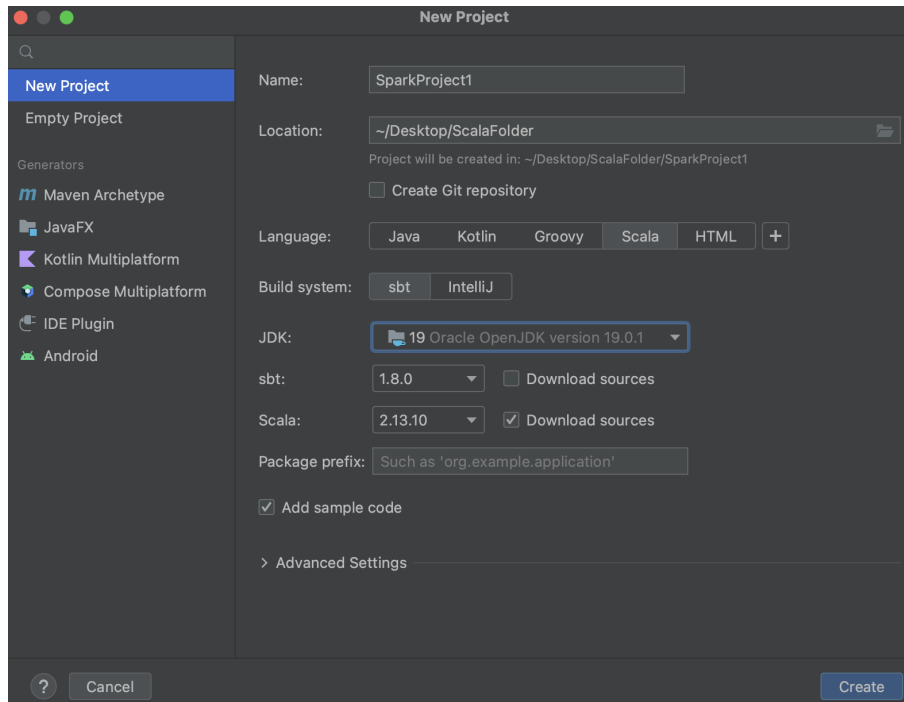# SBT VERSION

Project Configuration for Spark Scala

Step 1 - Create a new project

Step 2 - Configure a sbt file

```
ThisBuild / version := "0.1.0-SNAPSHOT"

ThisBuild / scalaVersion := "2.13.2"

lazy val root = (project in file("."))
 .settings(
   name := "SparkProject1"
 )

val sparkVersion = "3.3.1"
// Note the dependencies are provided
libraryDependencies += "org.apache.spark" %% "spark-core" % sparkVersion
libraryDependencies += "org.apache.spark" %% "spark-sql" % sparkVersion
```

Step 3 - Create a scala object to test the first Spark App

```scala
import org.apache.spark.sql.SparkSession

object Main {
 def main(args: Array[String]): Unit = {
   println("Hello world!")

   val spark =
SparkSession.builder.master("local").appName("SparkByExample").getOrCreate(
)

   println("First SparkContext:")
   println("APP Name :" + spark.sparkContext.appName)


 }
}
```
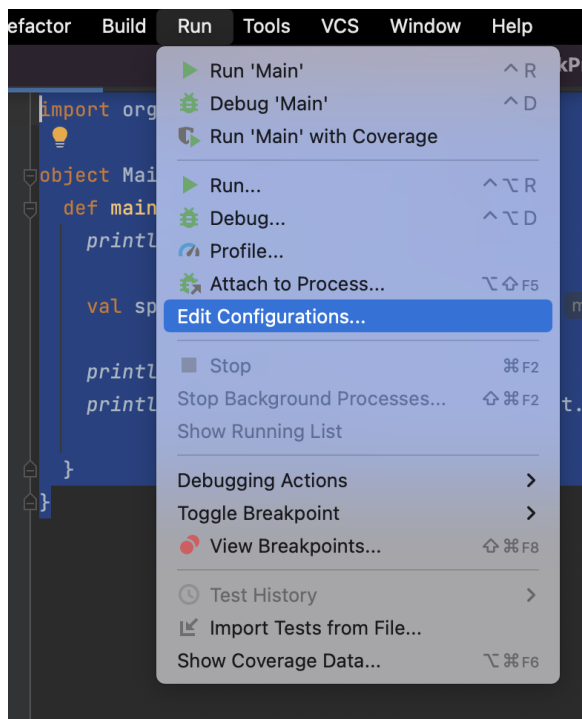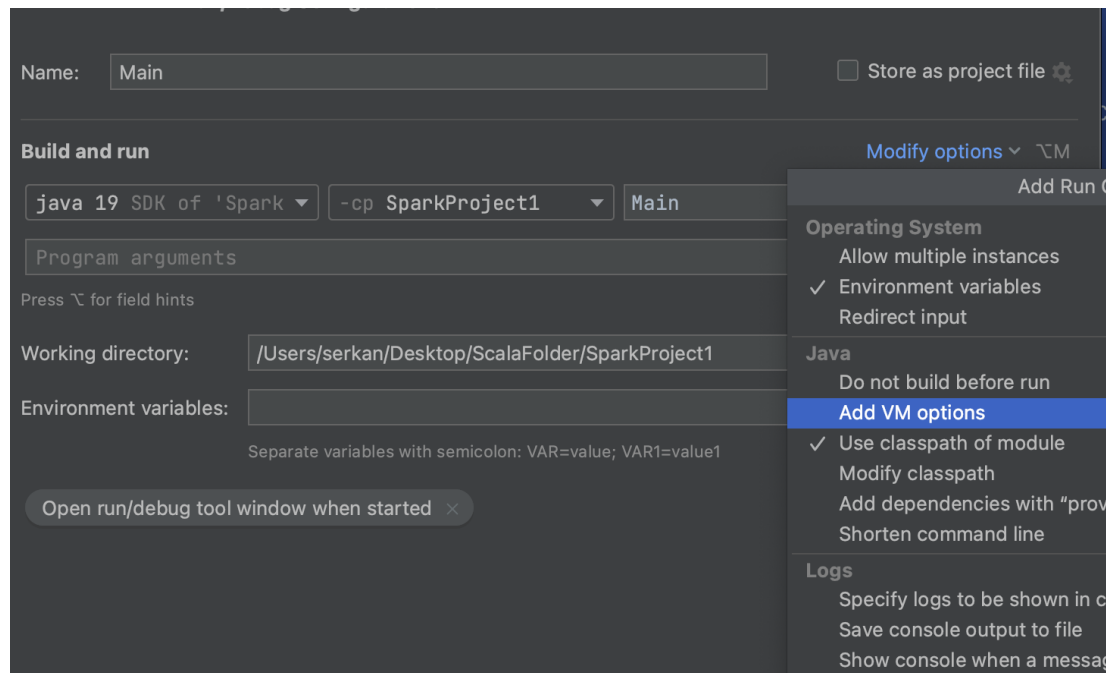
Step 4 - Change the run configuration
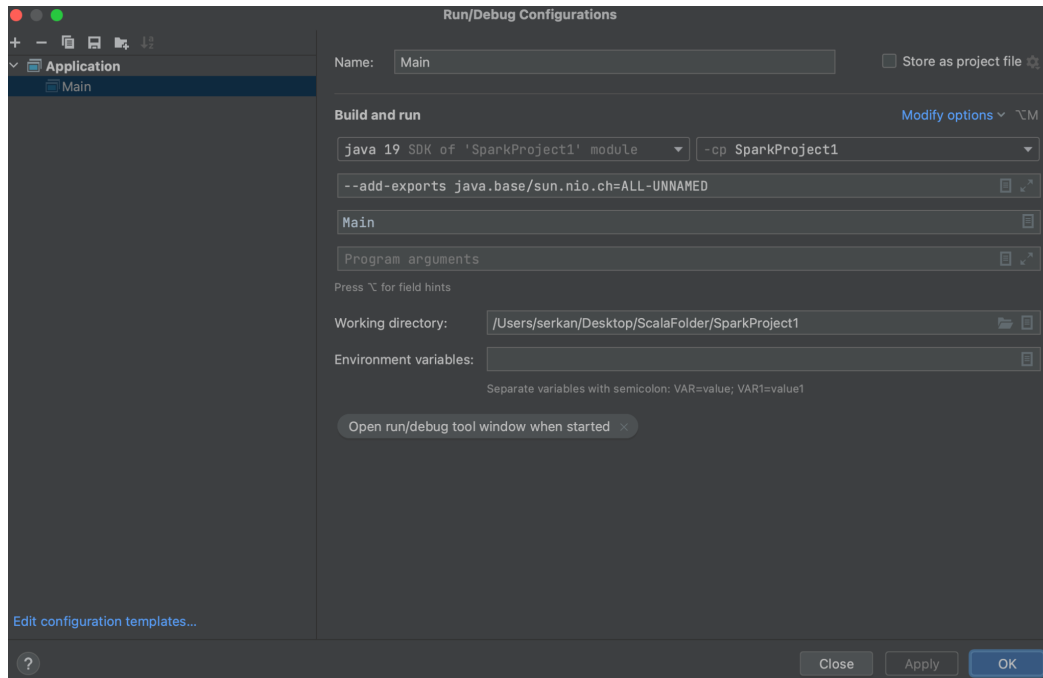
Click to the Edit Configuration

Select the Add VM Option

Add the following line

```
--add-exports java.base/sun.nio.ch=ALL-UNNAMED
```

Run the sample code

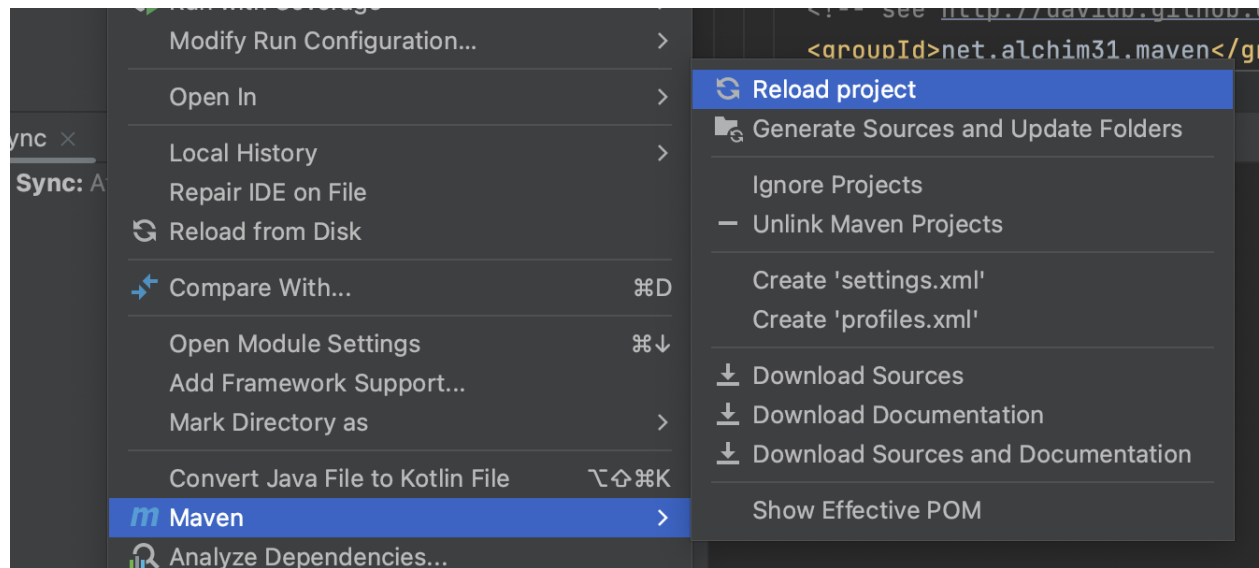```
First SparkContext:
APP Name :SparkByExample
```

# MAVEN VERSION

**Step 1** - Add respective spark core version

```xml
<dependency>
 <groupId>org.apache.spark</groupId>
 <artifactId>spark-core_2.12</artifactId>
 <version>3.0.0</version>
</dependency>
```

```xml
<dependency>
 <groupId>org.apache.spark</groupId>
 <artifactId>spark-sql_2.12</artifactId>
 <version>3.0.0</version>
</dependency>
```

Step 2 - Reload the project via Maven

Step 3 - Sample code

```scala
package org.example

import org.apache.spark.sql.SparkSession


object SparkHello {

 def main(args: Array[String]): Unit = {

   val logFile = "test" // Should be some file on your system
   val spark = SparkSession.builder.master("local").appName("Simple
Application").getOrCreate()
   val logData = spark.read.textFile(logFile).cache()
   val numAs = logData.filter(line => line.contains("a")).count()
   val numBs = logData.filter(line => line.contains("b")).count()
   println(s"Lines with a: $numAs, Lines with b: $numBs")
   spark.stop()

 }

}
```