CSC 301, Software Installation

Note: This document may be updated as needed throughout the semester

Overview

In this class we will use several tools for web development and web scraping using Python and R. In particular, we will work with the following (installation instructions are below):

1. *Python* (https://www.python.org/), a general purpose programming language
2. *Jupyter Notebook*s (https://jupyter.org/), a web-based platform for creating, explaining, and sharing code.
3. *R* (https://cran.rstudio.com/) and *R Studio* (https://rstudio.com/), a programming language and IDE for data visualization and analysis (and for creating dynamic web pages).

Text editors for web development

For **Windows** users, I recommend downloading Notepad++ (https://notepad-plus-plus.org/), a user-friendly text editor that has syntax highlighting for many languages. Notepad++ is what we will use in class.

For **Mac** users, I recommend downloading Brackets (http://brackets.io/) , a text editor designed for web developers.

Installing Jupyter Notebook and Spyder through the Anaconda Distribution

Install the Anaconda Distribution for your system by following the directions at the following link: https://www.anaconda.com/distribution/

The Anaconda Distribution comes with *Jupyter Notebook*.

Running Jupyter Notebook

Once *Jupyter Notebook* is installed, you can run Jupyter Notebook by double clicking the icon or by typing the following using your machine's **terminal** or **Anaconda Command Prompt**:

```
jupyter notebook
```

For more information, see https://jupyter.readthedocs.io/en/latest/install.html

Installing Selenium

Selenium (https://selenium.dev/) is a framework for browser automation (for writing code that opens and uses web browsers). In order to install selenium for Python, type the following command in your **terminal** or **Anaconda prompt**.

```
conda install -c conda-forge selenium
```

You will also need to install a web driver, for communicating with automated browsers. To install this, type the following in your **terminal** or **Anaconda prompt**.

```
conda install -c conda-forge geckodriver
```

The webdriver we will use is for the Firefox browser, and Firefox also must be installed: https://www.mozilla.org/en-US/firefox/new/

If installed correctly, you should be able to execute the following commands in Python without any errors. The code below uses Selenium to open Eastern's homepage.

```
from selenium import webdriver
driver = webdriver.Firefox()
driver.get('https://easternct.edu')
```

Installing R/RStudio

1. Install R for your system from https://cran.rstudio.com/
2. Install RStudio by clicking the download button on https://rstudio.com/products/rstudio/download/#download and following the instructions.

R package installation

In class we will be using the packages *dplyr* (for manipulating data), *shiny* (for development of web applications), and *rvest* (for web scraping).

These packages can be installed by running the commands below from within R Studio:

```
install.packages('dplyr')
install.packages('shiny')
install.packages('rvest')
```

Note that during installation you will receive red messages (and possibly warnings) showing the progress of the installation. This is normal. However, if you get an error, then the package was not installed and you should contact me to help troubleshoot the issue.

After a package is installed, it can be loaded using the *library* function, as in the code below:

```
library(dplyr)
library(shiny)
library(rvest)
```