**INSTALLATION AND SETTING UP APACHE SPARK IN WINDOWS**

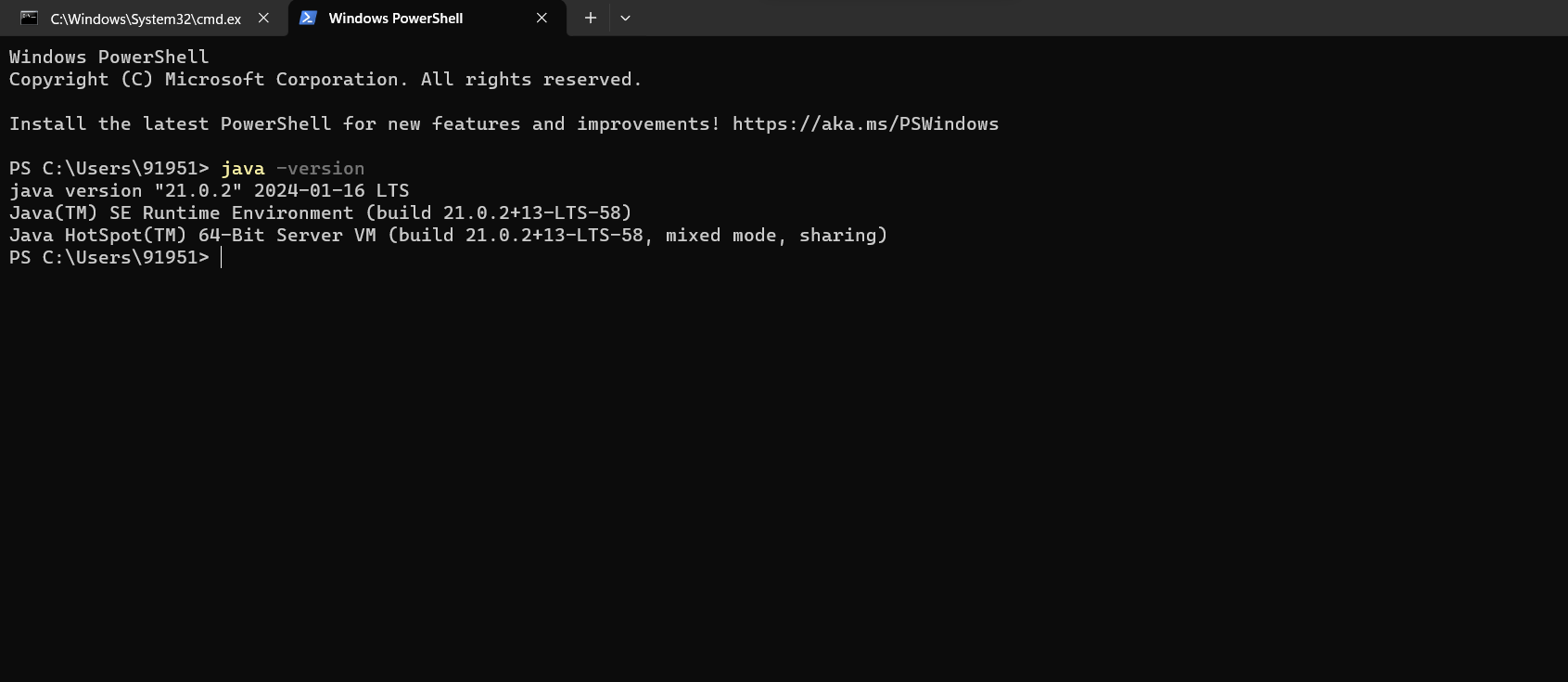
Here are the steps to Install Apache Spark

**Apache Spark:**

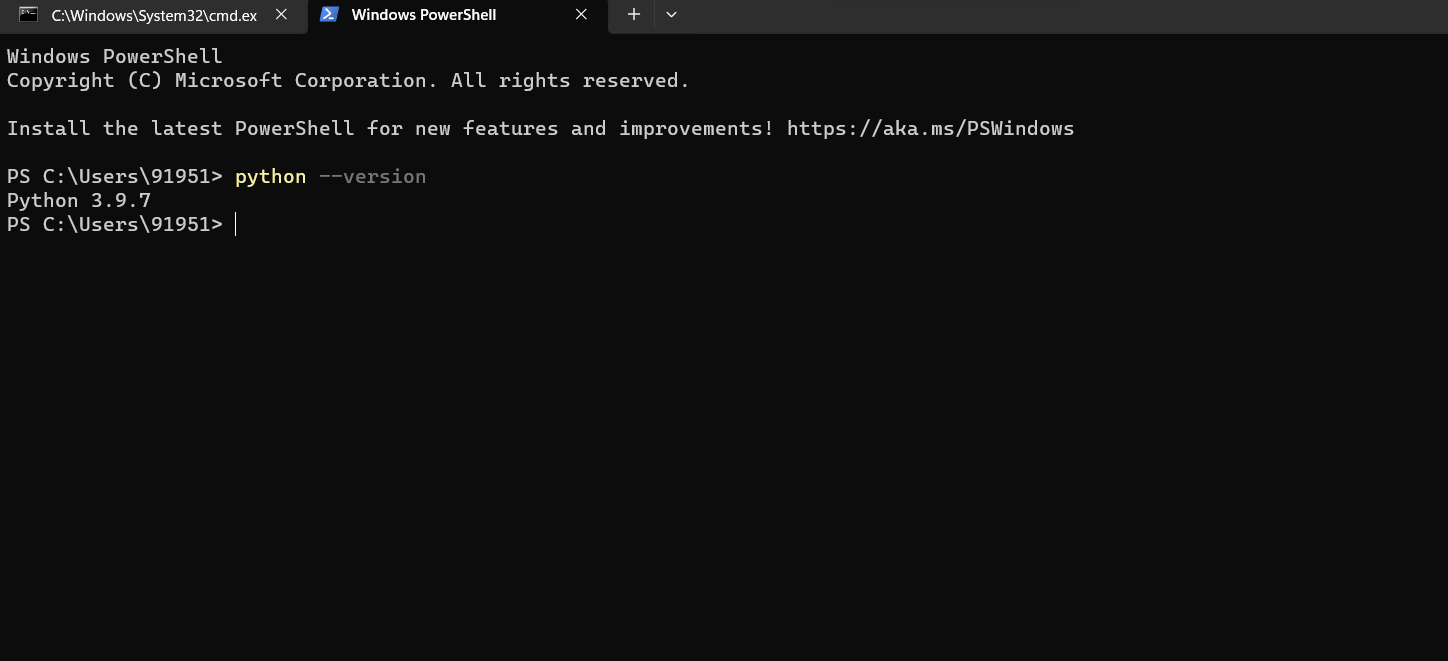
Apache Spark is an open-source, distributed processing system used for big data workloads. It utilizes in-memory caching, and optimized query execution for fast analytic queries against data of any size.

**Step 1:** Installing JAVA in the System

* Here we need have installed java jdk 21 in our system. We have downloaded it from browser and installed latest version.
* Once we have installed java, for checking the version we need to type java -version in command prompt so that it displays the version which we have installed
* Below is the picture that we have successfully installed the java latest version



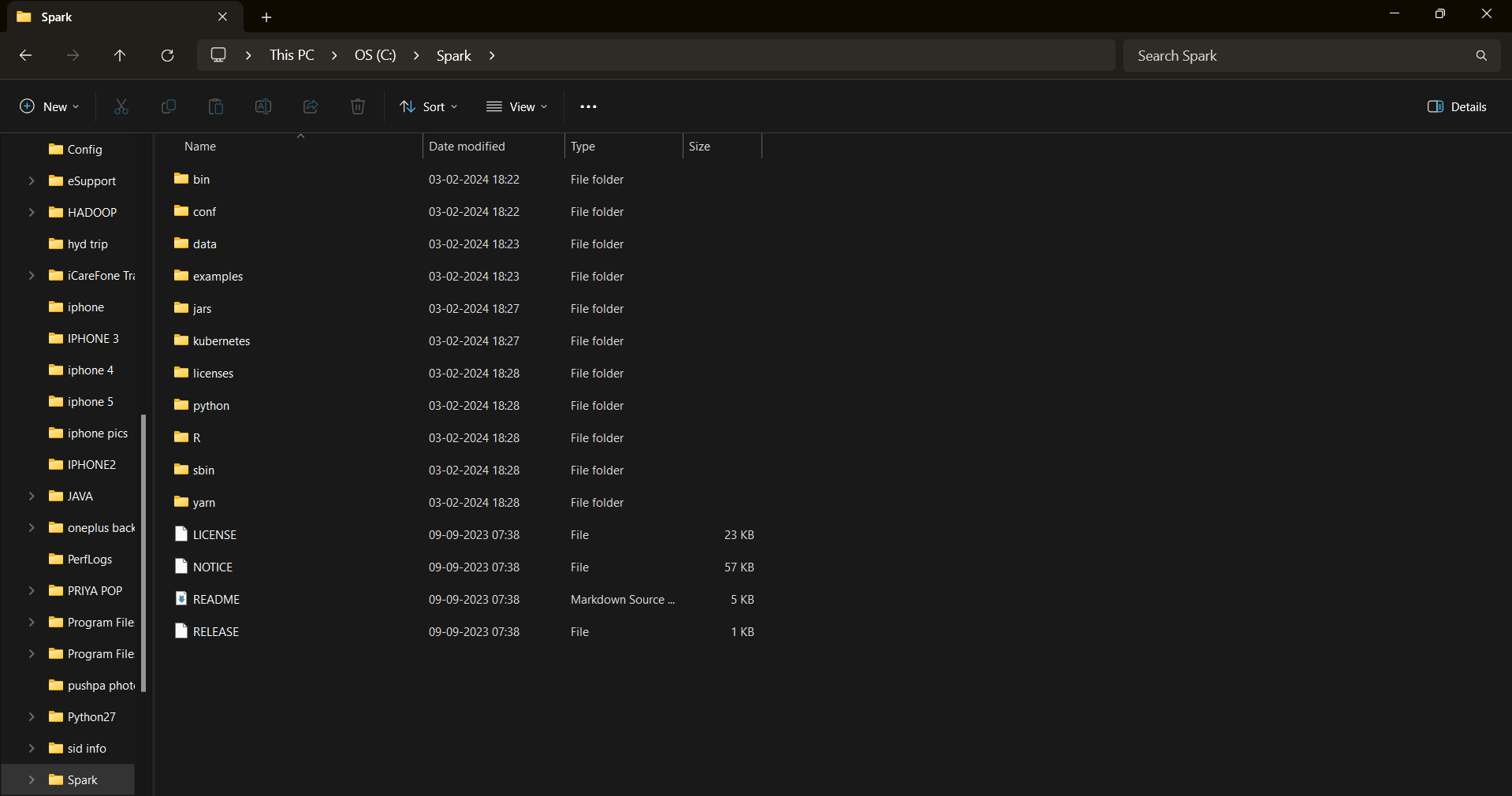
**Step 2:** Installing Python in our system

* Here we have already downloaded the python and installed it in our system. If you haven’t downloaded use the latest version and download it,
* After installing the python, we can check the version by typing python --version in the command prompt and which displays the version we have installed
* Below is the picture that we have successfully installed python 

**STEP 3:**

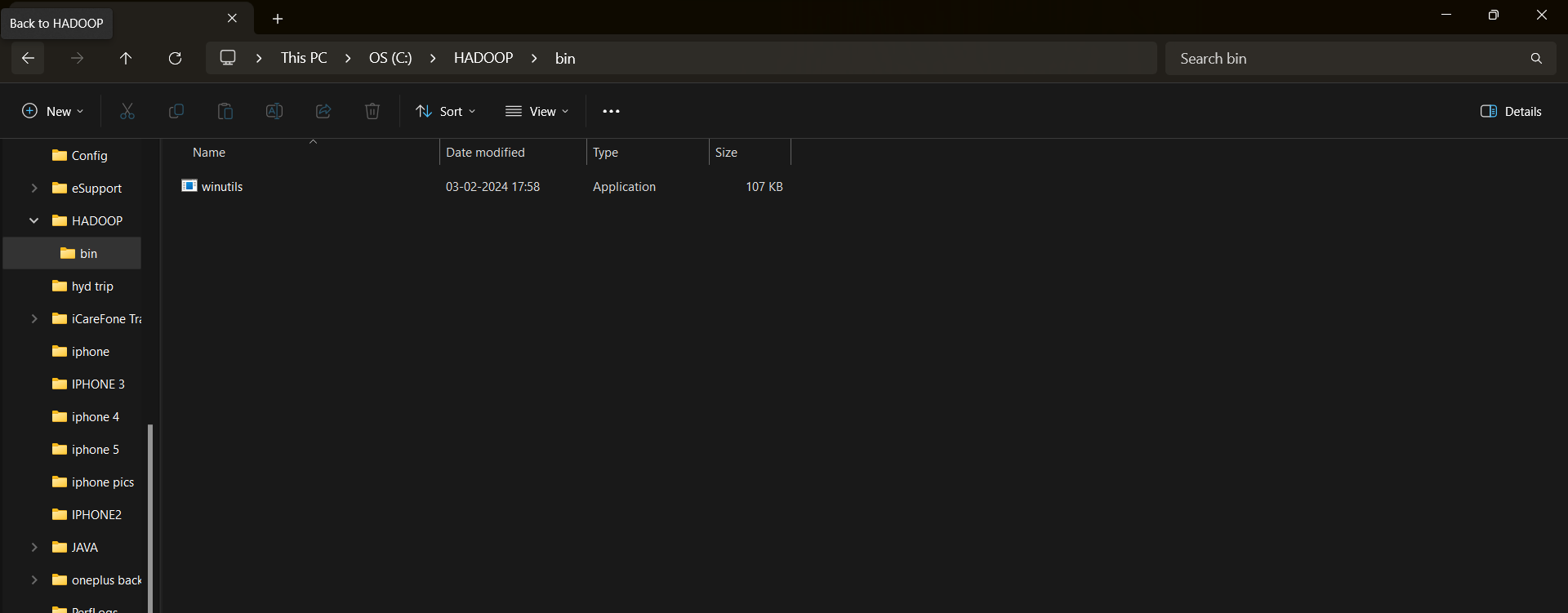
Now we need to download and install apache spark in our system

* Download Apache Spark from Official Website and download the latest version
* Download It with package for Hadoop.
* Extract the files in “C” Folder preferably.



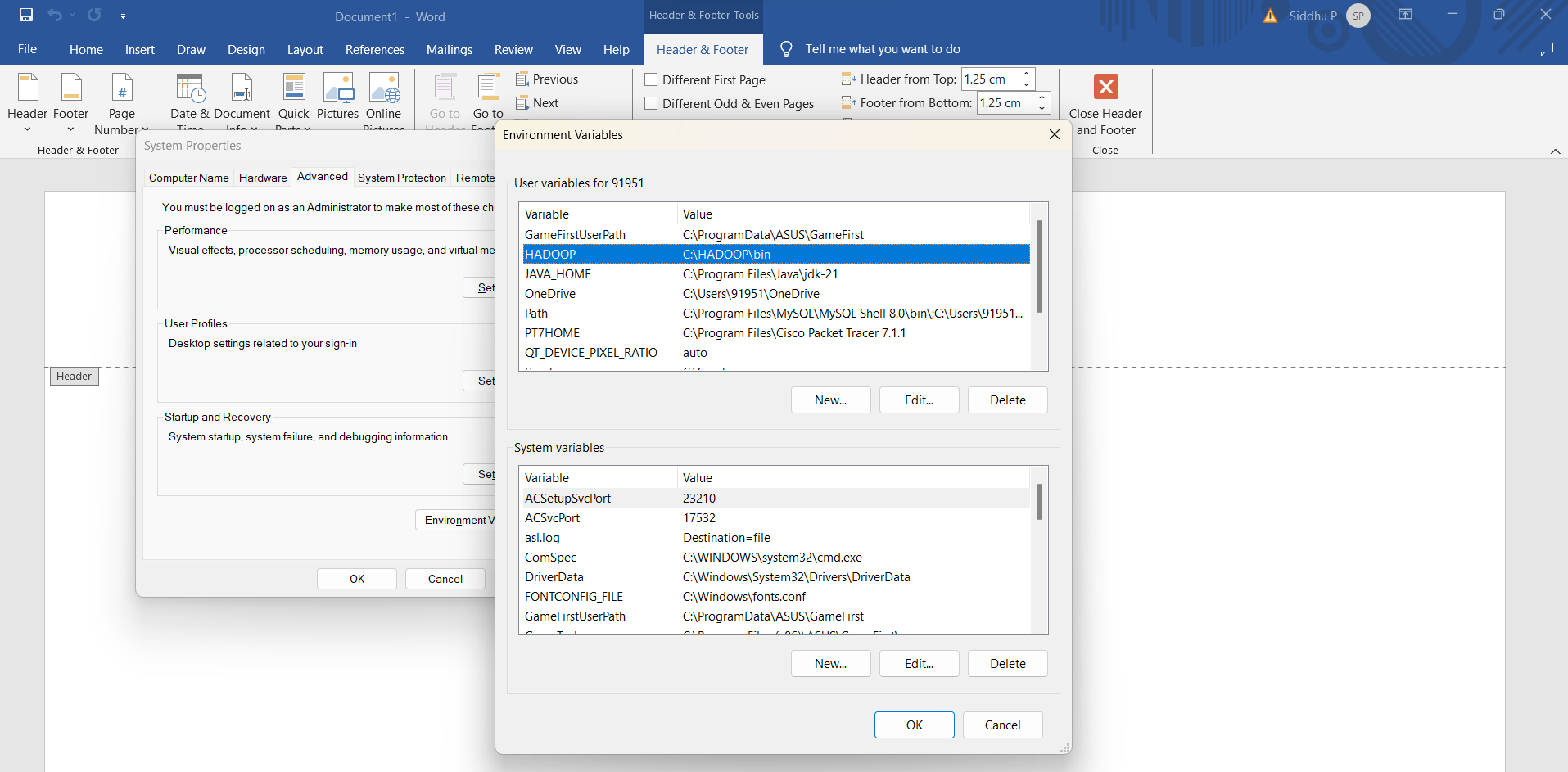
**STEP 4:**

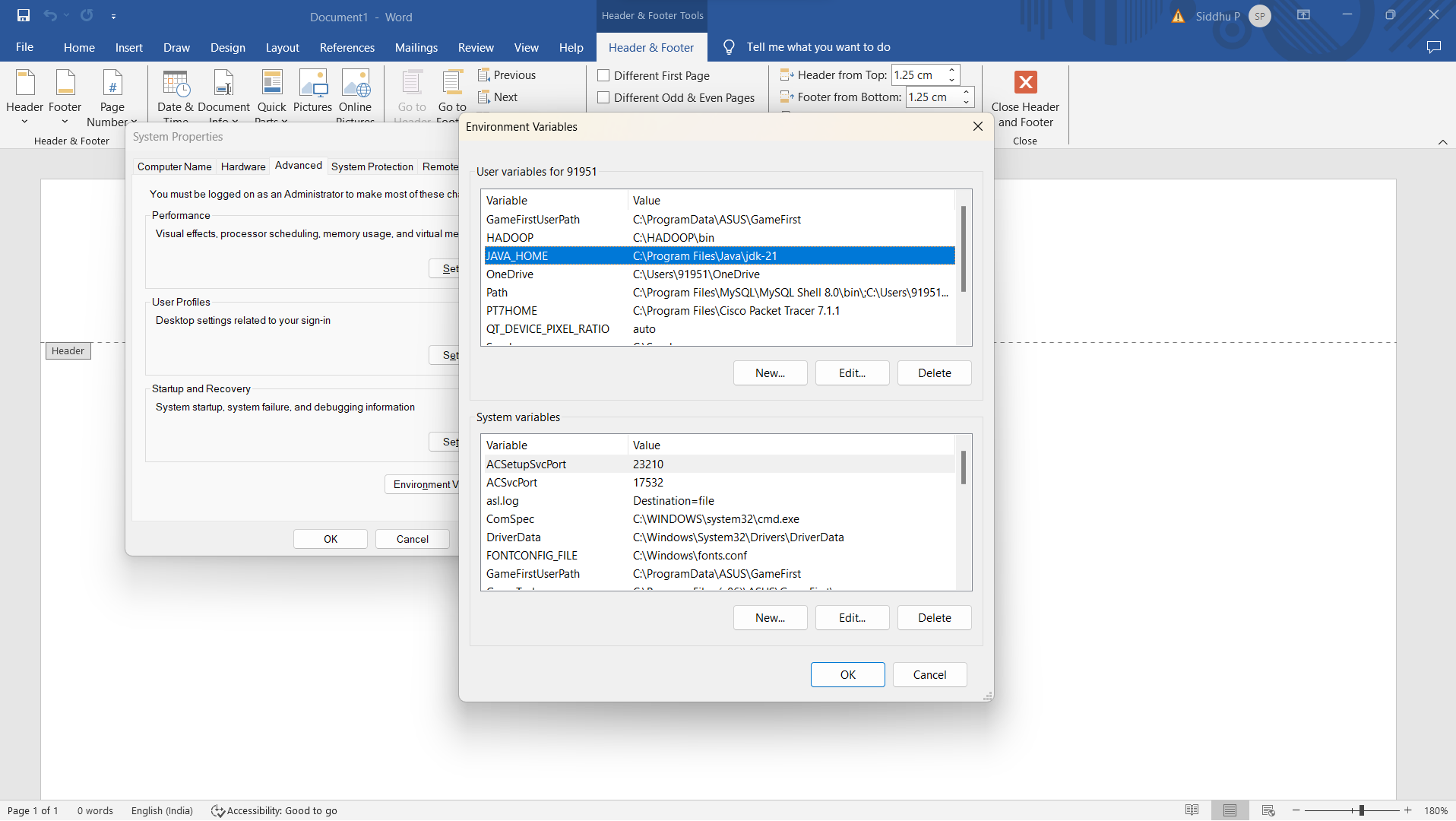
* Create Folder HADOOP in “C” Folder and download the Hadoop from the apache webpage
* Extract all the files In “C” Folder Preferably

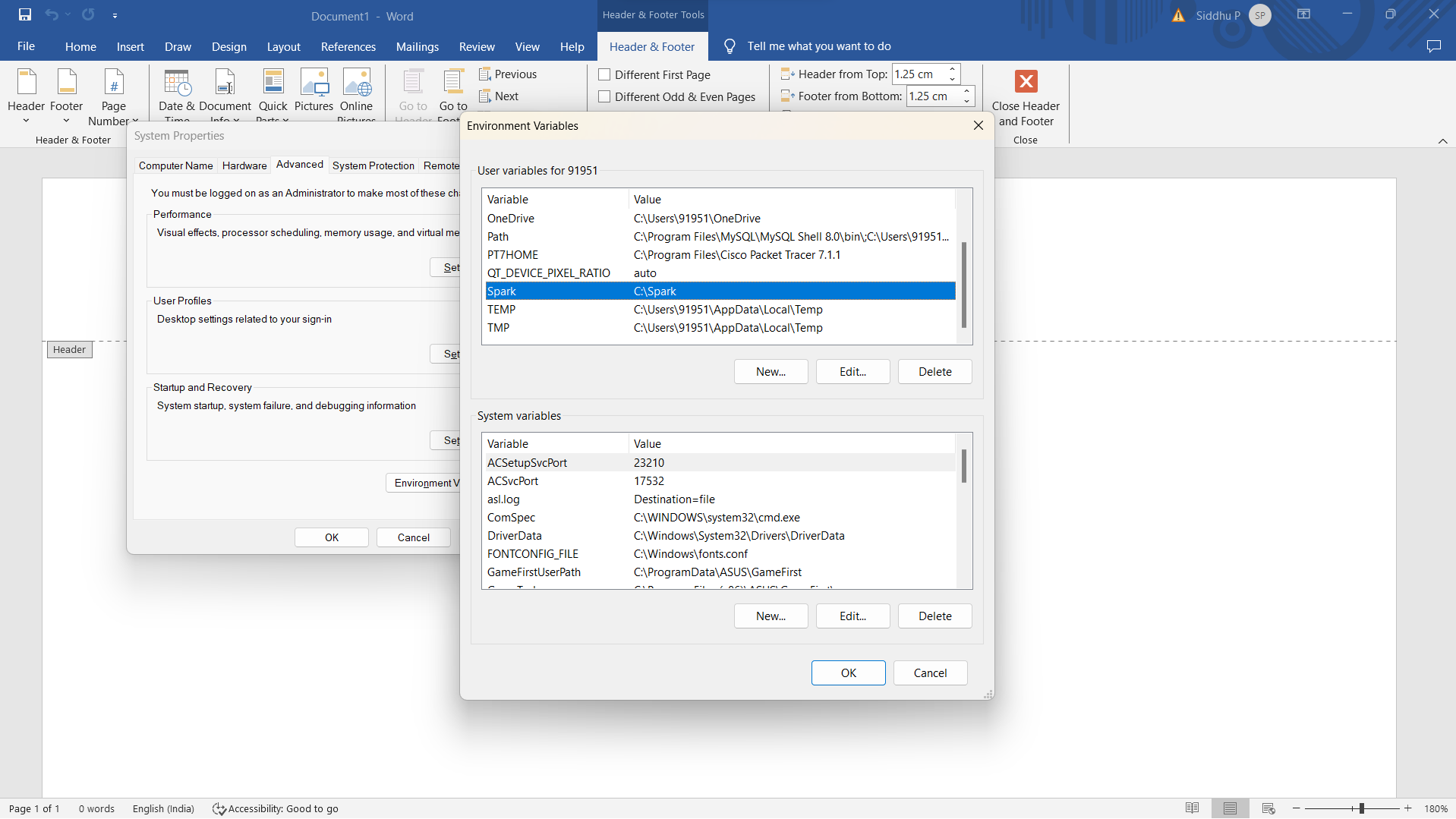


**STEP 5:**

Now we need to Set Environment Variables For JAVA, HADOOP AND SPARK as shown below





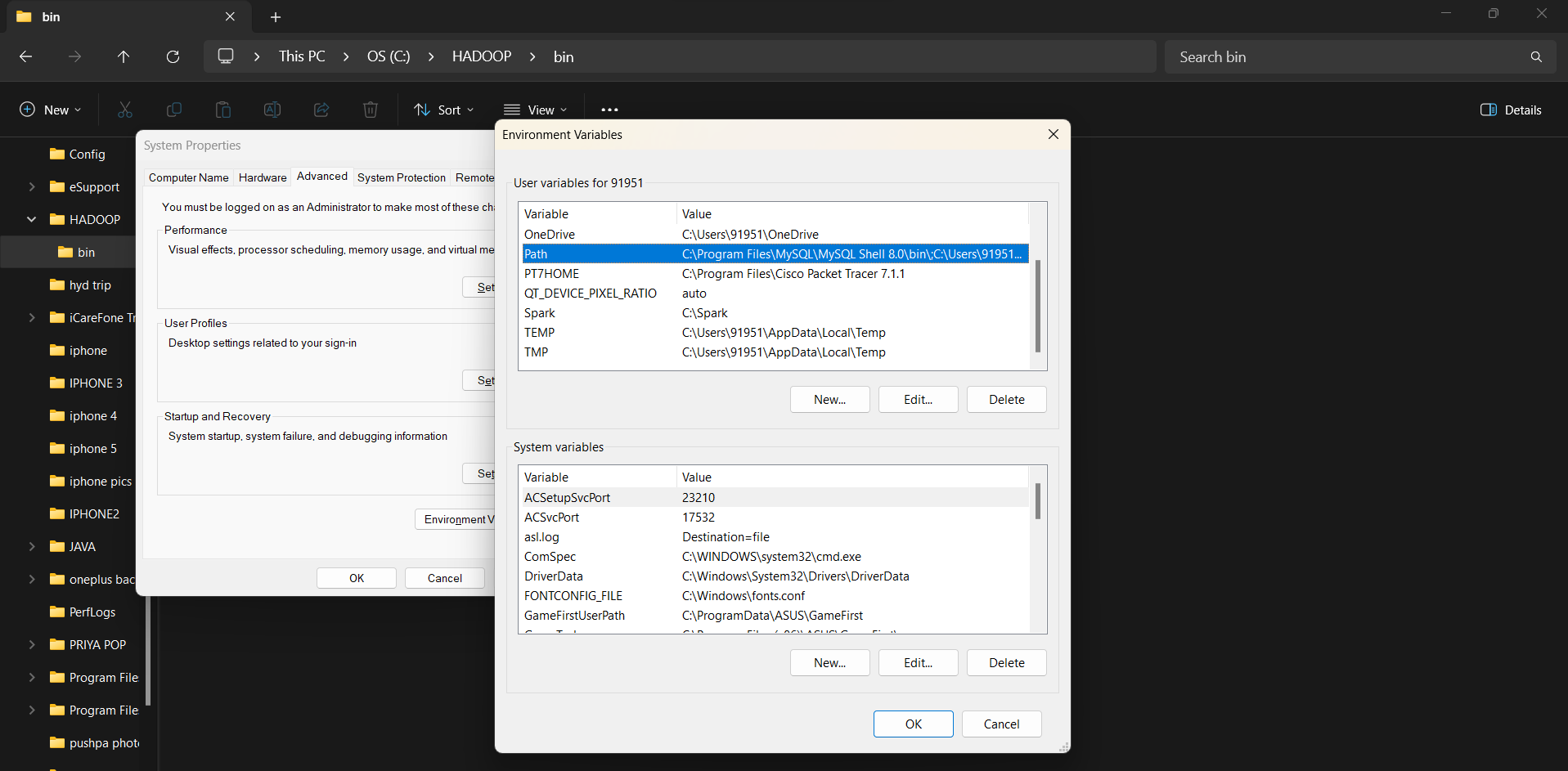


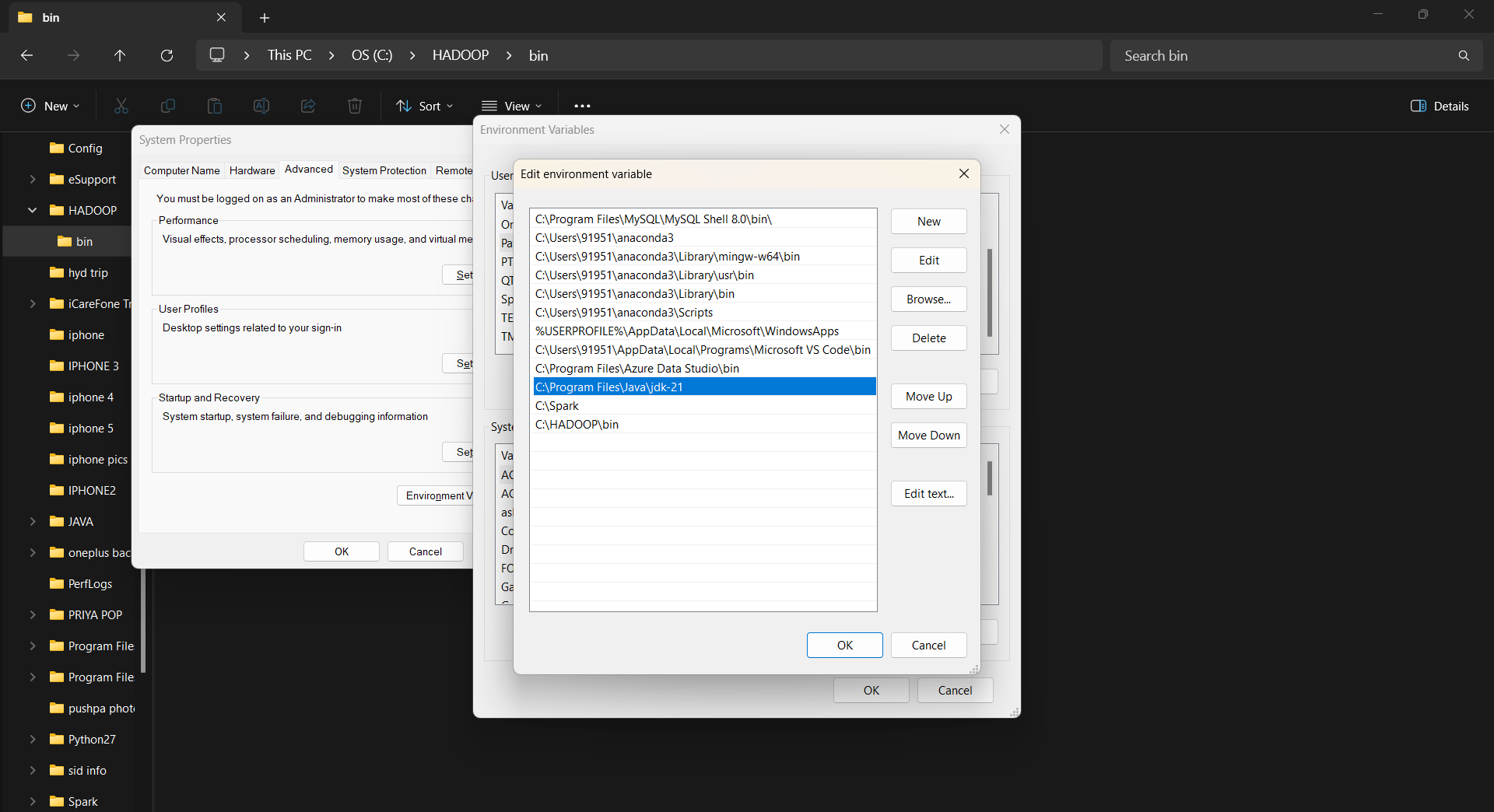
**STEP 6:**

Now we need to edit the Path and add the paths of JAVA, HADOOP AND SPARK in the Path Variable

Below You can find how to add them in the Path Variable

* Select Path from Variable
* You will Find New and click on it
* Now we need to add the paths of Java, Apache and Hadoop in that and click on OK

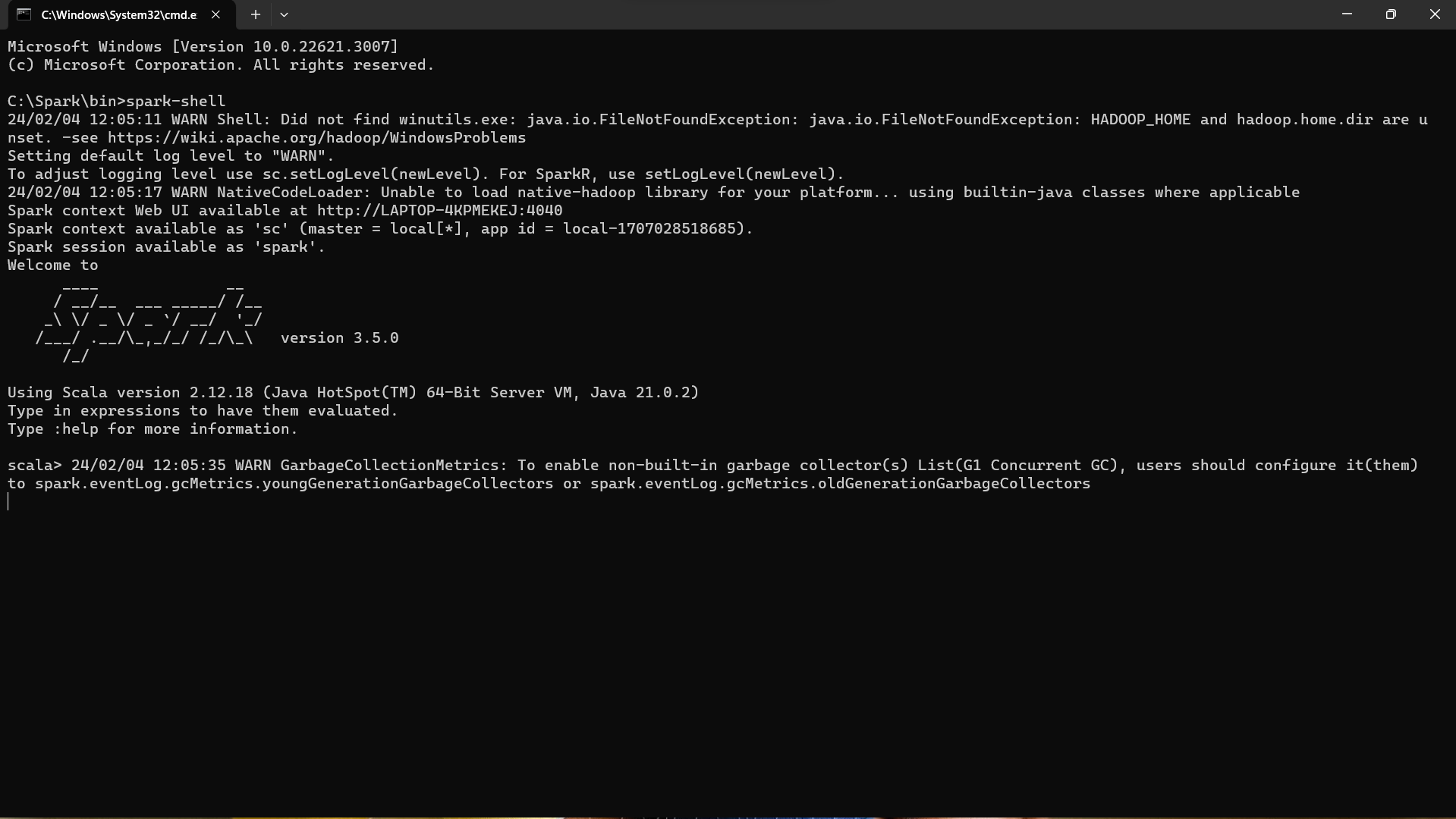




**Step 7:**

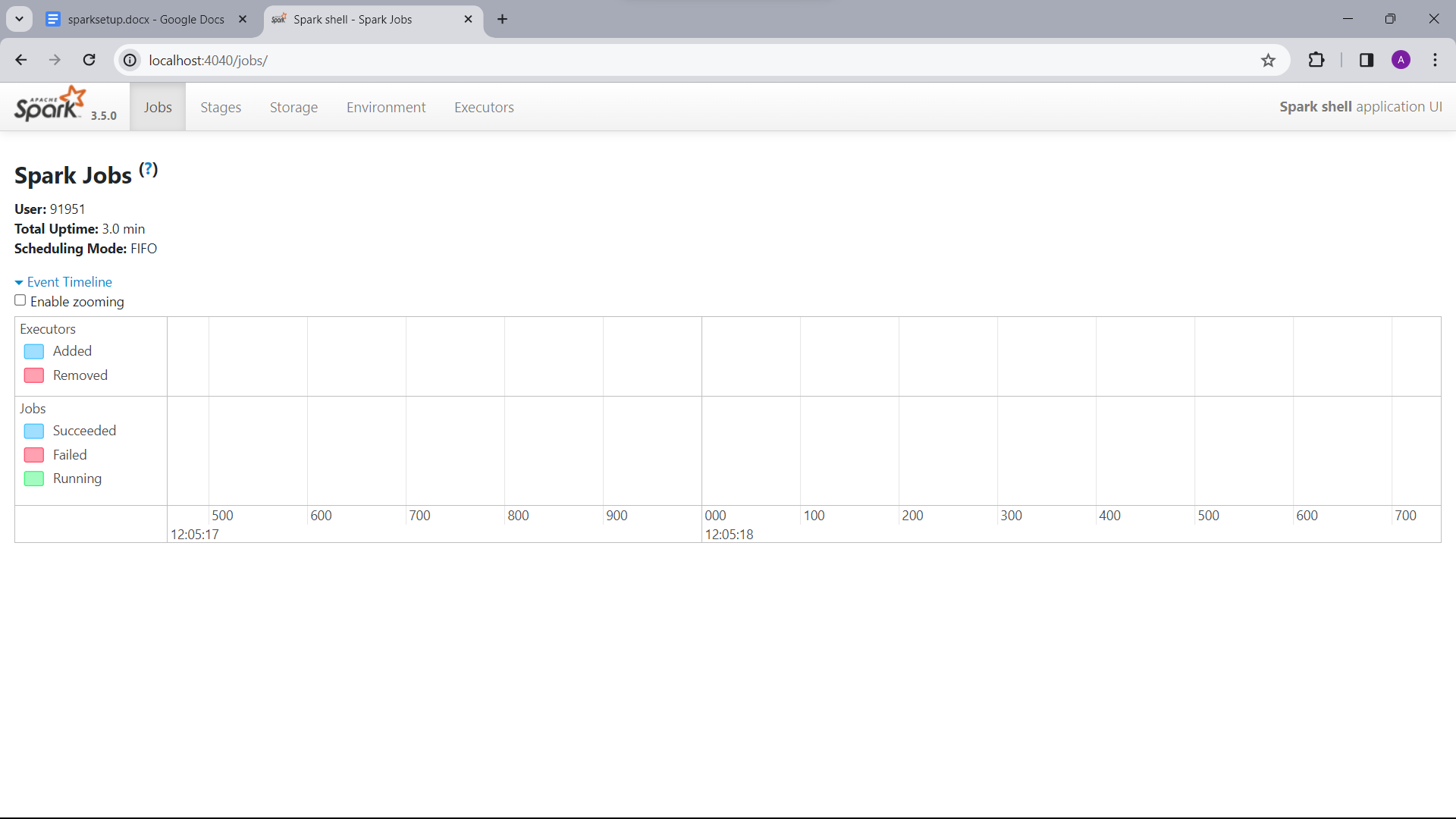
Now we will check whether the Spark is successful running or not using the command prompt

We will check by entering Spark-shell in command prompt and if it displays like the picture below, then we have successfully installed Spark and we have done the Setup.



**Step 8:**

Now lets check the setup which we have done by entering this link **“http://localhost:4040/jobs/ “** in our browser and if it displays the picture below then we have done all right.



We have successfully installed and we have done the setup of APACHE SPARK in our system.