PySpark installation(Macbook)

Name-Ajay Chaudhary
Batch-Data Engineering(Batch-1)

Steps for installation of Pyspark in macbook-

1. Install PySpark on Mac using Homebrew

```
/bin/bash -c "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

```
Last login: Sat Feb  3 16:59:14 on ttys000
ajaychaudhary@Ajays-MacBook-Air ~ % /bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"

==> Checking for `sudo` access (which may request your password)...
Password:
[==> This script will install:
/usr/local/bin/brew
/usr/local/share/doc/homebrew
/usr/local/share/man/man1/brew.1
/usr/local/share/zsh/site-functions/_brew
/usr/local/etc/bash_completion.d/brew
/usr/local/Homebrew
==> The following existing directories will be made group writable:
/usr/local/bin
==> The following existing directories will have their owner set to ajaychaudhary:
/usr/local/bin
==> The following existing directories will have their group set to admin:
/usr/local/bin
==> The following new directories will be created:
/usr/local/etc
/usr/local/include
/usr/local/lib
/usr/local/sbin
/usr/local/share
/usr/local/var
/usr/local/opt
/usr/local/share/zsh
/usr/local/share/zsh/site-functions
/usr/local/var/homebrew
/usr/local/var/homebrew/linked
/usr/local/Cellar
/usr/local/Caskroom
/usr/local/Frameworks
==> The Xcode Command Line Tools will be installed.
```

2. Install Java Version

```
# Install OpenJDK 11
brew install openjdk@11
```

```
ajaychaudhary@Ajays-MacBook-Air ~ % brew install openjdk11
==> Downloading https://ghcr.io/v2/homebrew/core/openjdk/11/manifests/11.0.22
################################################################## 100.0%
==> Fetching dependencies for openjdk@11: giflib, libpng, freetype, fontconfig, pcre2, gettext, glib, xorgproto, libau, libxdmcp, libxcb, libx11, libxext, libxrender, lzo, pixman, cairo, graphite2, icu4c
, harfbuzz, jpeg-turbo, xz, lz4, zstd, libtiff and little-cms2
==> Downloading https://ghcr.io/v2/homebrew/core/giflib/manifests/5.2.1
################################################################## 100.0%
==> Fetching giflib
==> Downloading https://ghcr.io/v2/homebrew/core/giflib/blobs/sha256:fa6adb4afc1
################################################################## 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/libpng/manifests/1.6.41
################################################################## 100.0%
==> Fetching libpng
==> Downloading https://ghcr.io/v2/homebrew/core/libpng/blobs/sha256:e0940248ae4
################################################################## 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/freetype/manifests/2.13.2-1
################################################################## 100.0%
==> Fetching freetype
==> Downloading https://ghcr.io/v2/homebrew/core/freetype/blobs/sha256:ecd47039b
|################################################################# 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/fontconfig/manifests/2.15.0
################################################################## 100.0%
==> Fetching fontconfig
```

3. Install Scala

```
# Install Scala (optional)
brew install scala
```

4. Install Python

```
# install Python
brew install python
```

```
ajaychaudhary@Ajays-MacBook-Air ~ % brew install python
==> Downloading https://formulae.brew.sh/api/formula.jws.json
#=#=#
==> Downloading https://formulae.brew.sh/api/cask.jws.json
#=#=#
==> Downloading https://ghcr.io/v2/homebrew/core/python/3.11/manifests/3.11.7_1
################################################################### 100.0%
==> Fetching dependencies for python@3.11: mpdecimal, ca-certificates, openssl@3, readline and sqlite
==> Downloading https://ghcr.io/v2/homebrew/core/mpdecimal/manifests/4.0.0-1
################################################################### 100.0%
==> Fetching mpdecimal
==> Downloading https://ghcr.io/v2/homebrew/core/mpdecimal/blobs/sha256:266a3f51
################################################################### 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/ca-certificates/manifests/2023-
################################################################### 100.0%
==> Fetching ca-certificates
==> Downloading https://ghcr.io/v2/homebrew/core/ca-certificates/blobs/sha256:5c
################################################################### 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/openssl/3/manifests/3.2.1
################################################################### 100.0%
==> Fetching openssl@3
==> Downloading https://ghcr.io/v2/homebrew/core/openssl/3/blobs/sha256:f2f4d0a1
################################################################### 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/readline/manifests/8.2.7
################################################################### 100.0%
==> Fetching readline
```

5. Install PySpark on Mac

```
# Install Apache Spark
brew install apache-spark
```

```
ajaychaudhary@Ajays-MacBook-Air ~ % brew install apache-spark
==> Downloading https://ghcr.io/v2/homebrew/core/apache-spark/manifests/3.5.0
################################################################### 100.0%
==> Fetching dependencies for apache-spark: openjdk@17
==> Downloading https://ghcr.io/v2/homebrew/core/openjdk/17/manifests/17.0.9
################################################################### 100.0%
==> Fetching openjdk@17
==> Downloading https://ghcr.io/v2/homebrew/core/openjdk/17/blobs/sha256:19b0e70
################################################################### 100.0%
==> Fetching apache-spark
==> Downloading https://ghcr.io/v2/homebrew/core/apache-spark/blobs/sha256:86def
################################################################### 100.0%
==> Installing dependencies for apache-spark: openjdk@17
==> Installing apache-spark dependency: openjdk@17
==> Downloading https://ghcr.io/v2/homebrew/core/openjdk/17/manifests/17.0.9
Already downloaded: /Users/ajaychaudhary/Library/Caches/Homebrew/downloads/6274f55d293c4208
==> Pouring openjdk@17--17.0.9.monterey.bottle.tar.gz
🍺 /usr/local/Cellar/openjdk@17/17.0.9: 635 files, 304.9MB
```

```
Disable this behaviour by setting HOMEBREW_NO_INSTALL_CLEANUP.
Hide these hints with HOMEBREW_NO_ENV_HINTS (see `man brew`).
ajaychaudhary@Ajays-MacBook-Air ~ % pyspark
Python 3.11.2 (v3.11.2:878ead1ac1, Feb  7 2023, 10:02:41) [Clang 13.0.0 (clang-1300.0.29.30)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
24/02/03 17:51:04 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /__ / .__/\_,_/_/ /_/\_\   version 3.5.0
      /_/

Using Python version 3.11.2 (v3.11.2:878ead1ac1, Feb  7 2023 10:02:41)
Spark context Web UI available at http://192.168.1.7:4040
Spark context available as 'sc' (master = local[*], app id = local-1706962866468).
SparkSession available as 'spark'.
ajaychaudhary@Ajays-MacBook-Air ~ % echo 'export PATH="/usr/local/opt/openjdk@11/bin:$PATH"' >> ~/.zshrc
ajaychaudhary@Ajays-MacBook-Air ~ %  export CPPFLAGS="-I/usr/local/opt/openjdk@11/include"
ajaychaudhary@Ajays-MacBook-Air ~ % brew install scala
==> Downloading https://formulae.brew.sh/api/formula.jws.json
################################################################### 100.0%
==> Downloading https://formulae.brew.sh/api/cask.jws.json
################################################################### 100.0%
==> Downloading https://ghcr.io/v2/homebrew/core/scala/manifests/3.3.1
################################################################### 100.0%
==> Fetching dependencies for scala: openjdk
==> Downloading https://ghcr.io/v2/homebrew/core/openjdk/manifests/21.0.2
################################################################### 100.0%
==> Fetching openjdk
```