and add the following libraries from maven as shown in the screenshot below:

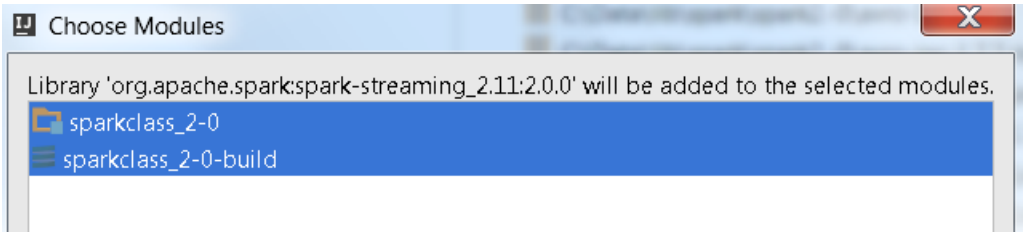


Libraries to add:

Package	Artifact ID
org.apache.spark:spark-core_2.11:2.0.0	spark-core_2.11
org.apache.spark:spark-sql_2.11:2.0.0	spark-sql_2.11
org.apache.spark:spark-mllib_2.11:2.0.0	spark-mllib_2.11
org.apache.spark:spark-streaming_2.11:2.0.0	spark-streaming_2.11
org.apache.spark:spark-graphx_2.11:2.0.0	spark-graphx_2.11

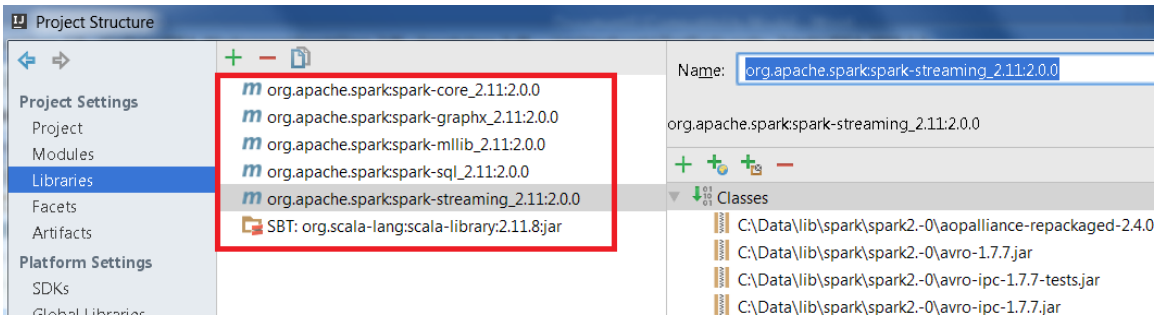
2. When adding each library add the libraries to the following projects:

   choose to add support for the following modules:



Validate your configuration:

After the addition of all of the libraries, your project libraries should look like this:



Test your setup

Code and run the following Scala test object to validate your configuration works

```
import org.apache.spark.{SparkConf, SparkContext}

/**
  * Created by lnahoom on 22/08/2016.
  */
object TestScala {

  def main(args: Array[String]): Unit = {
    val conf = new SparkConf()
    conf.setAppName("Datasets Test")
    conf.setMaster("local[2]")
    val sc = new SparkContext(conf)
    println(sc)
  }
}
```

Correct output should look like this:

Congratulations, you have been able to set up the Intellij with your Spark 2.0 support.