# Software specifications

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chapter number | Software required (With version) | Free/Proprietary | If proprietary, can code testing be performed using a trial version | If proprietary, then cost of the software | Download links to the software | **Hardware specifications** | OS required |
| 1 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 1 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 1 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 1 | IPython 3.2.1 | Free |  |  | http://ipython.org/install.html |  | All |
| 1 | VirtualBox 4.3.28-100309 | Free |  |  | https://www.virtualbox.org/wiki/Downloads |  |  |
| 1 | Vagrant 1.7.2 | Free |  |  | https://www.vagrantup.com/downloads.html |  |  |
| 1 | Docker 1.7.0 | Free |  |  | https://docs.docker.com/index.html |  |  |
| 1 | watermark 1.2.2 | Free |  |  | https://raw.githubusercontent.com/rasbt/watermark/master/watermark.py |  |  |
| 1 | AppDirs 1.4.0 | Free |  |  | https://pypi.python.org/pypi/appdirs/1.4.0 |  |  |
| 1 | pandas 0.16.2 | Free |  |  | http://pandas.pydata.org/pandas-docs/dev/install.html |  |  |
| 1 | scikit-learn 0.16.1 | Free |  |  | http://scikit-learn.org/dev/install.html |  |  |
| 1 | pep8 1.6.2 | Free |  |  | https://pep8.readthedocs.org/en/latest/intro.html |  |  |
| 1 | tabulate 0.7.5 | Free |  |  | https://pypi.python.org/pypi/tabulate |  |  |
| 1 | landslide 1.1.3 | Free |  |  | https://pypi.python.org/pypi/landslide |  |  |
| 2 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 2 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 2 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 3 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 3 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 3 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 4 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 4 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 4 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 5 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 5 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 5 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 6 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 6 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 6 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 7 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 7 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 7 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 8 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 8 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 8 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 9 | NumPy 1.9.2 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 9 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 9 | Matplotlib 1.4.3 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 10 | NumPy 1.10.1 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 10 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 10 | Matplotlib 1.5.0 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 11 | NumPy 1.10.1 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 11 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 11 | Matplotlib 1.5.0 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |
| 12 | NumPy 1.10.1 | Free |  |  | http://sourceforge.net/projects/numpy/files/ |  | All |
| 12 | SciPy 0.16.0 | Free |  |  | http://sourceforge.net/projects/scipy/files/ |  | All |
| 12 | Matplotlib 1.5.0 | Free |  |  | http://sourceforge.net/projects/matplotlib/files/ |  | All |

# Detailed installation steps (software-wise)

The steps should be listed in a way that it prepares the system environment to be able to test the codes of the book.

For all chapters Python 3 is required, as well as pip. To install Python go to https://wiki.python.org/moin/BeginnersGuide/Download TO install pip go to http://pip.readthedocs.org/en/latest/installing.html. Instructions to install software are given throughout the chapters. Mostly we need to run in admin mode the following command:

pip install <some software>

First, you need a Python 3 distribution. I recommend the full Anaconda distribution (how to install Anaconda is explained in Chapter 1 in the Setting up Anaconda recipe) as it comes with the majority of the software we need. I tested the code with Python 3.4 and the following packages:

* joblib 0.8.4
* IPython 3.2.1
* NetworkX 1.9.1
* NLTK 3.0.2
* Numexpr 2.3.1
* pandas 0.16.2
* SciPy 0.16.0
* seaborn 0.6.0
* sqlalchemy 0.9.9
* statsmodels 0.6.1
* matplotlib 1.5.0
* NumPy 1.10.1
* scikit-learn 0.17
* dautil 0.0.1a29

For some recipes you need to install extra software, but this is explained whenever the software is required. The following script installs most of the software (it assumes Anaconda):

export PATH=/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:${PATH}

apt-get install -y libgfortran3

conda config --set always\_yes True

conda install beautiful-soup bokeh=0.9.1 execnet=1.3.0 fastcache=1.0.2 \

joblib=0.8.4 jsonschema ipython=3.2.1 lxml mpmath=0.19 \

networkx=1.9.1 nltk=3.0.2 numba=0.22.1 numexpr=2.3.1 \

pandas=0.16.2 pyzmq scipy=0.16.0 seaborn=0.6.0 \

sqlalchemy=0.9.9 statsmodels=0.6.1 terminado=0.5 tornado

conda install matplotlib=1.5.0 numpy=1.10.1 scikit-learn=0.17

pip install dautil==0.0.1a29

pip install hiveplot==0.1.7.4

pip install landslide==1.1.3

pip install LiveStats==1.0

pip install mpld3==0.2

pip install pep8==1.6.2

pip install requests-cache==0.4.10

pip install tabulate==0.7.5

find /opt/conda -name \\_\_pycache\_\_ -depth -exec rm -rf {} \;

rm -r /opt/conda/pkgs/\*

mkdir -p /.cache/dautil/log

mkdir -p /.local/share/dautil