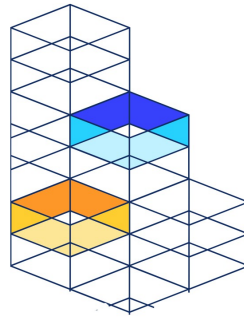




TECH
TALENT
SOUTH



Data Science Introduction & Setup Checklist

As prepared by GS Jackson

Package Manager Installations Regardless of Operating System

Installing and Usage of Conda, Anaconda & Jupyter Notebooks

ANACONDA/CONDA is a editor and emulator framework that includes Jupyter Notebooks that acts as a wiki / compiler for common programming languages and tools. It also includes a variety of other modules using around Big Data & Data Science.

For all operating systems you can download Anaconda or the mini version Conda at - <https://www.continuum.io/downloads>

The **Homebrew** (Mac OS, Linux) and **Chocolatey** (Windows) installation is below.

Installing Data Science Applications & Tools via Scripting Frameworks (Brew & Chocolatey)

Mac OS / Linux

The following programs are for installation and configuration on Mac OS X and Linux around work with Data Science. They include the frameworks for Big Data / Data Science including the programming languages necessary to configure.

For Mac OS X we will be using a scripting program called Homebrew.

The equivalent for Linux is called Linuxbrew.

Homebrew

Install Homebrew on MAC OS X, LINUX go to the website - <http://brew.sh/>

To install Homebrew –

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

If you get an error in the installation –

```
sudo chown -R $USER /usr/local; brew update
brew install wget
```

All installation locations are here -

```
$ cd /usr/local
```

```
$ find Cellar
Cellar/wget/1.16.1
Cellar/wget/1.16.1/bin/wget
Cellar/wget/1.16.1/share/man/man1/wget.1

$ ls -l bin
bin/wget -> ../Cellar/wget/1.16.1/bin/wget
```

Brew uses modules of programs called formulas. You can see view a list of these formulas here - <http://brewformulas.org/>

Install Git Using Brew

```
brew install git
```

Install Anaconda

```
brew cask install anaconda
```

Install Python, Numpy, Scipy, Matplotlib

```
brew tap homebrew/science && brew install python numpy
scipy matplotlib
```

You can learn more about these at their respective repositories:

[homebrew/science](#)

[samueljohn/python](#)

Install Ipython and Corresponding Packages via Brew & PIP

```
brew install zeromq
pip install jinja2
pip install tornado
pip install pyzmq
pip install ipython
```

Install Pandas

```
pip install pandas
```

Testing

Open up Ipython

```
import numpy
import scipy
import matplotlib
import pandas
```

Installing iJavascript for Jupyter Notebooks for Mac OS

iJavascript

```
ruby -e "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/master/install)"
brew install pkg-config node zeromq
sudo easy_install pip
pip install --upgrade pyzmq jupyter
npm install -g ijavascript
ijsinstall
```

To discover what programming languages and modules have been configured with Jupyter you can issue the following command:

```
jupyter kernelspec list
```

De-install Homebrew -

```
/bin/bash -c "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/master/uninstall.s
h)"
```

CHOCOLATEY

Windows Only

Install Windows Powershell

The PowerShell team has also [put together a script](#) you can invoke directly from PowerShell. It's a one-liner cmdlet that downloads and runs the setup wizard automatically. All you have to do is paste the code snippet in and hit the Enter key.

Fire up PowerShell and copy/paste the following cmdlet into the window:

```
iex "& { $(irm https://aka.ms/install-powershell.ps1) } -UseMSI"
```

Install Chocolatey

Running PowerShell as an Administrator cut and paste this into the command line -

```
Set-ExecutionPolicy Bypass -Scope Process -Force;  
[System.Net.ServicePointManager]::SecurityProtocol =  
[System.Net.ServicePointManager]::SecurityProtocol -bor 3072; iex  
( (New-Object  
System.Net.WebClient).DownloadString('https://chocolatey.org/install.  
ps1'))
```

If you don't see any errors, you are ready to use Chocolatey!

```
choco
```

Install Git Using Chocolatey

```
choco install git.install
```

Install Anaconda Using Chocolatey

```
choco install anaconda3
```

Install Python

```
choco install python
```

Install Numpy

```
choco instal numpy
```

Installation and Usage of Github Regardless of MacOS, Windows, or Linux

Signup for a GITHUB account –

<https://github.com/>

Our Tech Talent South Github location is -

<https://github.com/enterlifeonline/techtalentsouth>

Pull Github TTS Data Science Repository

Initialized it with a LICENSE and README.md

Create a local directory called GITHUB

Created a local directory called REPOSITORIES

Within the `./Repositories` directory

From the local command line:

```
git clone https://github.com/enterlifeonline/techtalentsouth.git
```

```
git config --global user.name "Your Name"
```

```
git config --global user.email your.email.domain.com
```

Then from the local command line:

```
git status
```

–

When you want to update your Tech Talent South repository -

Within the `./Repositories` directory

```
rm -rf ./TechTalentSouth
```

```
git clone https://github.com/enterlifeonline/techtalentsouth.git
```

Download Jupyter Notebooks for Mac OS

Download or fork exercise Jupyter notebooks for Javascript, R, Julia, and Python – <https://github.com/enterlifeonline/techtalentsouth>

Windows

The following programs are for installation and configuration on Windows around work with Data Science. They include the frameworks for Big Data / Data Science including the programming languages necessary to configure.

Chocolatey (like Homebrew)

Some version of Windows such as Windows 8 or earlier requires you to install install Powershell 3 - <http://www.microsoft.com/en-us/download/details.aspx?id=34595> - however Windows 10 and above normally have it pre-installed.

Scoop is a scripting framework that allows you to keep up to date on all the latest tools - <http://scoop.sh/>

Note: if you get an error you might need to change the execution policy (i.e. enable Powershell) `$ set-executionpolicy unrestricted -s cu`

Install Scoop

```
iex (new-object  
net.webclient).downloadstring('https://get.scoop.sh')
```

Install Core, Utilities and Tools using Scoop

```
scoop install 7zip coreutils curl git grep openssh  
sed wget vim grep
```

Install Java using Scoop

```
scoop install openjdk
```

(Optional) Install Oracle JDK but it defaults to OpenJDK

```
scoop bucket add extras  
scoop install oraclejdk  
scoop reset oraclejdk
```

Installing programming languages using Scoop


```
scoop install python ruby go nodejs
```

Installing Anaconda (Jupyter Notebooks) using Scoop for Windows

```
scoop install conda
```

Installing Anaconda (Jupyter Notebooks) without using Scoop for Windows

Or to install Anaconda download from here to run executable -

<https://www.anaconda.com/distribution/>

Installing iJavascript for Anaconda (Jupyter Notebooks) for Windows

See documentation - <https://github.com/n-riesco/ijavascript#installation>

More documentation - <https://hackernoon.com/data-science-for-javascript-developers-2cc3fb1fbefc>

Once installed, open the Anaconda prompt and run:

```
conda install nodejs  
npm install -g ijavascript  
ijsinstall
```

Then you can run `jupyter notebook` in your terminal to load Jupyter Notebook. When you create a new Jupyter Notebook, you should see the Javascript (Node) kernel available.

To discover what programming languages and modules have been configured with Jupyter you can issue the following command:

```
jupyter kernelspec list
```

Installation of Big Data & Data Science tools and technologies

–

```
scoop install maven
```

```
scoop install scala
```

Download Jupyter Notebooks for Windows

Download or fork exercise notebooks for Javascript, R, Julia, and Python –
<https://github.com/enterlifeonline/techtalentsouth>