

# Installation

This class will involve some coding, for which you will need some basic tools.

During the upcoming semester you will learn a lot more about these tools, so don't worry if this is all new and perhaps a bit frightening right now.

## Gephi

Gephi is an open-source software for network visualization. We will use it during the data exercises tomorrow.

To install Gephi go to <https://gephi.org/>, download the installation file and follow the installation instructions.

## Python and Jupyter Notebook

[Python](#) is a free programming language. We will use the distribution called [Anaconda](#) as it comes with the most essential for working with data, statistical computing and visualizations. We will use Anaconda for Python 3.7 throughout the course, so:

### **EXERCISE;**

#### **Download and install Anaconda:**

Anaconda can be downloaded [here](#), for Windows, Mac or Linux. If you want to watch a step-by-step tutorial on how to install Anaconda and Jupyter Notebook for your machine see the guides here:

Install Anaconda for Windows by [following these steps](#) or [watch this video](#)

Install Anaconda for Mac by [following these steps](#) or [watch this video](#)

Since the vast majority of our coding will be in the Python language, we will use an integrated development environment ([IDE](#)). IDEs integrate text editing, syntax highlighting, and version control, simplifying the coding process.

These packages are automatically included in Anaconda Jupyter Notebook. [Jupyter Notebook](#): It's free and modern, and if you're new to Python this will make it much easier to get started.

All Python coding will be done in Jupyter Notebook.

## Verifying the installation

After installation of Python please to execute a number of commands in the shell. You need to first open your local shell, this is the **Anaconda Prompt** on Windows or **terminal** on Linux/Mac. See details [here](#). Once the shell is open type **python** - this will start Python. Once you have Python started please try the following two commands to verify that it's working.

```
1+2
```

```
>>> 3
```

```
print('Welcome to Introduction to Social Data Science')
```

```
>>> Welcome to Introduction to Social Data Science
```

## Welcome to open source

To know core Python is powerful in itself, but the great potentials lie in the huge community of developers and researchers contributing to a shared pool of software packages. A programming language is as powerful as the community that surrounds it. Especially in the field of machine learning, the Python community is leading the way, allowing you to share code with top researchers from the field and industry, among others Google's top engineering teams. Tapping into these vast resources is made easy by the Conda distribution and the pip package manager. Just open your shell/command-line/terminal and type the following: **conda install [name of package]**

or if conda does not support it directly use the more generic package manager: **pip install [name of package]**

## Using Jupyter notebook

The [Jupyter Notebook App](#) can be launched by clicking on the Jupyter Notebook icon installed by Anaconda in the start menu (only Windows) or by typing the following either in a terminal (on Mac/Linux) or cmd (on Windows).

```
jupyter notebook
```

This will launch a new browser window (or a new tab) showing the Notebook Dashboard, a sort of control panel that allows (among other things) to select which notebook to open.

## Getting friendly with Jupyter

Try to spend a little time familiarizing yourself with the Jupyter framework. For instance, try learning a few of your editor's keyboard shortcuts; see our post [here](#). The point is to be as productive as possible when working with the computer. Karl

Broman, a professor of biostatistics and medical informatics at the University of Wisconsin-Madison, gives some great advice for working with code: