

Python packages

Machine learning with Python for finance professionals



Packages

- Packages are libraries of code written to solve a particular set of problems
- In Python, there are many related packages relevant to data analytics and machine learning: pandas, scikit-learn, NumPy, etc
- These are installed with pip or conda
 - `pip install <package_name>`
 - `conda install <package_name>`

Tutorial on managing and installing packages with Anaconda:

<https://docs.anaconda.com/anaconda/navigator/tutorials/manage-packages/>

What packages do I need?



pandas: manipulate data



SciPy/NumPy: scientific computing and numerical calculations



scikit-learn: machine learning



matplotlib/Seaborn: data visualisation



spacy/nltk: natural language processing



statsmodels: statistical tests



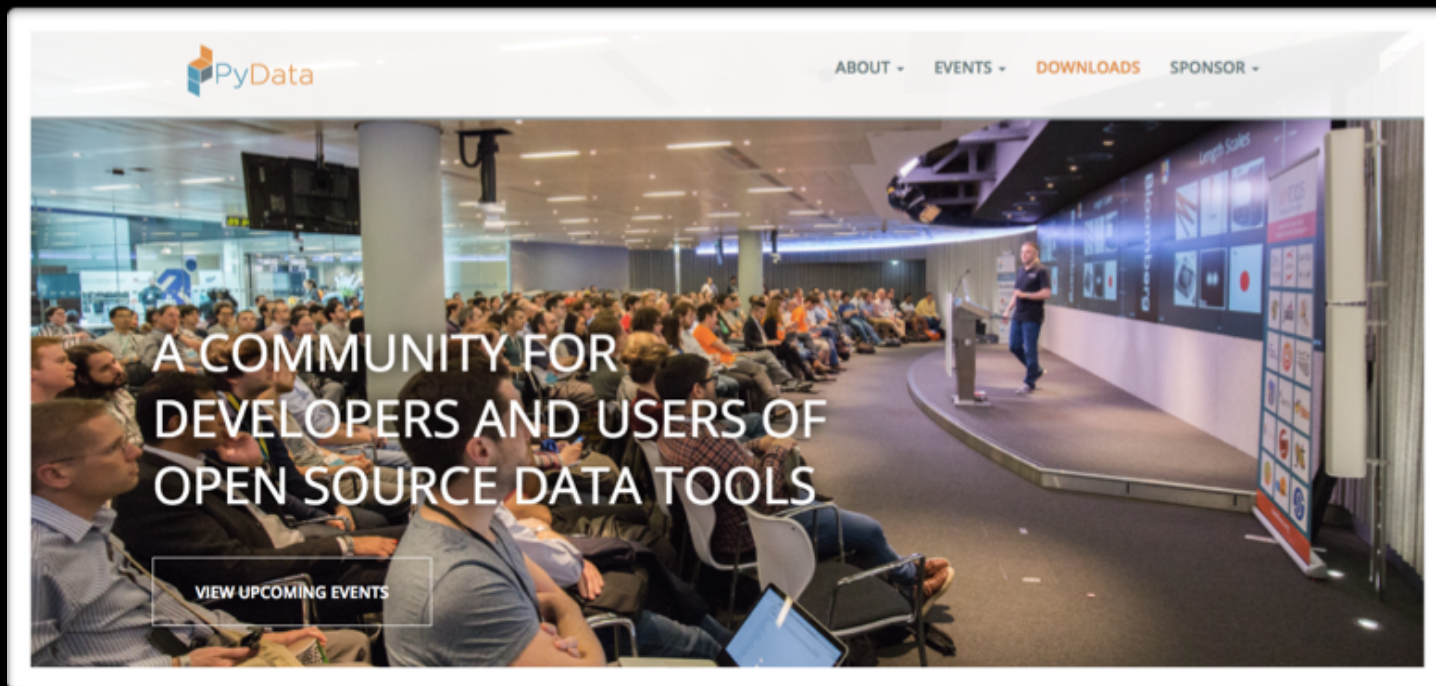
Beautiful Soup: HTML/XML data & web scrapers



Jupyter: interactive programming environment

The PyData Ecosystem

“The global PyData network promotes discussion of best practices, new approaches, and emerging technologies for data management, processing, analytics, and visualization.”



<https://pydata.org/>

The PyData Ecosystem

- 170k+ members
- 184 meetups
- 66 countries
- Promotes sharing ideas & learning from each other



PyData

66 countries • 184 groups • 177634 members



Upcoming events Groups Near Me



PyData

Members

177,634

Groups

184

Countries

66

<https://www.meetup.com/pro/pydata/>



Think Ahead

© ACCA

Public

