Assignment 0 – Mukesh Ravichandran - 50380788

Apache Spark Installation on my MAC

[Figure 1 - Website->Java Downloads->latest version->ARM64 DMG Installer 2](#_Toc191575279)

[Figure 2- Terminal display of JAVA 23 version on mac 3](#_Toc191575280)

[Figure 3- Apache Spark Website For downloading installer 4](#_Toc191575281)

[Figure 4- Installing Java 11 4](#_Toc191575282)

[Figure 5- pyspark put into test 5](#_Toc191575283)

[Figure 6 Example program 6](#_Toc191575284)

[Figure 7 Apache pyspark working 6](#_Toc191575285)

# Java Installation

Apache Spark runs on java, so installation of JAVA SDK , Since my PC Is M2 MAC 2022 ,

I will have to choose ARM64 which fits for my macOS.

Website->Java Downloads->latest version->ARM64 DMG Installer





Figure 1 - Website->Java Downloads->latest version->ARM64 DMG Installer

# Java Installation

Now open the downloads and just follow the next steps, it will automatically do the rest of the installation , once installed , you may move the installer to trash .  
Go to terminal and type the command

“ java –version”

A screenshot of a computer

AI-generated content may be incorrect.

Figure 2- Terminal display of JAVA 23 version on mac

# Download Apache SPARK

Go to apache spark downloads, , choosing the latest release and package type and download the .tgz file .  
I was pondering what is this new file type , I understood later that   
A .tgz file is a compressed archive that combines multiple files into one using TAR (Tape Archive) and then compresses them using Gzip (.gz).

# 

Figure 3- Apache Spark Website For downloading installer

# Java Version Trouble

After successful installation of apache spark , I encountered with the a problem,

While figuring it out refering the apache spark documentation

Java latest version is incombatible, hence I need to downgrade to java 8/11.

A screen shot of a computer

AI-generated content may be incorrect.

Figure 4- Installing Java 11

# Even After Java 11 I had trouble

This was draining me, I am now having compatible version, then as well encountred issue, I asked Chatgpt , to troubleshoot , It then asked me to check which java version is set as default and apache spark is pointing to , it was still pointing the latest one. Then changed the home to java 11 from java 23

# Finally

After following all the steps, “pyspark” and the command

# A screenshot of a computer program AI-generated content may be incorrect.

Figure 5- pyspark put into test

# Example with PYSPARK I would like to try with simple example from online to check if apache spark is really installed and running fine on my local machine.

# A black background with green text AI-generated content may be incorrect.

Figure 6 Example program

# Hurray !

# A screenshot of a computer program AI-generated content may be incorrect.

Figure 7 Apache pyspark working