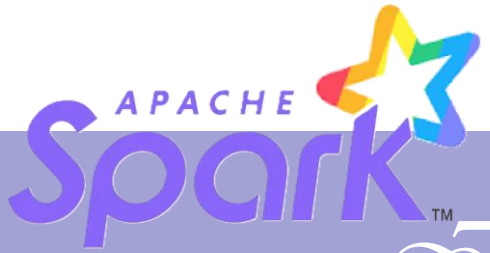
BigData Analytics

Dr. SOHAIL IMRAN



Install

sohailimran2@yahoo.com

Ref: Web

Login and Download

1

Cloud Account

Sign in to Cloud

Sign Up for Free Cloud Tier

Oracle Account

Sohail IMRAN

Help

Sign Out

2 https://www.oracle.com/java/technologies/downloads/#java17-windows

ORACLE Products Industries Resources Customers Partners Developers Company

Java downloads Tools and resources Java archive

Java SE Development Kit 17.0.16 downloads

Java SE subscribers will receive JDK 17 updates until at least **September 2029**.

The Oracle JDK 17 license changed in October 2024

The Oracle Technology Network License Agreement for Oracle Java SE used for JDK 17 updates 17.0.13 and greater is substanti uses, such as personal use and development use, at no cost -- but other uses authorized under prior Oracle JDK licenses may and using this product. FAQs are available here.

Commercial license and support are available for a low cost with Java SE Universal Subscription.

JDK 17.0.16 checksums and OL 9 GPG Keys for RPMs

Linux **macOS** **Windows**

Product/file description	File size	3 Download
x64 Compressed Archive	172.93 MB	jdk-17.0.16_windows-x64_bin.zip
x64 Installer	154.10 MB	jdk-17.0.16_windows-x64_bin.exe
x64 MSI Installer	152.85 MB	jdk-17.0.16_windows-x64_bin.msi

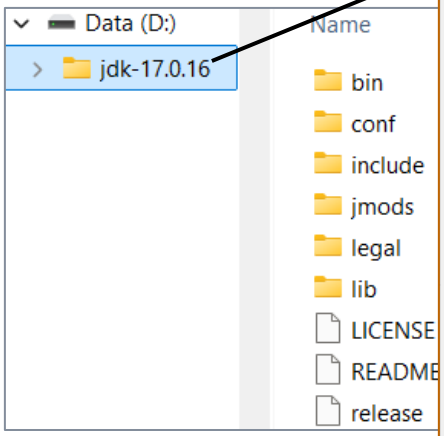
Recent download history

jdk-17.0.16_windows-x64_bin.zip
173 MB • 6 minutes ago

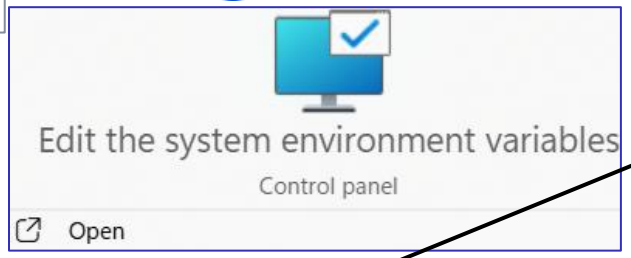
Setup JAVA_HOME

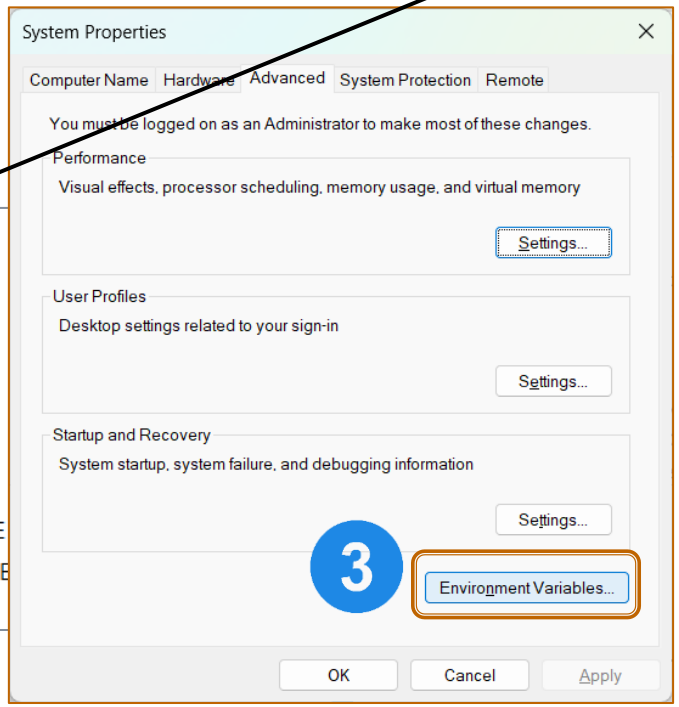
Extract

1



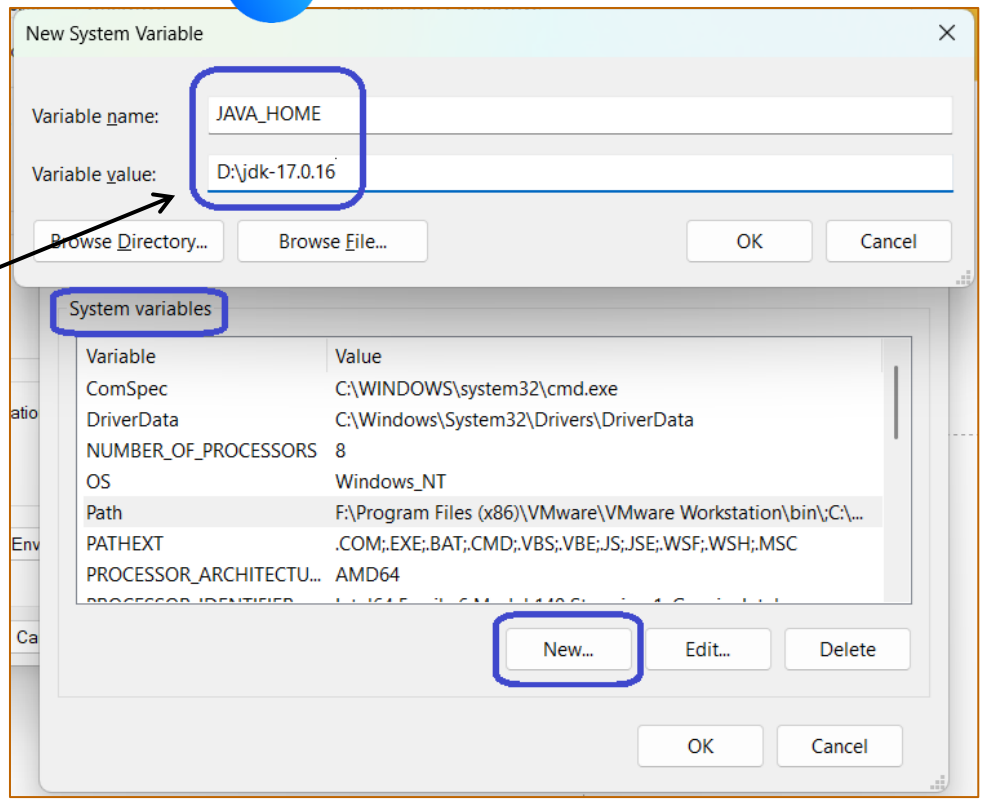
2



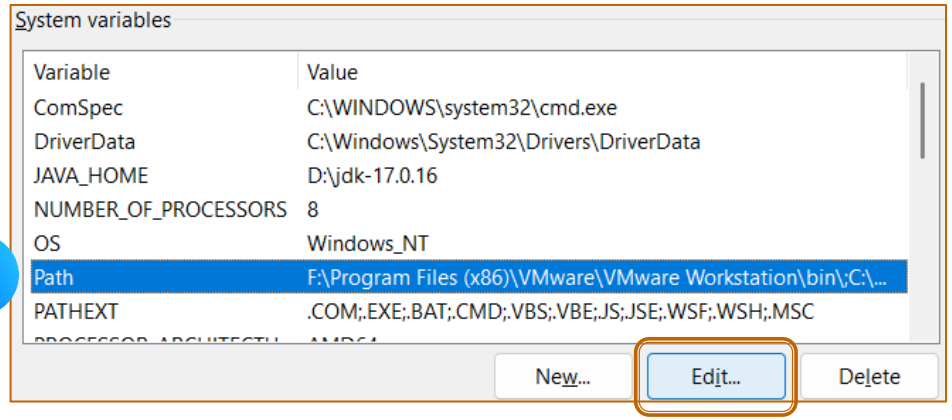


3

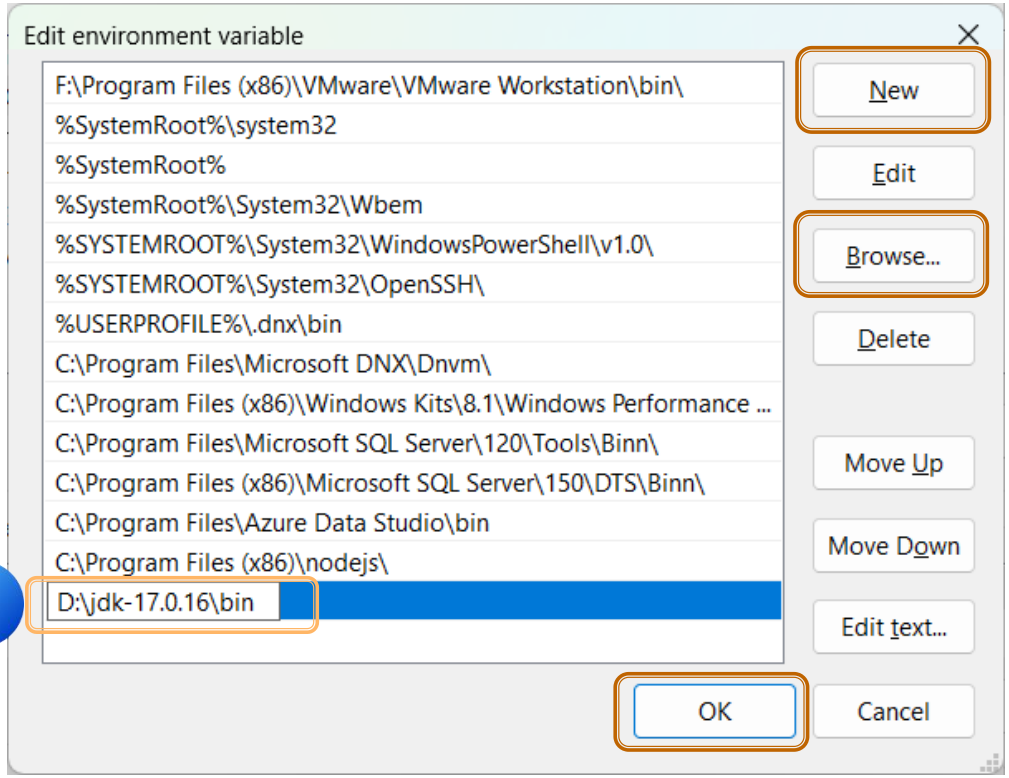
4



Setup - PATH



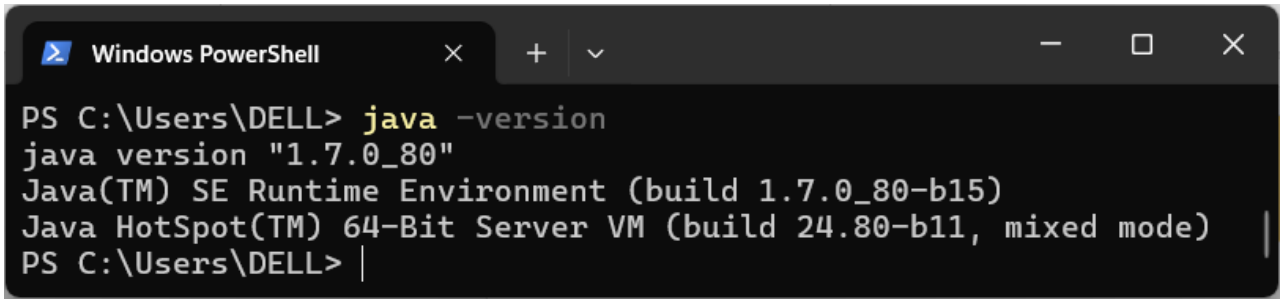
1



2

3

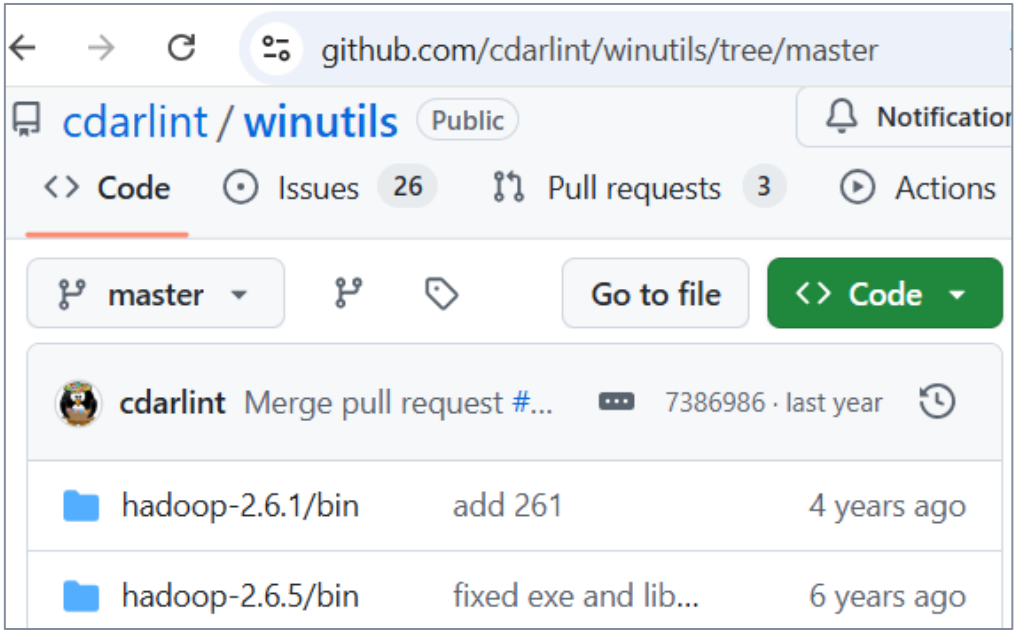
java -version

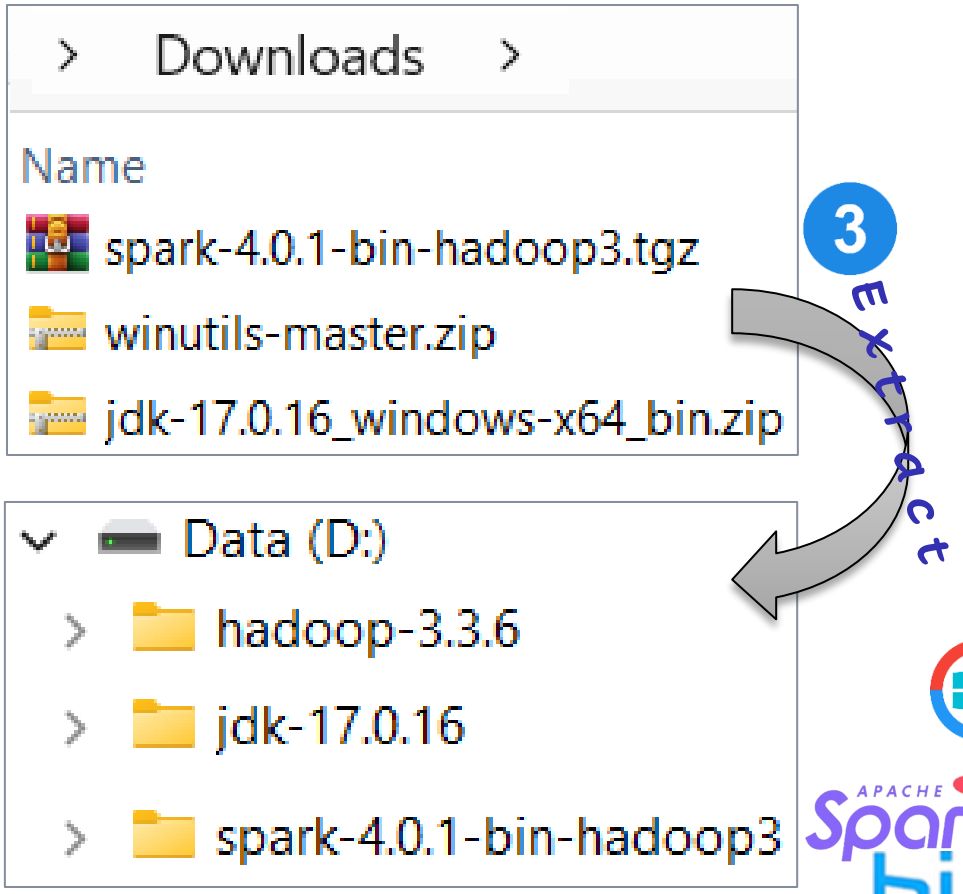


Download & Extract

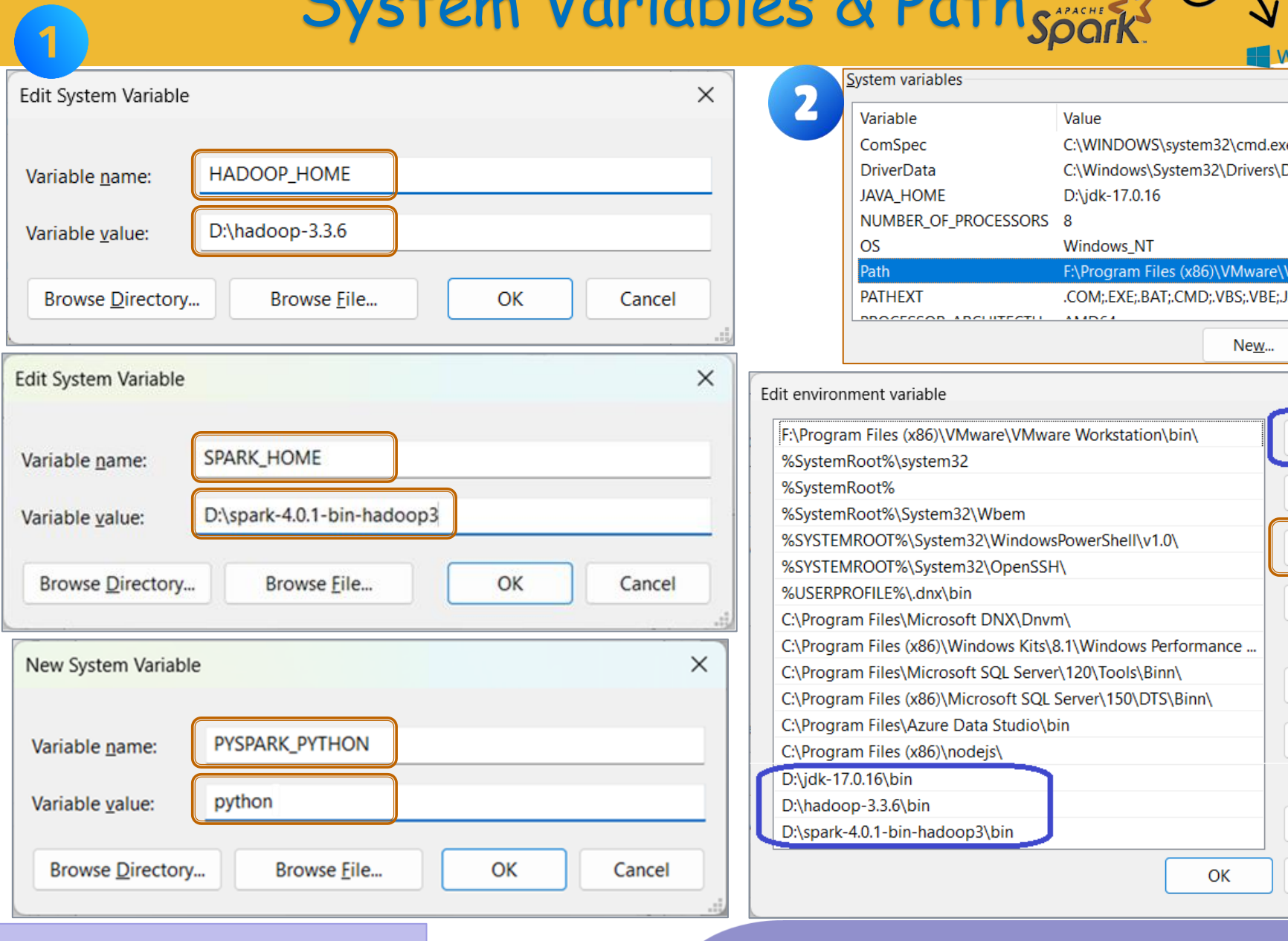
1. https://dlcdn.apache.org/spark/spark-4.0.1/spark-4.0.1-bin-hadoop3.tgz

2. https://codeload.github.com/cdarlint/winutils/zip/refs/heads/master





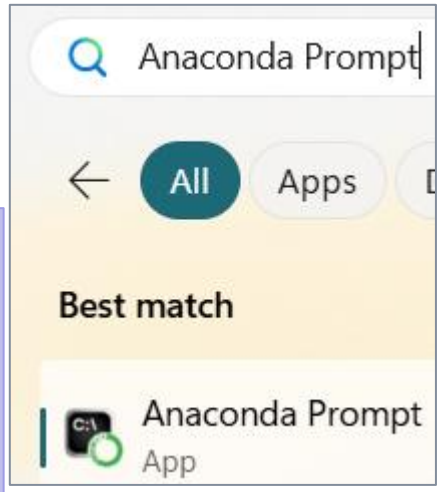
System Variables & Path



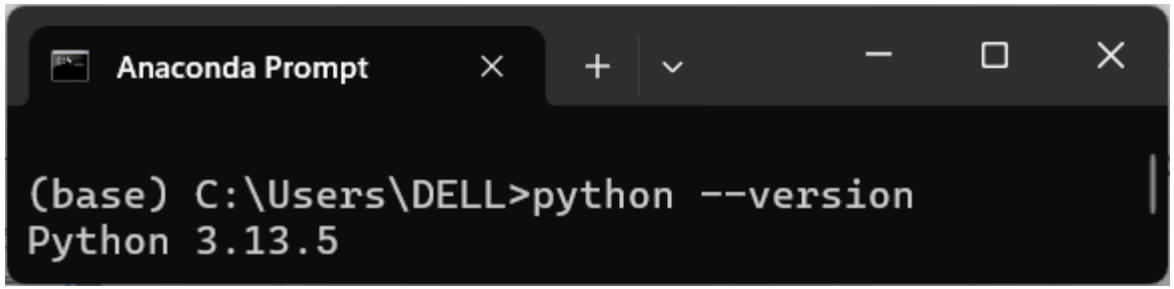
spark-shell

spark-shell

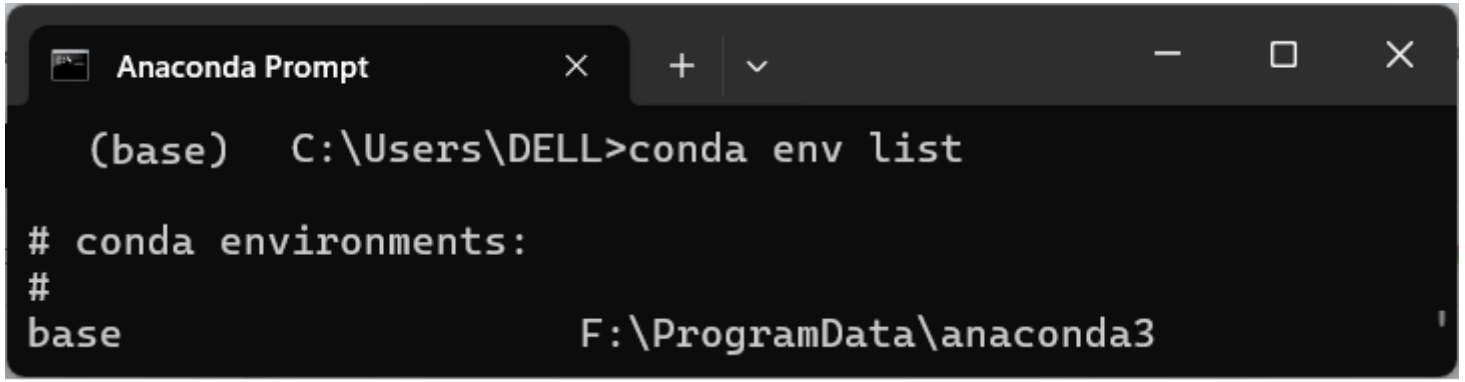
```
Windows PowerShell
PS C:\Users\DELL> spark-shell
WARNING: Using incubator modules: jdk.incubator.vector
Using Spark's default log4j profile: org/apache/spark/log4j2-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /___/ .__/\_,_/_/ /_/\_\   version 4.0.1
      /_/

Using Scala version 2.13.16 (Java HotSpot(TM) 64-Bit Server VM, Java 17.0.16)
Type in expressions to have them evaluated.
Type :help for more information.
Spark context Web UI available at http://DESKTOP-9VAOBA3:4040
Spark context available as 'sc' (master = local[*], app id = local-1761139955553).
Spark session available as 'spark'.

scala>
```



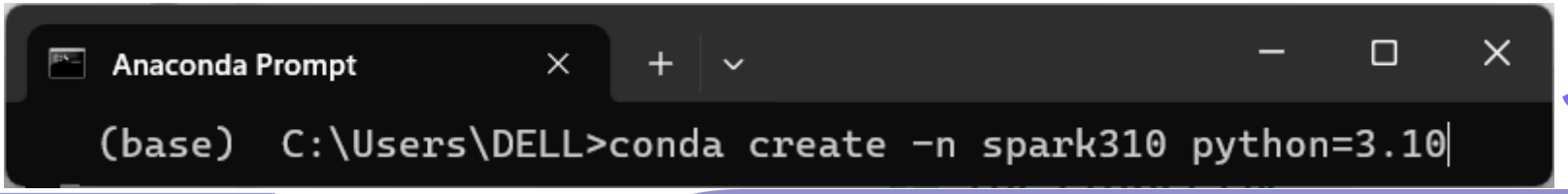
python --version



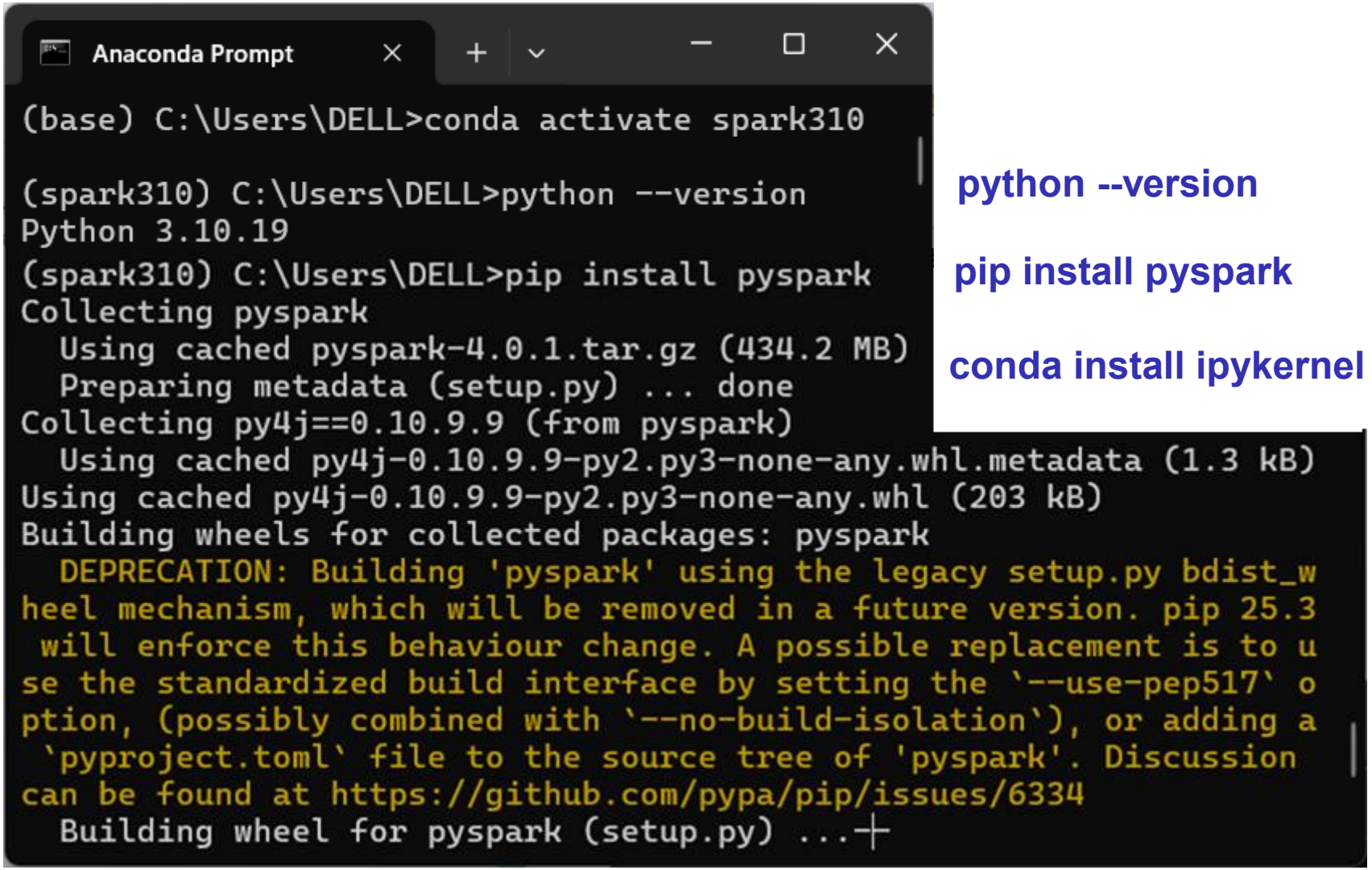
conda env list



conda create –n spark310 python=3.10



conda activate spark310



python --version

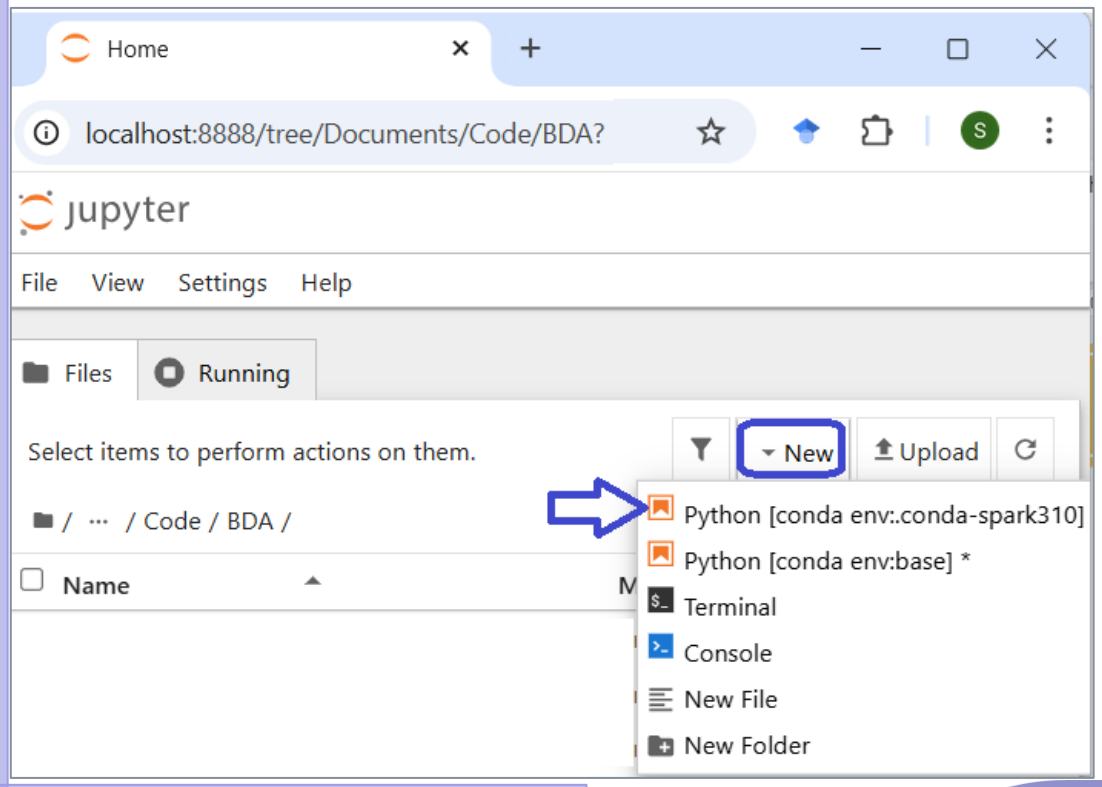
pip install pyspark

conda install ipykernel

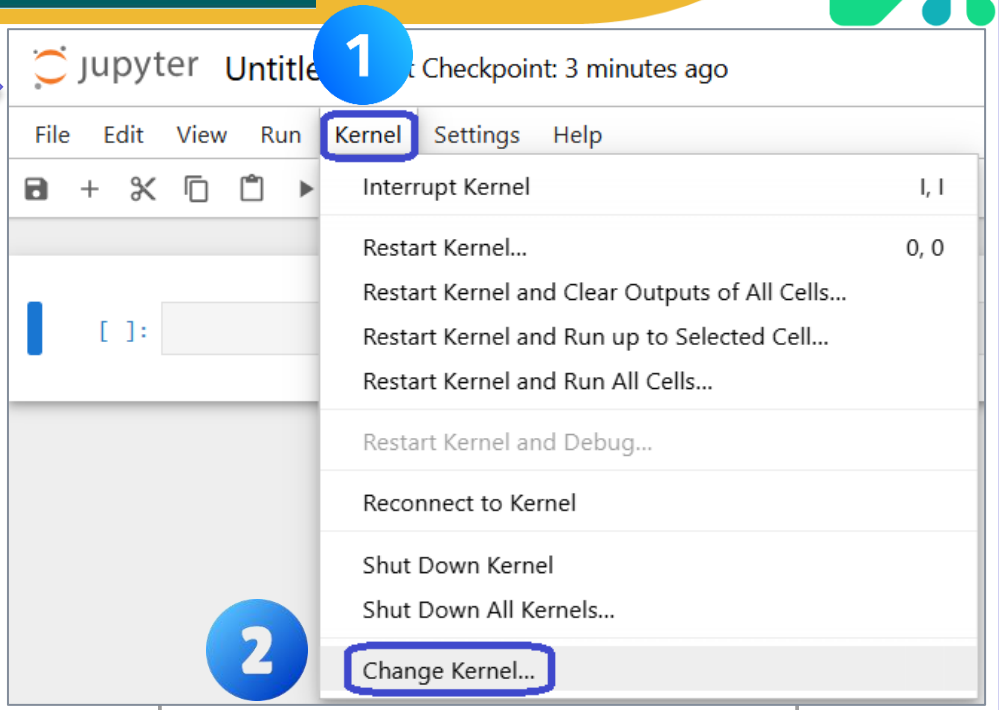
Alternate Method to Connect

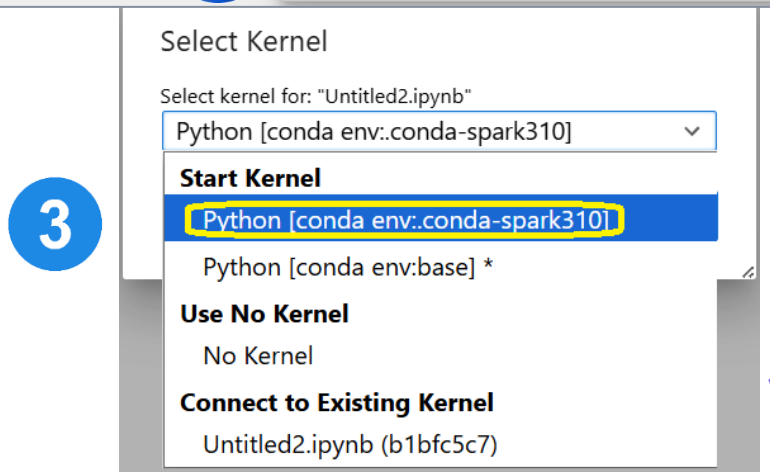
Next Page

New File

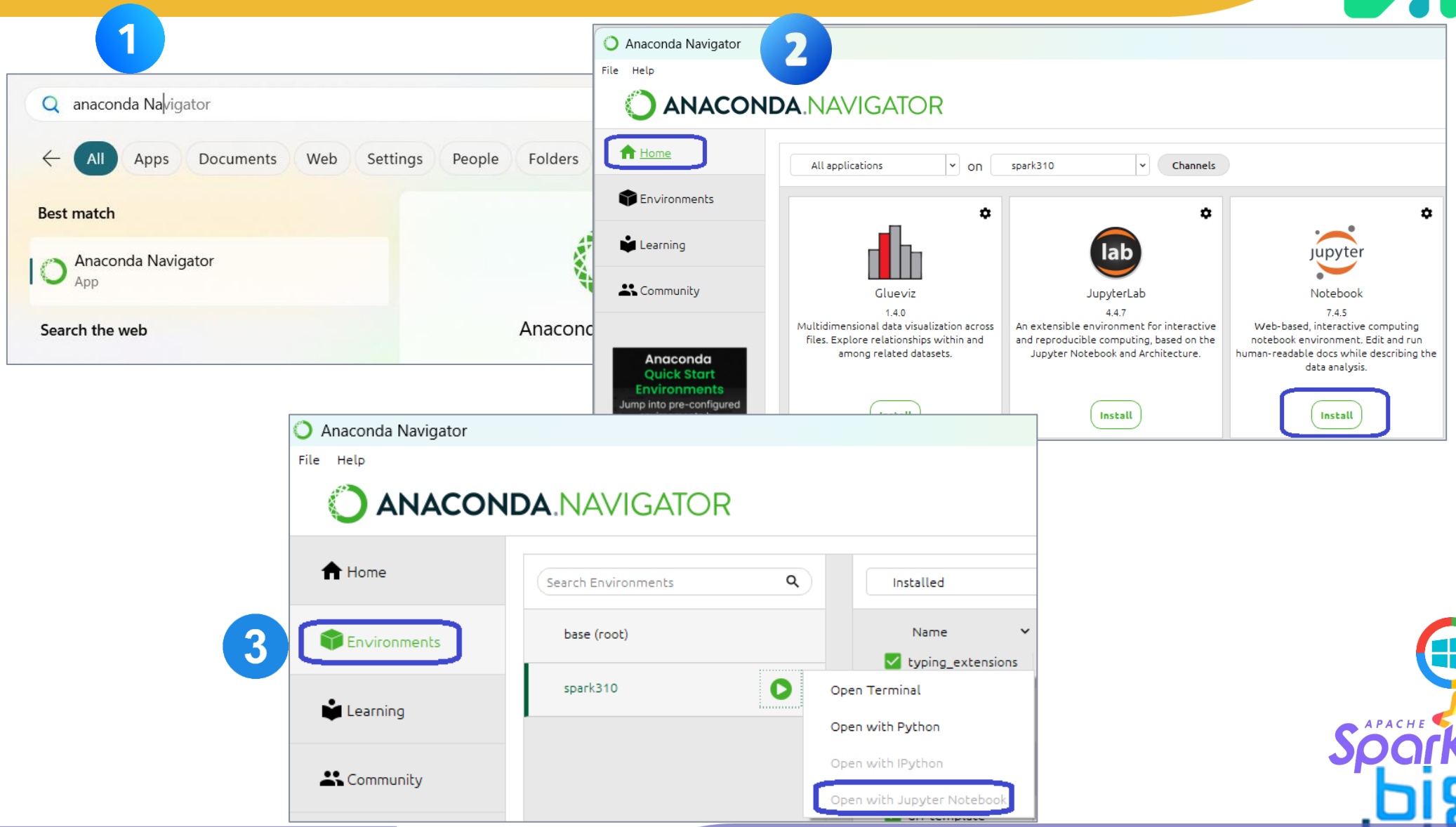


Existing File





Alternate Method to Connect



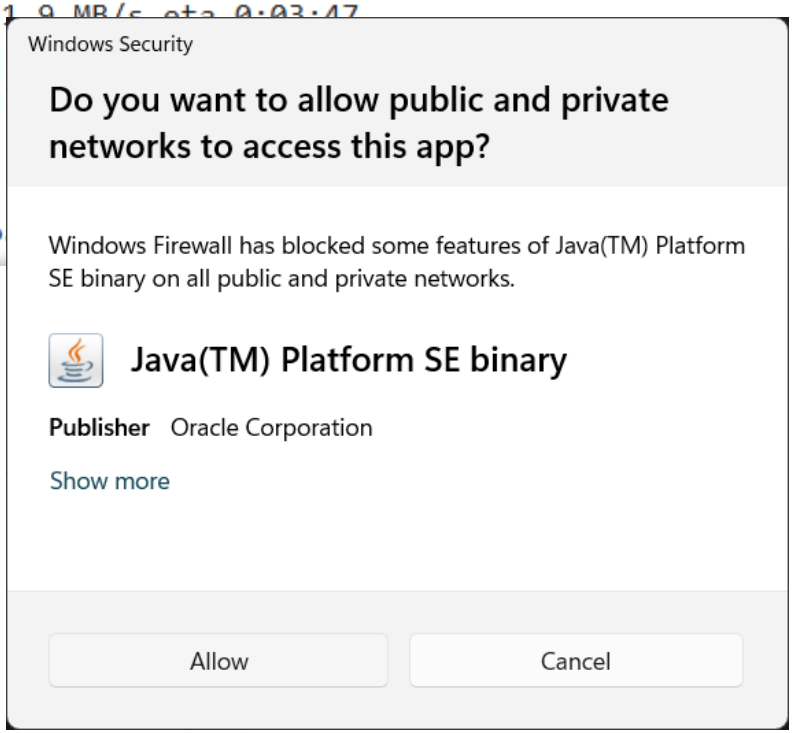
```
from pyspark.sql import SparkSession

spark = SparkSession.builder.appName('pySp0').getOrCreate()
```

```
[1]: from pyspark.sql import SparkSession

[2]: spark = SparkSession.builder.appName('pySp0').getOrCreate()

[3]: spark

[3]: SparkSession - in-memory

     SparkContext

     Spark UI

     Version          v4.0.1
     Master           local[*]
     AppName          pySp0
```

install pyspark

```
pip install pyspark
```

```
Defaulting to user installation because normal site-packages is not writeable
Collecting pyspark
  Downloading pyspark-4.0.1.tar.gz (434.2 MB)
     ---------------------------------------- 0.0/434.2 MB ? eta -:--:--
     ---------------------------------------- 0.3/434.2 MB ? eta -:--:--
     ---------------------------------------- 0.8/434.2 MB 2.3 MB/s eta 0:03:11
     ---------------------------------------- 1.0/434.2 MB 1.9 MB/s eta 0:03:47
     ------------------- -------------------- 1/2 [pyspark]
     ---------------------------------------- 2/2 [pyspark]

  Successfully installed py4j-0.10.9.9 pyspark-4.0.1
  Note: you may need to restart the kernel to use updated p
```



delete py environment

conda deactivate spark310

```
conda remove --name spark310 --all
```