Installing Spark on Windows 10.
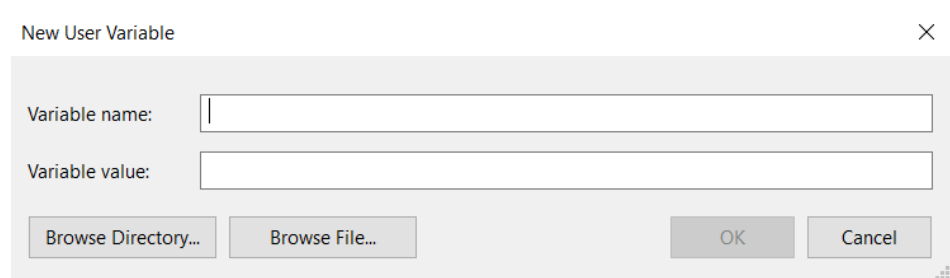
1. **Install Scala:**
   a. Download Scala from the link: http://downloads.lightbend.com/scala/2.11.8/scala-2.11.8.msi
   b. Install Scala (I installed it to C:\Program Files (x86)\scala)
   c. Set environmental variables:
      i. Find the Environmental button in System Properties
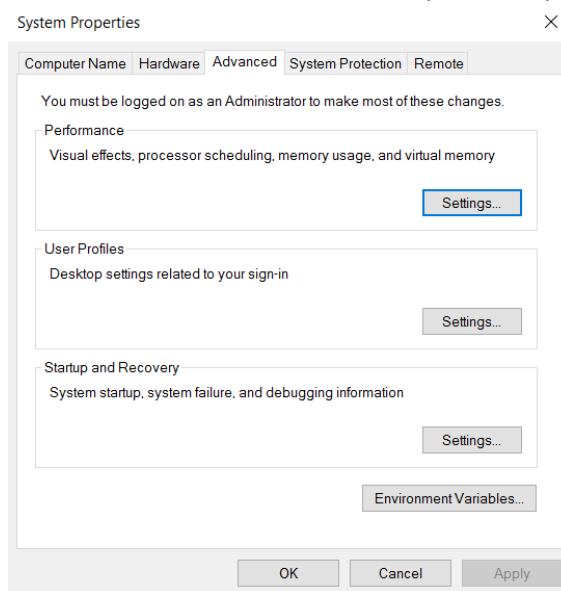
      ii. User variable:

      - Click New. Below window will pop out

      - Variable: SCALA_HOME

      - Value: C:\Program Files (x86)\scala

      System variable:

      - Click PATH

      - Then click EDIT

      - Then click NEW

      - Paste C:\Program Files (x86)\scala\bin

Chia-Lun Chen

2. **Install Java 8**:
   a. Download Java 8 from the link.
      http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html
   b. Install Java (I installed it to C:\Program Files (x86)\Java)
   c. Set environmental variables:
      User variable:

      - Variable: JAVA_HOME

      - Value: C:\Program Files\Java\jdk1.8.0_101

      System variable:

      - Click PATH

      - Then click EDIT

      - Then click NEW

      - Paste C:\Program Files\Java\jdk1.8.0_101

3. **Install Eclipse:**
   a. Download and install Eclipse. https://eclipse.org/downloads/
   b. I installed it to C drive (C:\Users\endmy\eclipse\java-neon\eclipse)
   c. Set environmental variables:
      User variable:

      - Variable: ECLIPSE_HOME

      - Value: C:\Users\endmy\eclipse\java-neon\eclipse

      System variable:

      - Eclipse Neon does not have bin. If your Eclipse is a version that has bin folder. Paste the directory under PATH – EDIT - NEW (e.g. C:\eclipse \bin)
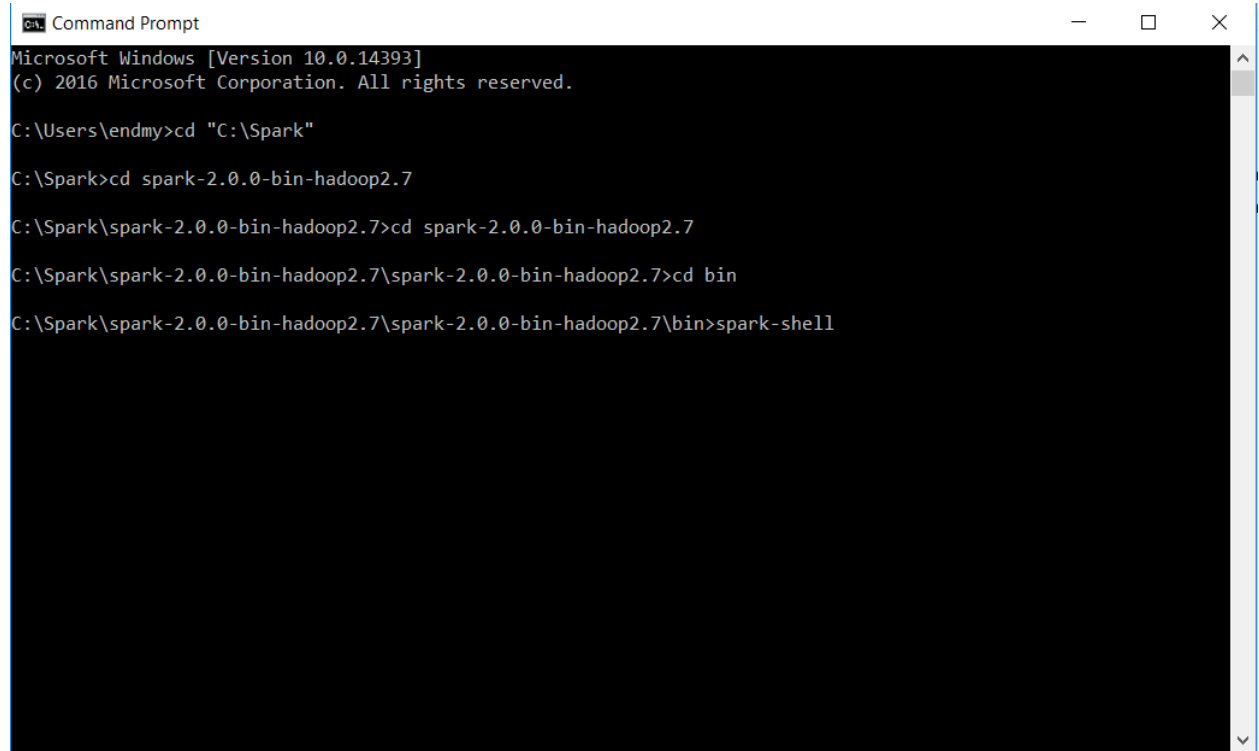
4. **Install Spark 2.0.0**.
   a. Download it from the following link: http://spark.apache.org/downloads.html
   b. Unzip the tgz file (I unzipped it to C drive)
   c. Unzip the tar file (The final dir for mine looks like this: C:\Spark\spark-2.0.0-bin-hadoop2.7\spark-2.0.0-bin-hadoop2.7)
   d. Set environmental variables:
      User variable:

      - Variable: SPARK_HOME

      - Value: C:\Spark\spark-2.0.0-bin-hadoop2.7\spark-2.0.0-bin-hadoop2.7

      System variable:

      - Click PATH

Chia-Lun Chen

• Then click EDIT

• Then click NEW

• Paste C:\Spark\spark-2.0.0-bin-hadoop2.7\spark-2.0.0-bin-hadoop2.7\bin

5. **Download Windows Utilities:**
   a. Download it from the link (winutils.exe).
      https://github.com/steveloughran/winutils/tree/master/hadoop-2.6.0/bin
   b. And paste it in C:\Spark\spark-2.0.0-bin-hadoop2.7\spark-2.0.0-bin-hadoop2.7\bin

6. **Execute Spark on cmd:**

Result:

Chia-Lun Chen

Chia-Lun Chen