单机运行spark程序（IDEA）

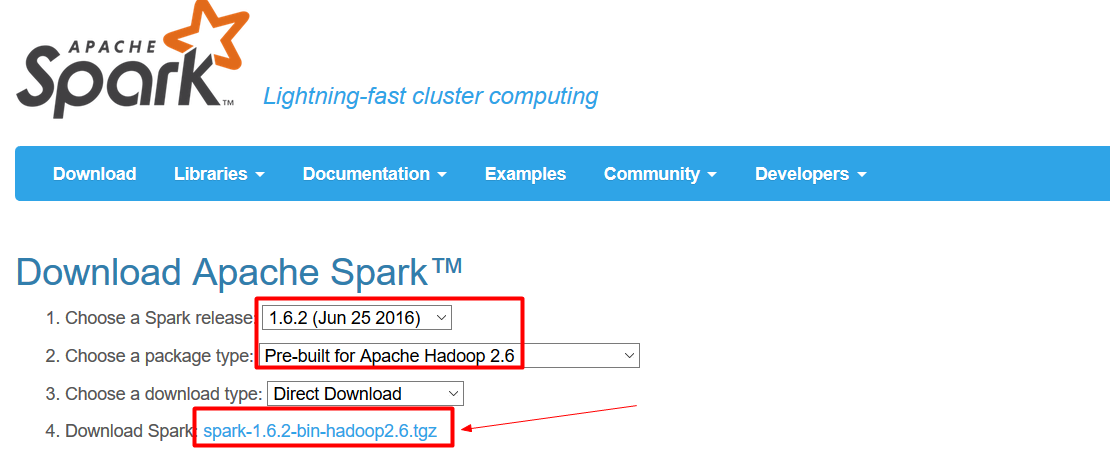
# 环境准备

## Jdk：1.8

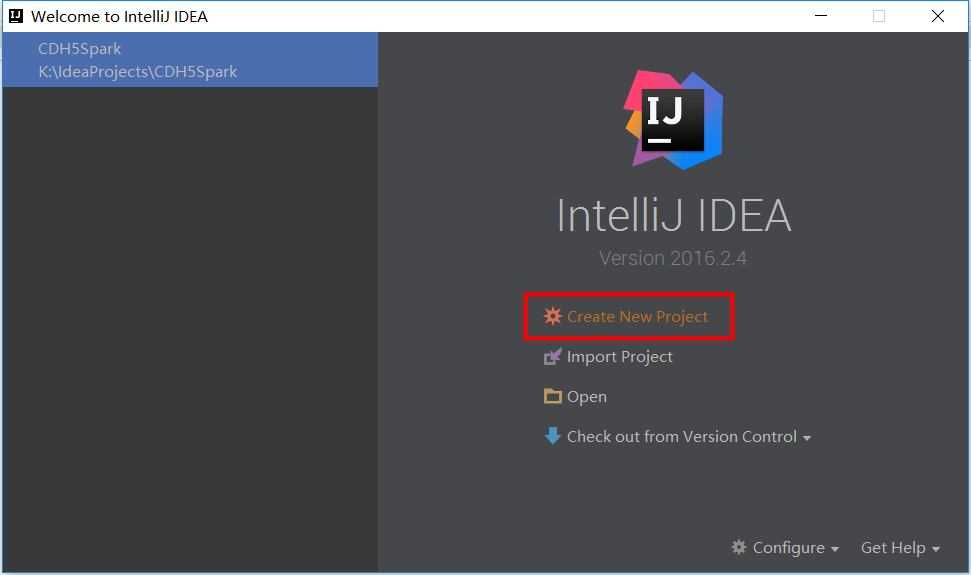
## Scala：2.10.6——不支持2.11，2.10.1有问题

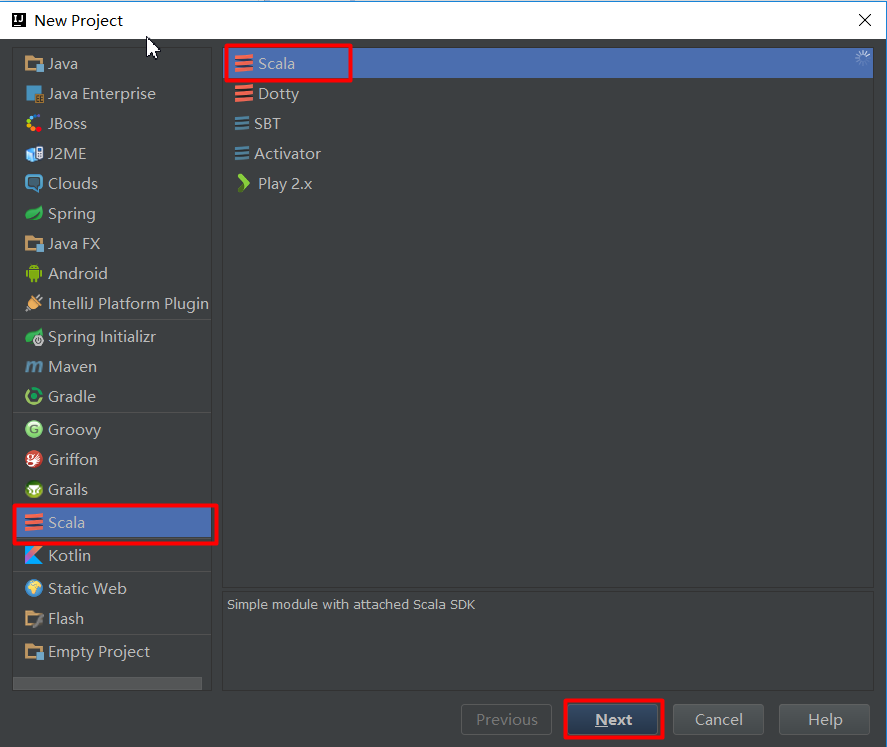
## Hadoop和Spark——安装完添加到环境变量

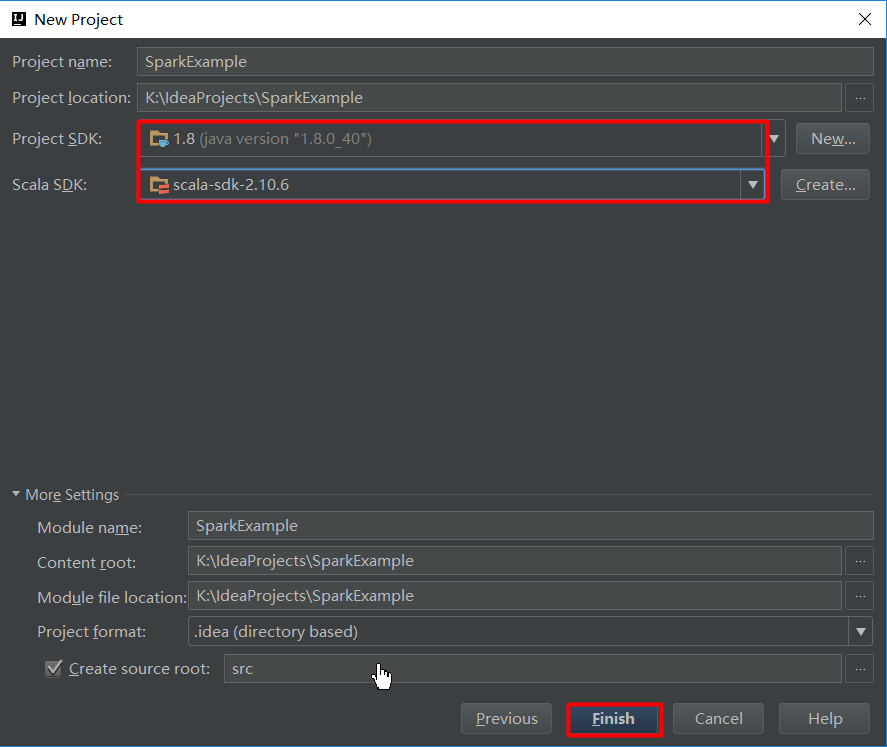
|  |  |
| --- | --- |
| Hadoop | <http://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-2.6.5/hadoop-2.6.5.tar.gz> |
| Spark | http://spark.apache.org/downloads.html |

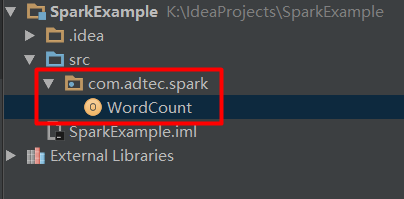


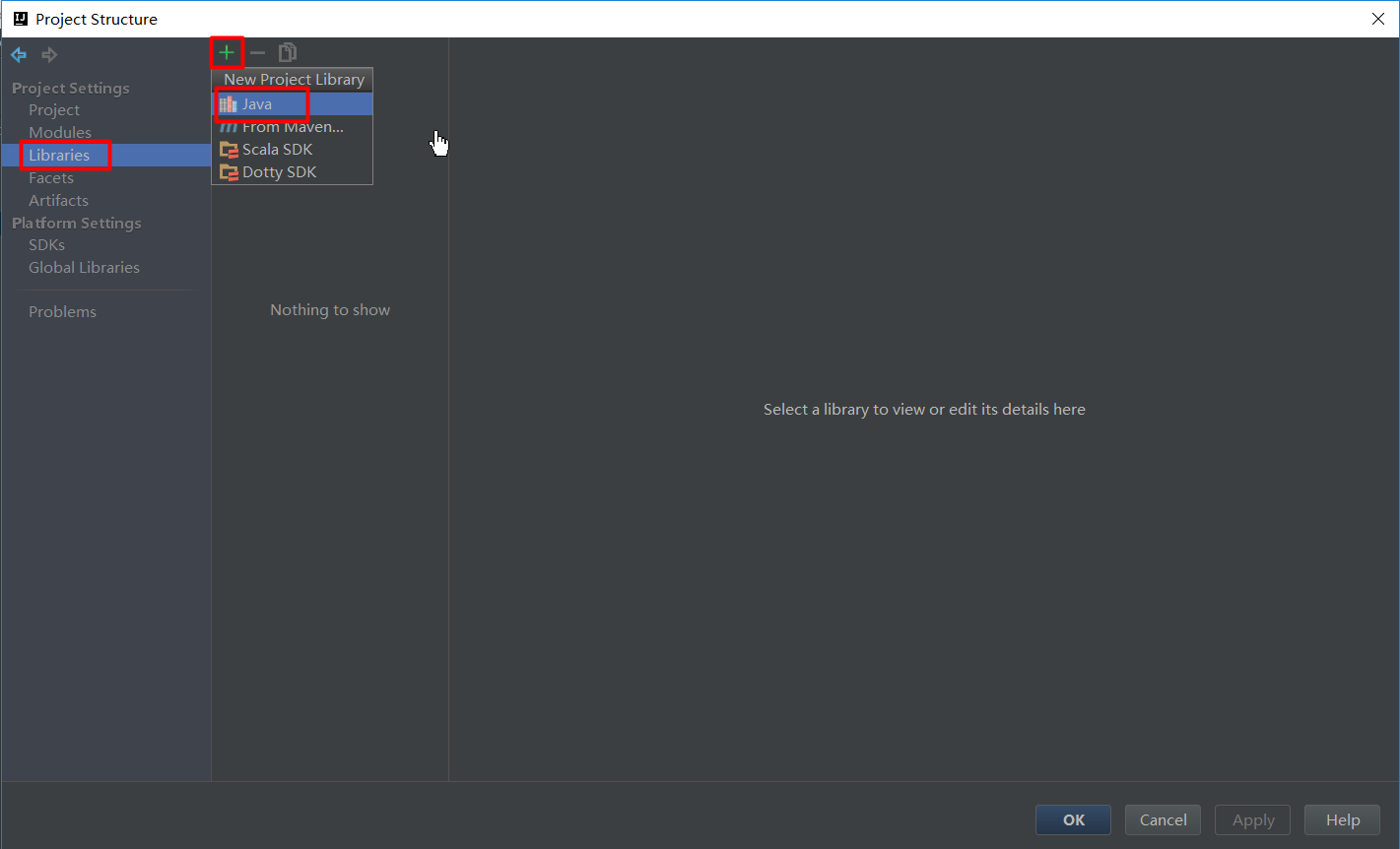
# 创建项目

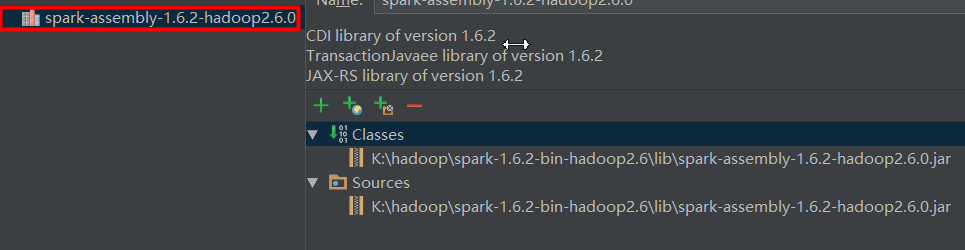












# 编写代码

**def** main(args: Array[String]): Unit = {**val** conf = **new** SparkConf().setAppName("WordCount").setMaster("local")**val** sc = **new** SparkContext(conf)  
  
 **val** lines = sc.textFile("J:\\ abc.txt")  
// 拆分单词，并过滤掉空格，当然还可以继续进行过滤，如去掉标点符号

**val** words = lines.flatMap(\_.split(" ")).filter(\_ != " ")   
  
 // 在单词拆分的基础上对每个单词实例计数为1, 也就是 word => (word, 1)

**val** pairs = words.map((\_, 1))   
  
 // 在每个单词实例计数为1的基础之上统计每个单词在文件中出现的总次数, 即key相同的value相加

**val** wordscount = pairs.reduceByKey(\_ + \_)   
  
 wordscount.collect.foreach(*println*) // 打印结果，

sc.stop() // 释放资源  
}

# 运行

