Mac Users will need to follow the below instructions:

<https://sparkbyexamples.com/spark/install-apache-spark-on-mac/>

Windows Installation:

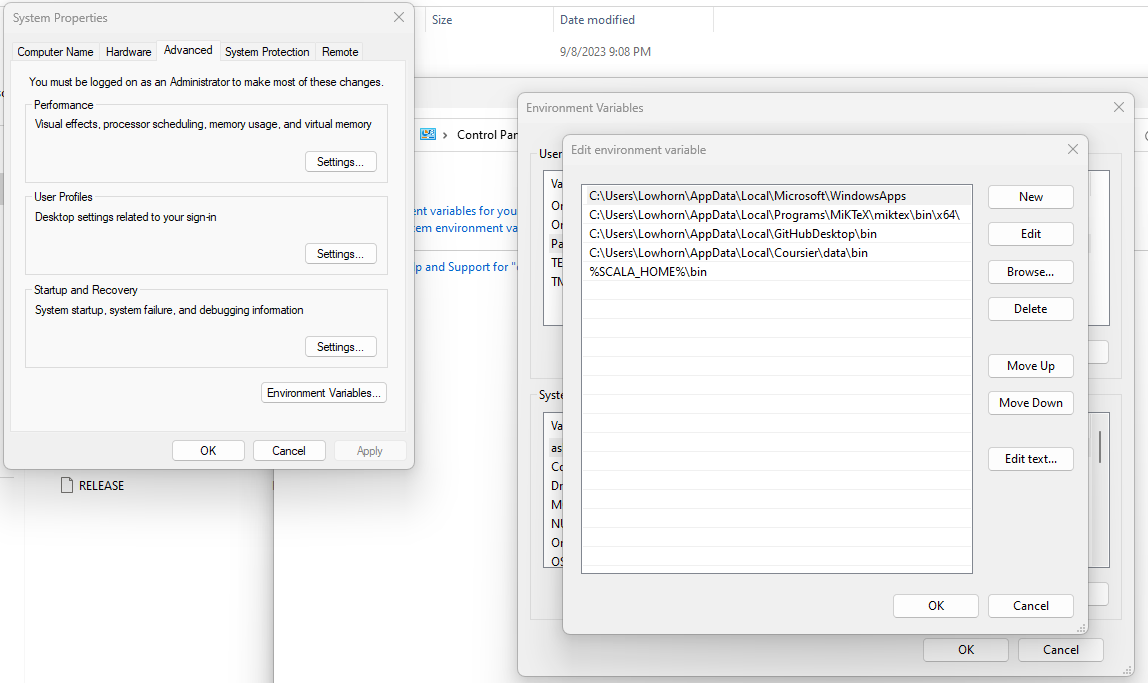
1. [Install Java 7 or later](http://www.oracle.com/technetwork/java/javase/downloads/index.html). To test java installation is complete, open command prompt type java and hit enter. If you receive a message 'Java' is not recognized as an internal or external command. You need to configure your environment variables, JAVA\_HOME and PATH to point to the path of jdk.
2. [Download and install Scala](http://www.scala-lang.org/download/).
   1. A screenshot of a computer

      Description automatically generated

Set SCALA\_HOME in Control Panel\System and Security\System . On the screen shot notice the relative path to Scala <\AppData\Local\Coursier\data\bin.> A screenshot of a computer

Description automatically generated

Go to "Adv System settings" and add %SCALA\_HOME%\bin in PATH variable in environment variables.



1. Install Python 2.6 or later from [Python Download link](https://www.python.org/downloads/windows/).
   1. If you have Anaconda and Python already installed you can skip this link.
   2. I highly recommend 3.9 or later.
2. [Download SBT](http://www.scala-sbt.org/download.html). Install it and set SBT\_HOME as an environment variable with value as <<SBT PATH>>.
3. Install Hadoop
   1. Download Hadoop on its own (if coming from Spark)
   2. Download matching Hadoop version files for Windows from https://github.com/cdarlint/winutils.
   3. Extract the files (e.g. C:\hadoop). Make sure the directory structure is similar to this C:\hadoop\bin\winutils.exe
   4. Set environment variable HADOOP\_HOME to C:\hadoop
   5. Restart your computer
   6. \*\*\*You may need to copy the Hadoop.dll file to C:\Windows\System32
   7. Add Hadoop to Path env variable: %HADOOP\_HOME%\binA screenshot of a computer

      Description automatically generated
4. A screenshot of a computer

   Description automatically generated
5. We will be using a pre-built Spark package, so choose a Spark pre-built package for [Hadoop Spark download](http://spark.apache.org/downloads.html). Download and extract it.

Set SPARK\_HOME and add %SPARK\_HOME%\bin in PATH variable in environment variables.

1. Run command: spark-shell
2. Open http://localhost:4040/ in a browser to see the SparkContext web UI.