**Steps to configure the Apache spark on window environment**

1. **Download spark from the URL :** [**https://spark.apache.org/downloads.html**](https://spark.apache.org/downloads.html)

**(a). Select the version** [spark-3.3.1-bin-hadoop3.tgz](https://www.apache.org/dyn/closer.lua/spark/spark-3.3.1/spark-3.3.1-bin-hadoop3.tgz) or which version you want to install

even this is the latest version of spark.

**(b).** unzip the spark file and put this folder into your desired location like C: or D Drive.

**2**. Download the **winutils.exe** files for **spark 64 bit.**

**(a)** Create a folder with name **Hadoop** inside **bin** this folder and put this file in this folder-

**i.e hadoop/bin/winutils.exe**

**3. install java in your system Java 8 or above version into your system-**

**4. After completing all these steps set the environment variables in window-**

**(A)** To setting up environment variable for Spark use: SPARK\_HOME- Location/path of your

Folder – **[SPARK\_HOME** - **C:\spark-3.2.3-bin-hadoop3.2-scala2.13 ]**

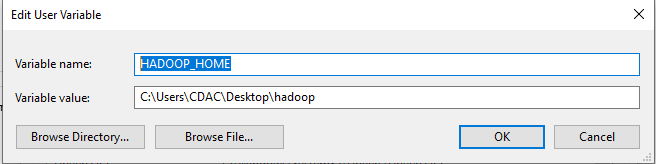
**(B)** To setting up environment variable for java use : JAVA\_HOME- Location/path of your

Folder [**JAVA\_HOME**- C:\Program Files\Java\jdk-11.0.17]

**( C)** To use the pyspark **add one more path in environment variable-**

**PYSPARK\_HOME (variable name) - C:\spark-3.2.3-bin-hadoop3.2-scala2.13**  (path value)

**(D) To use the winutils.exe** file: set the path in environment variables-



To come out from spark-shell use Ctrl+D or quit

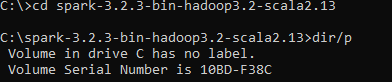
To come out from pyspark use Ctrl+z or exit()

**After completing all the steps spark has been configured on your system-**

**Now open command prompt and type the following commands-**

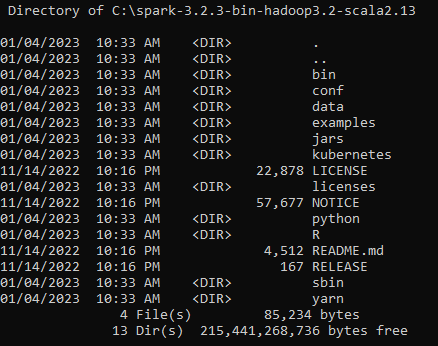
**1.cd...**

**2. cd** spark-3.2.3-bin-hadoop3.2-scala2.13

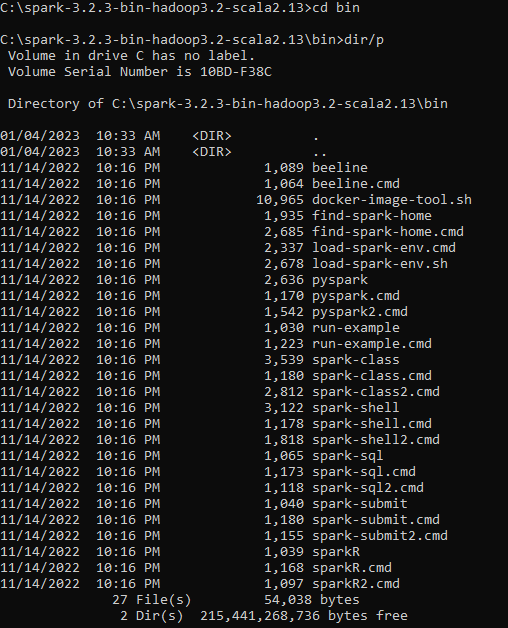


3. C:\spark-3.2.3-bin-hadoop3.2-scala2.13>**dir/p.**

Will show all the directories available in your spark package.



**4. cd bin**



**5. Start your spark by using spark-shell command**

