## Настройка Spark с помощью Scala и запуск в IntelliJ

Я покажу вам, как настроить Apache Spark с помощью Scala и запустить его в IntelliJ.  
IntelliJ IDEA — наиболее часто используемая среда IDE для запуска приложений Spark, написанных на Scala, благодаря хорошему завершению кода Scala. В этой статье я объясню, как настроить и запустить приложение Apache Spark, написанное на Scala, с использованием Apache Maven и IntelliJ IDEA.

1. Установите JDK

Возможно, вы знаете, что Spark был создан на языке Scala, а Scala — это язык JVM, для запуска которого требуется JVM, следовательно, для компиляции и выполнения приложений Spark вам необходимо, чтобы в вашей системе была установлена Java.

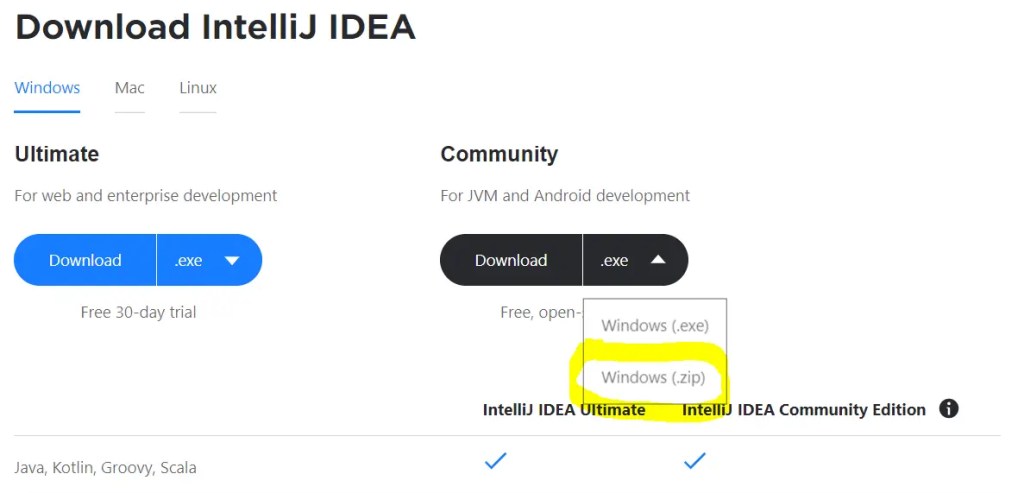
Загрузите и установите Java 8 или более позднюю версию с [Oracle.com](https://www.java.com/en/download/) или OpenJDK.

1. Настройка IntelliJ IDEA для Spark

Большинство инженеров Spark используют IntelliJ IDEA для запуска приложений Spark, написанных на Scala, из-за его хорошей совместимости со Scala, поэтому лучше настроить среду разработки с использованием IntelliJ.

IntelliJ IDEA поставляется с версией для сообщества и окончательной версией. Для запуска приложения Spark, написанного на Scala, нам достаточно версии для сообщества, поэтому скачайте версию [IntelliJ IDEA для сообщества](https://www.jetbrains.com/idea/download/#section=windows).

* 1. Вы можете загрузить установщик Windows (.exe) или сжатый zip-файл (.zip), в зависимости от вашего удобства. Я скачал ZIP-файл.

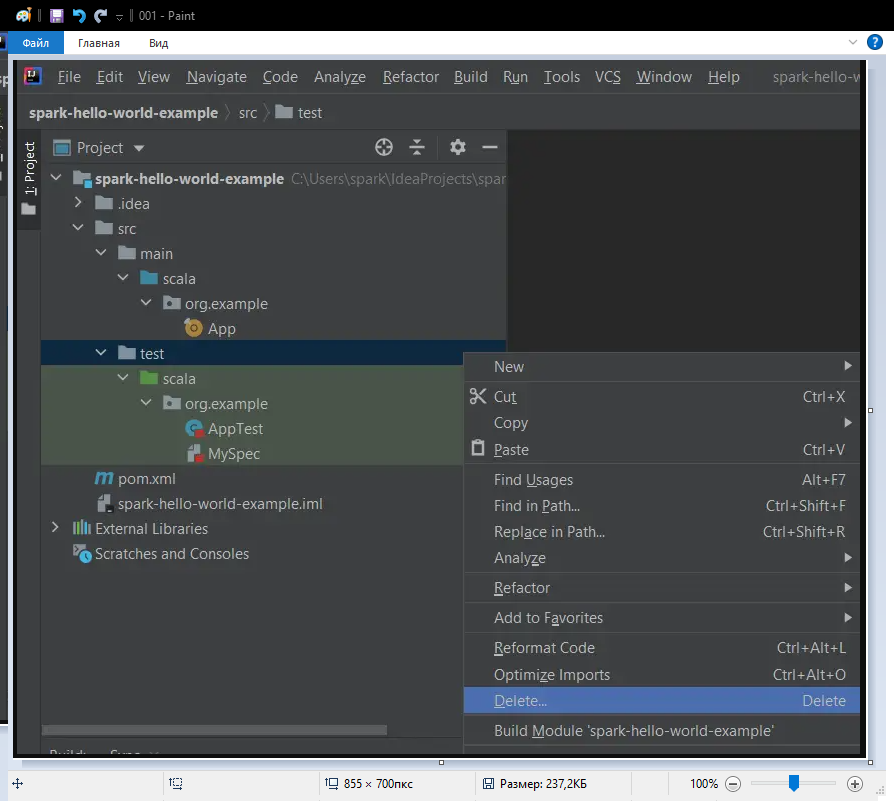


* 1. Теперь давайте распакуем архив с помощью Winzip, 7-Zip или любых других имеющихся у вас архивов. Я использовал 7-Zip для извлечения содержимого в папку.
  2. Переместите извлеченную папку из «Загрузки» в рабочую папку. В моем случае я перемещаю его в c:\apps\.
  3. Запустите IntelliJ IDE, запустив idea64.exe из C:\apps\ideaIC-2020.2.1.win\bin\idea64.exe.

1. Создайте проект Scala в IntelliJ

После запуска IntelliJ IDEA IDE вы увидите экран приветствия с различными опциями.

* 1. Выберите «Новый проект», чтобы открыть окно нового проекта.



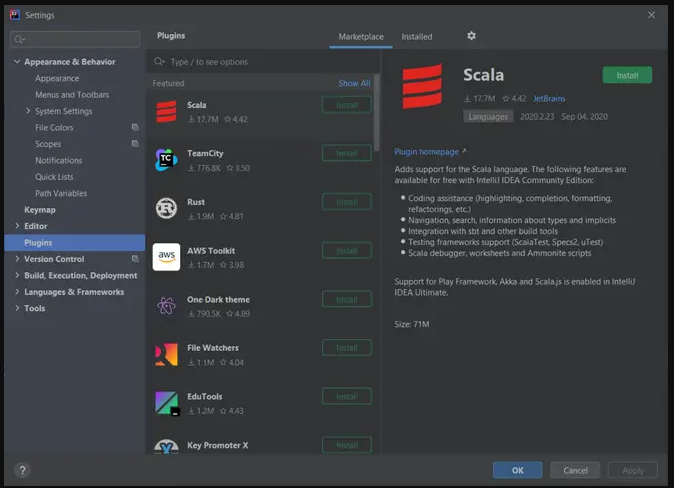
* 1. Выберите Maven на левой панели.
  2. Установите флажок «Создать из архетипа».( **Create from archetype**)
  3. Выберите org.scala-tools.archetypes:scala-archetypes-simple.  
     Архетип — это своего рода шаблон, который создает правильную структуру каталогов и загружает необходимые зависимости по умолчанию. Поскольку мы выбрали архетипы Scala, он загружает все зависимости Scala и позволяет IntelliJ писать код Scala.
  4. В следующем окне введите название проекта. Я называю свой проект «**spark-hello-world-example**».
  5. На следующем экране просмотрите параметры идентификатора артефакта и идентификатора группы.
  6. Выберите Готово.(**Finish**.)

При этом создается проект на IntelliJ, и на левой панели проекта отображается следующая структура проекта. (вам необходимо разворачивать проект, пока не увидите файл App.scala.

1. Установите плагин Scala

Теперь перейдите к:

* Откройте **File** > **Settings**  (или используйте клавиши быстрого доступа Ctrl + Alt + s). В **macOS** используйте **IntellijIDEA -> Preferences**
* Выберите параметр **Plugins** на левой панели. Вы попадете на панель функций.
* Нажмите **Install**, чтобы установить плагин Scala.



* После установки плагина перезапустите IntelliJ IDE.

1. Настройка Scala SDK

https://sparkbyexamples.com/spark/spark-setup-run-with-scala-intellij/