

# Python Packages Cheat Sheet

## List of Popular Packages

### Overview

If you are using Anaconda to manage your Python packages, for any of the following packages that need to be installed, copy the code to install into the Anaconda Prompt. When prompted, enter **y** to confirm yes to continue. Most of the packages should be pre-installed with a core installation of Anaconda; however, the ones highlighted in the summary table will most likely not be installed with a fresh installation of Anaconda. For a list of all packages that are pre-installed with a new installation of Anaconda please see: <https://docs.anaconda.com/anaconda/packages/pkg-docs/>

```
(newBase) C:\Users\bogda>conda install -c conda-forge tabula-py
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: C:\Users\bogda\Anaconda3\envs\newBase

added / updated specs:
- tabula-py

The following packages will be downloaded:

package | build | size | channel
-----|-----|-----|-----
openjdk-11.0.8 | 0 | 167.3 MB | conda-forge
tabula-py-1.4.1 | py38h32f6830_2 | 10.0 MB | conda-forge
-----|-----|-----|-----
Total: | | 177.3 MB |

The following NEW packages will be INSTALLED:

openjdk | conda-forge/win-64::openjdk-11.0.8-0
tabula-py | conda-forge/win-64::tabula-py-1.4.1-py38h32f6830_2
vs2013_runtime | pkgs/main/win-64::vs2013_runtime-12.0.21005-1

Proceed ([y]/n)? y
```

### Popular Packages

In the next two pages we have summarized 25 of the top Python packages used by business professionals that are also taught in the various Python courses at [Marquee](#). The packages are listed in alphabetical order in the summary tables and also by usage category below:

- **Data manipulation:** Numpy; Pandas
- **Web scraping:** BeautifulSoup; Requests; Selenium; Urllib3
- **Visualization:** Bokeh; Matplotlib; Plotly; Seaborn
- **Dashboarding:** Dash; Streamlit
- **Statistical analysis:** SciPy; Statsmodels
- **File management:** Camelot; OS; Pathlib; Pillow; Shutil; Tabula
- **Machine learning:** NLTK; OpenCv; PyTesseract; Scikit-learn; SpaCy

# Python Packages Cheat Sheet

## List of Popular Packages

Package Name	Conda Code to Install	Category	Additional Information
beautifulsoup4	<code>conda install -c anaconda beautifulsoup4</code> Should already come preinstalled with Anaconda	Web Scraping	Links: <a href="https://www.crummy.com/software/BeautifulSoup/bs4/doc/">https://www.crummy.com/software/BeautifulSoup/bs4/doc/</a> <a href="https://anaconda.org/anaconda/beautifulsoup4">https://anaconda.org/anaconda/beautifulsoup4</a>
bokeh	<code>conda install -c bokeh bokeh</code> Should already come preinstalled with Anaconda	Visualization	Links: <a href="https://docs.bokeh.org/en/latest/index.html">https://docs.bokeh.org/en/latest/index.html</a> <a href="https://anaconda.org/bokeh/bokeh">https://anaconda.org/bokeh/bokeh</a>
camelot-py	<code>conda install -c conda-forge camelot-py</code>	File Manipulation (reading PDF files)	Links: <a href="https://camelot-py.readthedocs.io/">https://camelot-py.readthedocs.io/</a> <a href="https://anaconda.org/conda-forge/camelot-py">https://anaconda.org/conda-forge/camelot-py</a>  Note: Tkinter and Ghostscript applications are also needed for more advanced uses of Camelot. Instructions on installing them can be found here: <a href="https://camelot-py.readthedocs.io/en/master/user/install-deps.html">https://camelot-py.readthedocs.io/en/master/user/install-deps.html</a>
dash	<code>conda install -c conda-forge dash</code>	Dashboard	Links: <a href="https://anaconda.org/conda-forge/dash">https://anaconda.org/conda-forge/dash</a> <a href="https://dash.plotly.com/layout">https://dash.plotly.com/layout</a>
matplotlib	<code>conda install -c conda-forge matplotlib</code> Should already come preinstalled with Anaconda	Visualization	Links: <a href="https://matplotlib.org/contents.html">https://matplotlib.org/contents.html</a> <a href="https://anaconda.org/conda-forge/matplotlib">https://anaconda.org/conda-forge/matplotlib</a>
nltk	<code>conda install -c anaconda nltk</code> Should already come preinstalled with Anaconda	Machine Learning (NLP)	Links: <a href="http://www.nltk.org/">http://www.nltk.org/</a> <a href="https://anaconda.org/anaconda/nltk">https://anaconda.org/anaconda/nltk</a>
numpy	<code>conda install -c anaconda numpy</code> Should already come preinstalled with Anaconda	Data Manipulation	Links: <a href="https://numpy.org/doc/stable/">https://numpy.org/doc/stable/</a> <a href="https://anaconda.org/anaconda/numpy">https://anaconda.org/anaconda/numpy</a>
opencv	<code>conda install -c conda-forge opencv</code>	Machine Learning (OCR)	Links: <a href="https://anaconda.org/conda-forge/opencv">https://anaconda.org/conda-forge/opencv</a>
os	<i>*part of the standard Python Library</i>	File Management	Link: <a href="https://docs.python.org/3/library/os.html">https://docs.python.org/3/library/os.html</a>
pandas	<code>conda install -c anaconda pandas</code> Should already come preinstalled with Anaconda	Data Manipulation	Links: <a href="https://pandas.pydata.org/docs/">https://pandas.pydata.org/docs/</a> <a href="https://anaconda.org/anaconda/pandas">https://anaconda.org/anaconda/pandas</a>
pathlib	<i>*part of the standard Python Library</i>	File Management	Link: <a href="https://pathlib.readthedocs.io/en/pep428/">https://pathlib.readthedocs.io/en/pep428/</a>
pillow	<code>conda install -c conda-forge pillow</code> Should already come preinstalled with Anaconda	File Management (reading image files)	Links: <a href="https://pillow.readthedocs.io/en/stable/">https://pillow.readthedocs.io/en/stable/</a> <a href="https://anaconda.org/conda-forge/pillow">https://anaconda.org/conda-forge/pillow</a>
plotly	<code>conda install -c plotly plotly</code>	Visualization	Links: <a href="https://plotly.com/python/plotly-express/">https://plotly.com/python/plotly-express/</a> <a href="https://anaconda.org/plotly/plotly">https://anaconda.org/plotly/plotly</a>
pytesseract	<code>conda install -c conda-forge pytesseract</code>	Machine Learning (OCR)	Links: <a href="https://github.com/madmaze/pytesseract">https://github.com/madmaze/pytesseract</a> <a href="https://anaconda.org/conda-forge/pytesseract">https://anaconda.org/conda-forge/pytesseract</a>

# Python Packages Cheat Sheet

## List of Popular Packages

Package Name	Conda Code to Install	Category	Additional Information
requests	<code>conda install -c anaconda requests</code> Should already come preinstalled with Anaconda	Web Scraping	Links: <a href="https://requests.readthedocs.io/en/master/">https://requests.readthedocs.io/en/master/</a> <a href="https://anaconda.org/anaconda/scipy">https://anaconda.org/anaconda/scipy</a>
scikit-learn	<code>conda install -c anaconda scikit-learn</code> Should already come preinstalled with Anaconda	Statistical Analysis	Links: <a href="https://scikit-learn.org/">https://scikit-learn.org/</a> <a href="https://anaconda.org/anaconda/scikit-learn">https://anaconda.org/anaconda/scikit-learn</a>
scipy	<code>conda install -c anaconda scipy</code> Should already come preinstalled with Anaconda	Statistical Analysis	Links: <a href="https://www.scipy.org/">https://www.scipy.org/</a> <a href="https://anaconda.org/anaconda/statsmodels">https://anaconda.org/anaconda/statsmodels</a>
seaborn	<code>conda install -c anaconda seaborn</code> Should already come preinstalled with Anaconda	Visualization	Links: <a href="https://seaborn.pydata.org/">https://seaborn.pydata.org/</a> <a href="https://anaconda.org/anaconda/seaborn">https://anaconda.org/anaconda/seaborn</a>
selenium	<code>conda install -c conda-forge selenium</code>	Web Scraping	Links: <a href="https://selenium-python.readthedocs.io/installation.html">https://selenium-python.readthedocs.io/installation.html</a> <a href="https://anaconda.org/conda-forge/selenium">https://anaconda.org/conda-forge/selenium</a>  Note: the corresponding Chrome driver should also be downloaded from here: <a href="https://sites.google.com/a/chromium.org/chromedriver/downloads">https://sites.google.com/a/chromium.org/chromedriver/downloads</a>
shutil	<i>*part of the standard Python Library</i>	File Management	Link: <a href="https://docs.python.org/3/library/shutil.html">https://docs.python.org/3/library/shutil.html</a>
spacy	<code>conda install -c conda-forge spacy</code>	Machine Learning (NLP)	Links: <a href="https://spacy.io/">https://spacy.io/</a> <a href="https://anaconda.org/conda-forge/spacy">https://anaconda.org/conda-forge/spacy</a>
spacy-model-en_core_web_sm	<code>conda install -c conda-forge spacy-model-en_core_web_sm</code>	Machine Learning (NLP)	Links: <a href="https://anaconda.org/conda-forge/spacy-model-en_core_web_sm">https://anaconda.org/conda-forge/spacy-model-en_core_web_sm</a>
statsmodels	<code>conda install -c anaconda statsmodels</code> Should already come preinstalled with Anaconda	Statistical Analysis	Links: <a href="https://www.statsmodels.org/stable/index.html">https://www.statsmodels.org/stable/index.html</a> <a href="https://anaconda.org/anaconda/statsmodels">https://anaconda.org/anaconda/statsmodels</a>
streamlit	<code>pip install streamlit</code>	Dashboard	Links: <a href="https://docs.streamlit.io/en/stable/">https://docs.streamlit.io/en/stable/</a> <a href="https://docs.streamlit.io/en/stable/troubleshooting/clean-install.html">https://docs.streamlit.io/en/stable/troubleshooting/clean-install.html</a>
tabula-py	<code>conda install -c conda-forge tabula-py</code>	File Manipulation (reading PDF files)	Links: <a href="https://tabula-py.readthedocs.io/en/latest/">https://tabula-py.readthedocs.io/en/latest/</a> <a href="https://anaconda.org/conda-forge/tabula-py">https://anaconda.org/conda-forge/tabula-py</a>
urllib3	<code>conda install -c anaconda urllib3</code> Should already come preinstalled with Anaconda	Web Scraping	Links: <a href="https://urllib3.readthedocs.io/en/latest/">https://urllib3.readthedocs.io/en/latest/</a> <a href="https://anaconda.org/anaconda/urllib3">https://anaconda.org/anaconda/urllib3</a>