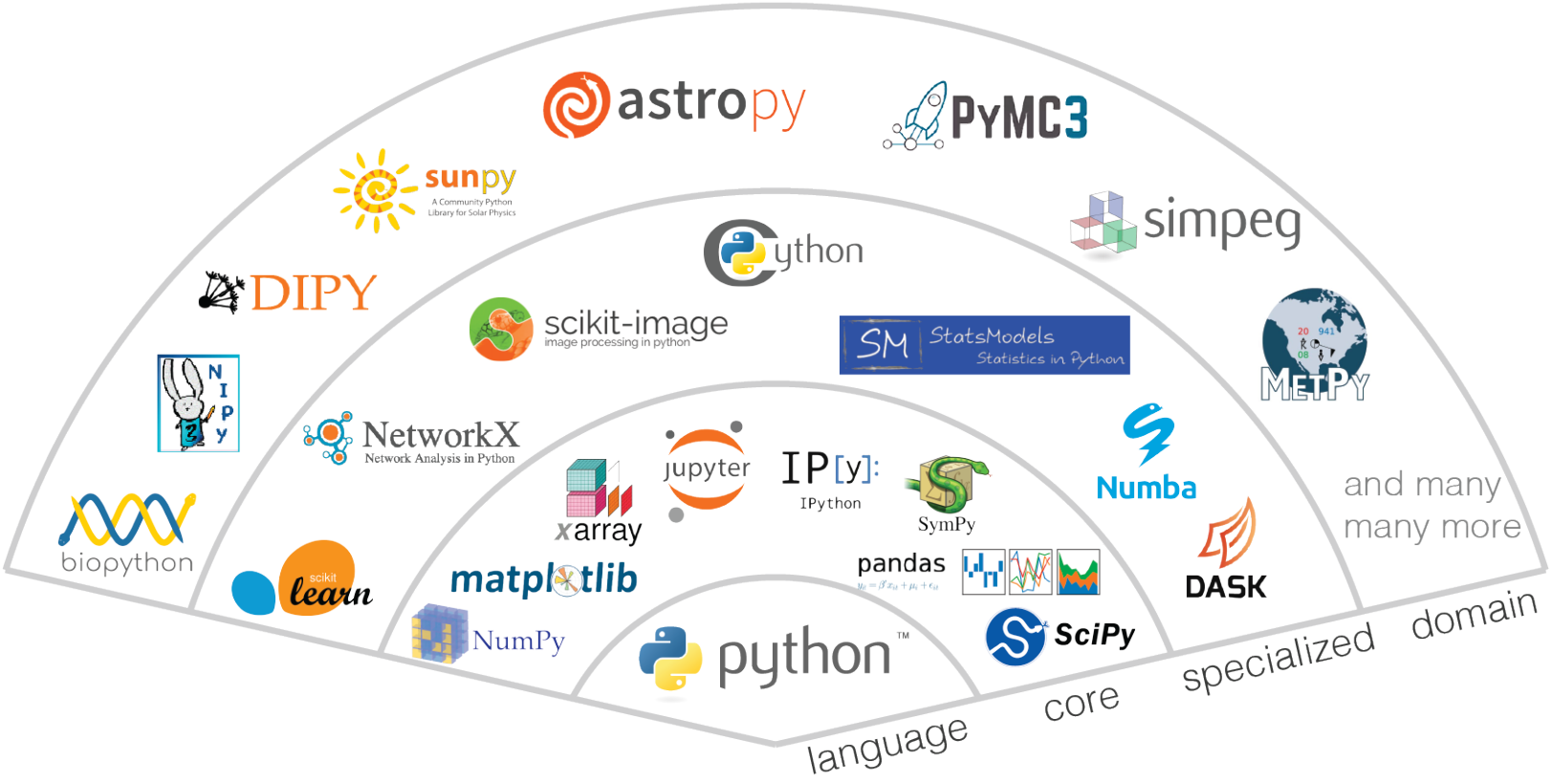
De la bibliothèque standard à l'univers

Des modules disponibles :

`https://docs.python.org/3/library/`

Un écosystème stabilisé et interdépendant pour le traitement de données



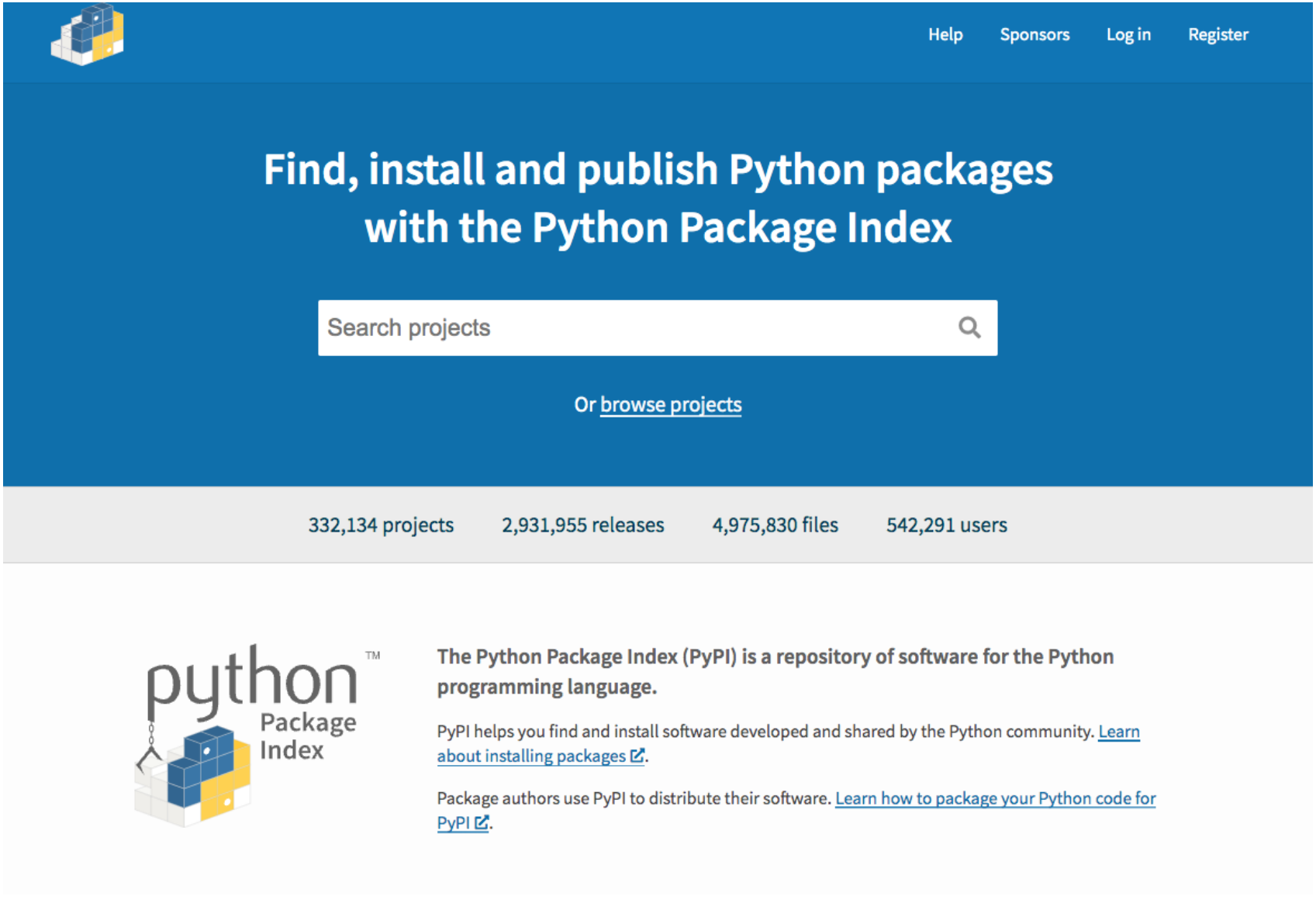
`https://jupytearth.org/jupyter-resources/introduction/ecosystem.html`

Particularité de Python : quelques bibliothèques structurantes bien maintenues.

Et toutes les autres bibliothèques existantes

De la supernovae aux petites comètes passagères...

Accès aux bibliothèques



`https://pypi.org/`

Se repérer dans la jungle

- Des bibliothèques stables, bien intégrées et maintenues
- Des bibliothèques qui viennent, qui partent
- Et tous les autres codes disponibles...

Donc :

1. Développer des compétences sur des bibliothèques stables
2. ... son sens de l'orientation dans l'*open source*
3. ... regarder régulièrement ce qui se fait.

Et surtout : lire la documentation et les exemples !

Quelques bibliothèques

- Pandas, pour la manipulation de tableaux : ce sera la séance de la semaine prochaine
- Matplotlib, qui fixe une grammaire des visualisations 2D, la semaine d'après
- Numpy, caché dans presque toutes les bibliothèques, pour le calcul numérique

Pandas

pandas About us Getting started Documentation Community Contribute

pandas

pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool,
built on top of the Python programming language.

Install pandas now!

Getting started

- Install pandas
- Getting started

Documentation

- User guide
- API reference
- Contributing to pandas
- Release notes

Community

- About pandas
- Ask a question
- Ecosystem

With the support of:

NUMFOCUS OPEN CODE = BETTER SCIENCE · TWO SIGMA · URSA LABS · d-fine · TIDELIFT · Chan Zuckerberg Initiative

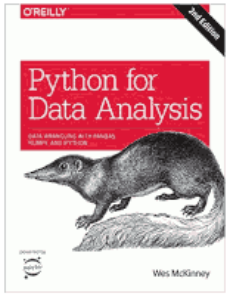
Latest version: 1.4.2

- What's new in 1.4.2
- Release date: Apr 02, 2022
- Documentation (web)
- Documentation (pdf)
- Download source code

Follow us

Follow @pandas_dev

Get the book



Previous versions

- 1.4.1 (Feb 12, 2022)

Scipy

Fundamental algorithms for scientific computing in Python

GET STARTED

SciPy 1.7.2 released 2021-11-05

FUNDAMENTAL ALGORITHMS

SciPy provides algorithms for optimization, integration, interpolation, eigenvalue problems, algebraic equations, differential equations, statistics and many other classes of problems.

BROADLY APPLICABLE

The algorithms and data structures provided by SciPy are broadly applicable across domains.

FOUNDATIONAL

Extends NumPy providing additional tools for array computing and provides specialized data structures, such as sparse matrices and k-dimensional trees.

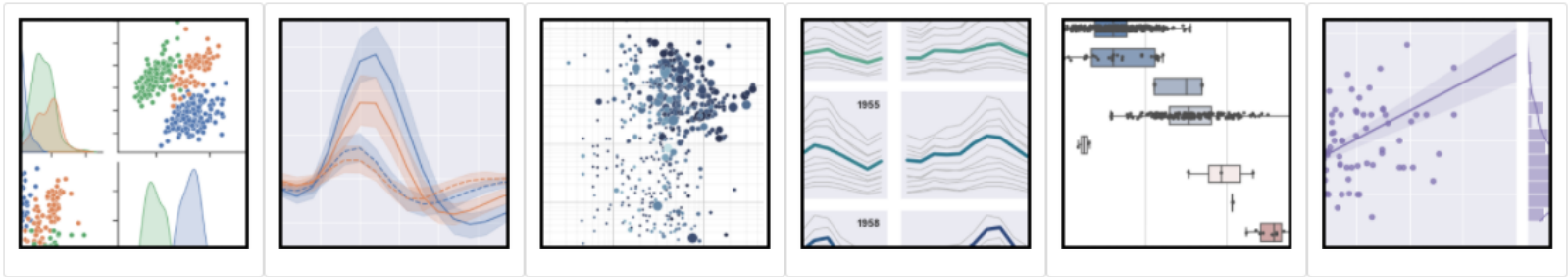
PERFORMANT

SciPy wraps highly-optimized implementations written in low-level languages like Fortran, C, and C++. Enjoy the flexibility of Python with the speed of compiled code.

Seaborn

seaborn 0.11.2 Gallery Tutorial API Site Page Search

seaborn: statistical data visualization



Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

For a brief introduction to the ideas behind the library, you can read the introductory notes or the paper. Visit the installation page to see how you can download the package and get started with it. You can browse the example gallery to see some of the things that you can do with seaborn, and then check out the tutorial or API reference to find out how.

To see the code or report a bug, please visit the GitHub repository. General support questions are most at home on stackoverflow or discourse, which have dedicated channels for seaborn.

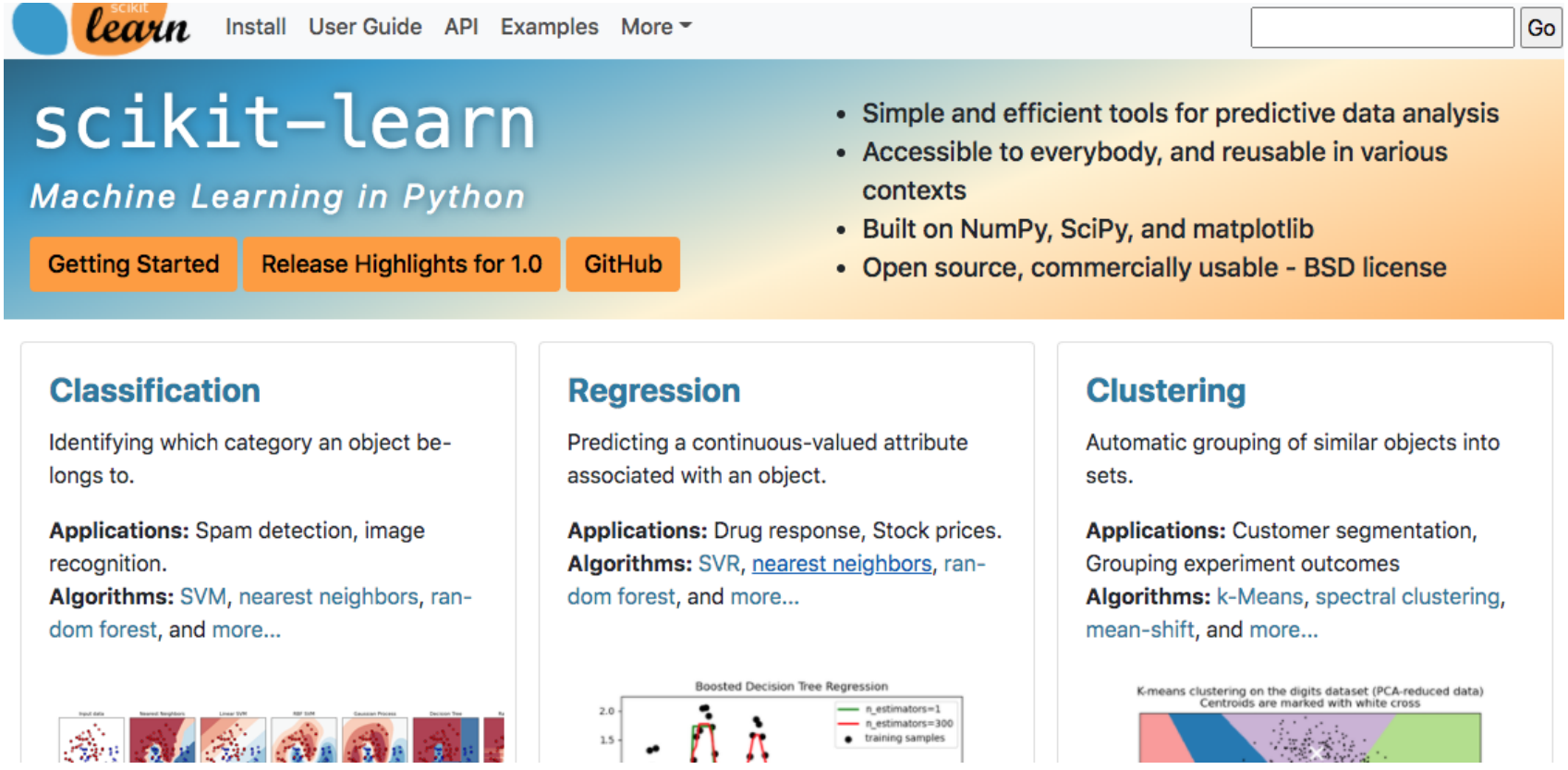
Contents

- Introduction
- Release notes
- Installing
- Example gallery
- Tutorial
- API reference

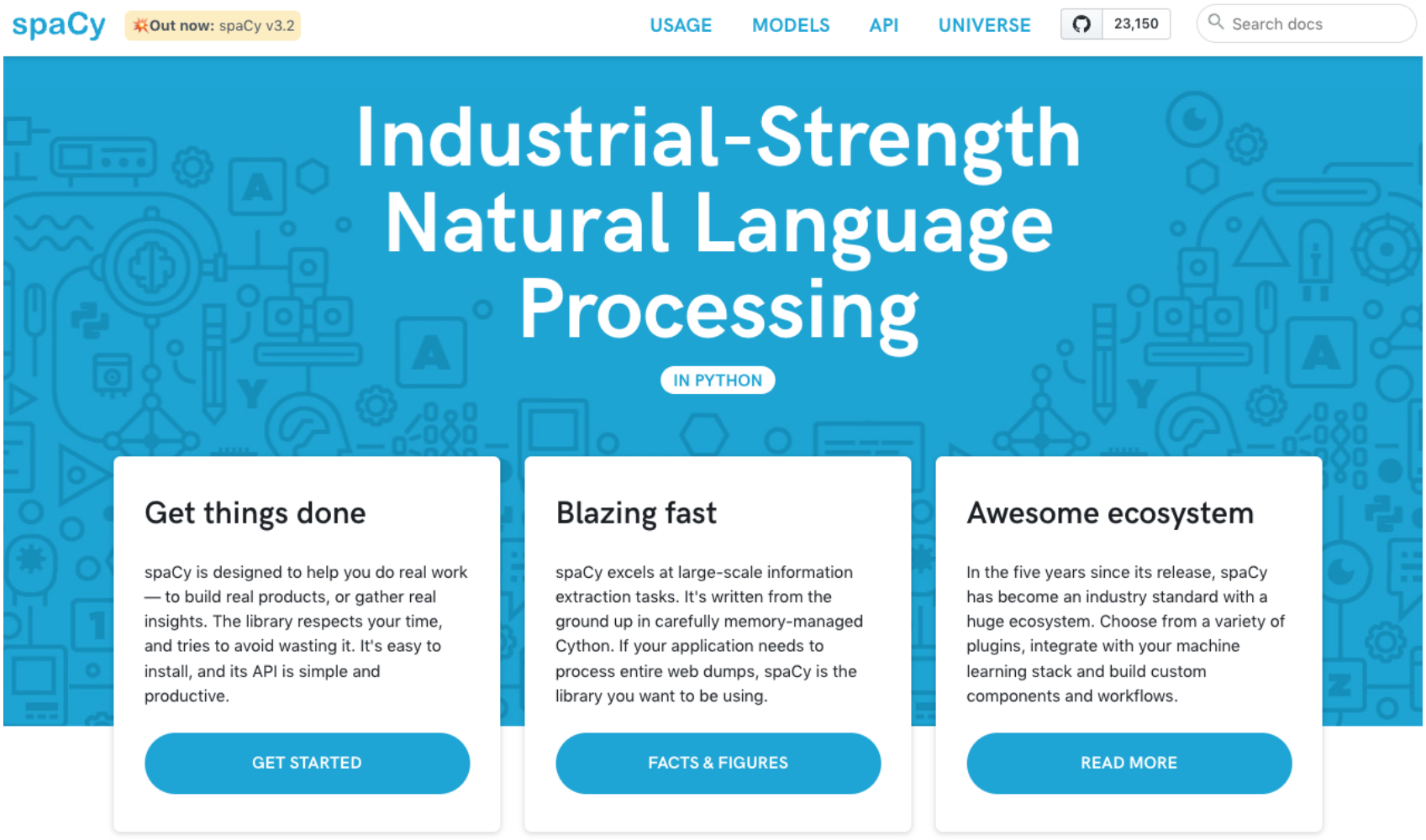
Features

- Relational: API | Tutorial
- Distribution: API | Tutorial
- Categorical: API | Tutorial
- Regression: API | Tutorial
- Multiples: API | Tutorial
- Style: API | Tutorial
- Color: API | Tutorial

Scikit-learn



SpaCy


spaCy
Out now: spaCy v3.2
USAGE
MODELS
API
UNIVERSE
23,150
Search docs
Industrial-Strength Natural Language Processing
IN PYTHON
Get things done
spaCy is designed to help you do real work — to build real products, or gather real insights. The library respects your time, and tries to avoid wasting it. It's easy to install, and its API is simple and productive.
GET STARTED
Blazing fast
spaCy excels at large-scale information extraction tasks. It's written from the ground up in carefully memory-managed Cython. If your application needs to process entire web dumps, spaCy is the library you want to be using.
FACTS & FIGURES
Awesome ecosystem
In the five years since its release, spaCy has become an industry standard with a huge ecosystem. Choose from a variety of plugins, integrate with your machine learning stack and build custom components and workflows.
READ MORE

NLTK

NLTK

Search

NLTK Documentation
API Reference
Example Usage
Module Index
Wiki
FAQ

Installation
Installing NLTK
Installing NLTK Data

More
Release Notes
Contributing to NLTK
NLTK Team

Documentation

Natural Language Toolkit

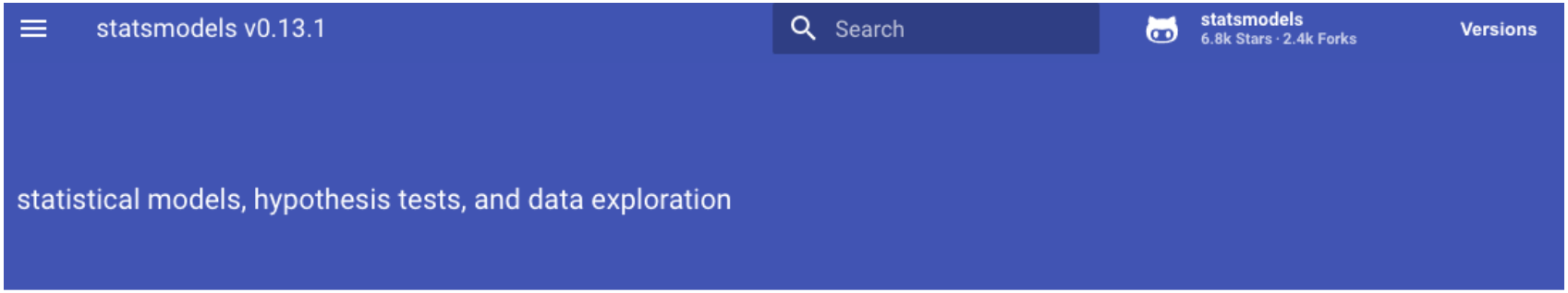
NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to **over 50 corpora and lexical resources** such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active **discussion forum**.

Thanks to a hands-on guide introducing programming fundamentals alongside topics in computational linguistics, plus comprehensive API documentation, NLTK is suitable for linguists, engineers, students, educators, researchers, and industry users alike. NLTK is available for Windows, Mac OS X, and Linux. Best of all, NLTK is a free, open source, community-driven project.

NLTK has been called “a wonderful tool for teaching, and working in, computational linguistics using Python,” and “an amazing library to play with natural language.”

Natural Language Processing with Python provides a practical introduction to programming for language processing. Written by the creators of NLTK, it guides the reader through the fundamentals of writing Python programs, working with corpora, categorizing text, analyzing linguistic structure, and more. The online version of the book has been been updated for Python 3 and NLTK 3. (The original Python 2 version is still available at **https://www.nltk.org/book_1ed**.)

Statsmodel



statsmodels

statsmodels is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration. An extensive list of result statistics are available for each estimator. The results are tested against existing statistical packages to ensure that they are correct. The package is released under the open source Modified BSD (3-clause) license. The online documentation is hosted at statsmodels.org.

Contents

Pingouin

pingouin 0.5.0 Functions Guidelines FAQ What's new Contribute

pypi package 0.5.0 | conda-forge v0.5.0 | license GPL-3.0 | Python tests passing | codecov 99%
downloads/month 259k | JOSS 10.21105/joss.01026 | chat on gitter

Pingouin is an open-source statistical package written in Python 3 and based mostly on Pandas and NumPy. Some of its main features are listed below. For a full list of available functions, please refer to the API documentation.

1. ANOVAs: N-ways, repeated measures, mixed, ancova
2. Pairwise post-hocs tests (parametric and non-parametric) and pairwise correlations
3. Robust, partial, distance and repeated measures correlations
4. Linear/logistic regression and mediation analysis
5. Bayes Factors
6. Multivariate tests
7. Reliability and consistency

BeautifulSoup

Beautiful Soup Documentation

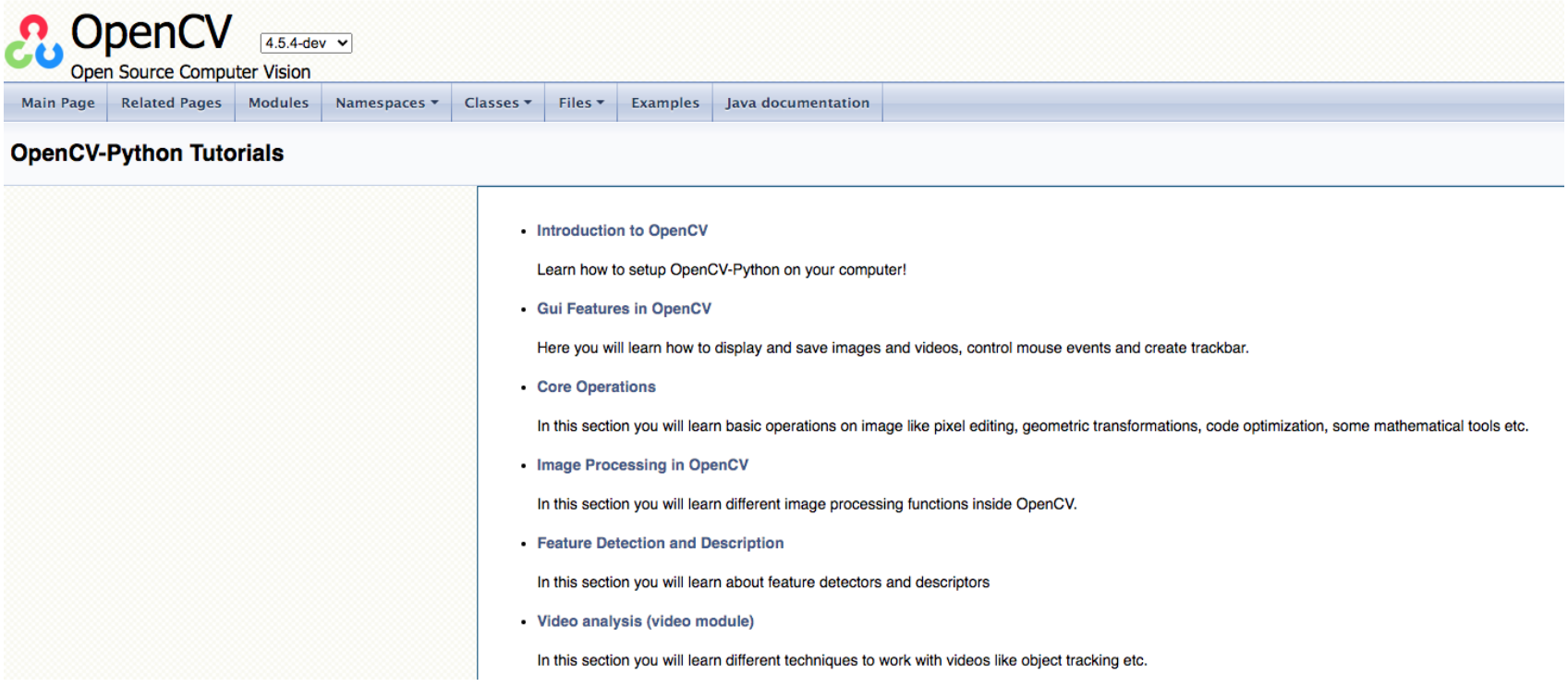
Beautiful Soup is a Python library for pulling data out of HTML and XML files. It works with your favorite parser to provide idiomatic ways of navigating, searching, and modifying the parse tree. It commonly saves programmers hours or days of work.

These instructions illustrate all major features of Beautiful Soup 4, with examples. I show you what the library is good for, how it works, how to use it, how to make it do what you want, and what to do when it violates your expectations.

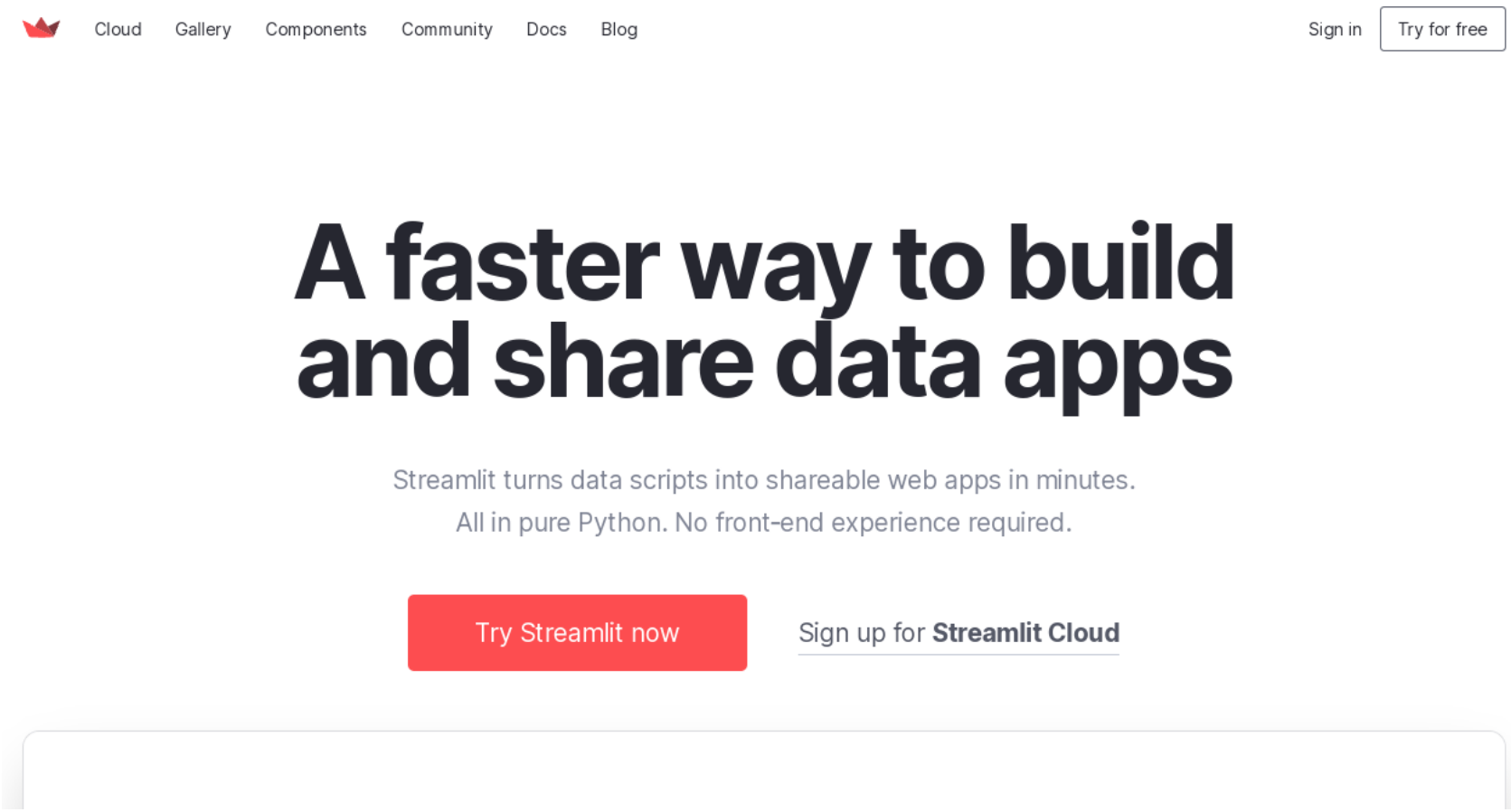
This document covers Beautiful Soup version 4.8.1. The examples in this documentation should work the same way in Python 2.7 and Python 3.2.

You might be looking for the documentation for Beautiful Soup 3. If so, you should know that Beautiful Soup 3 is no longer being developed and that support for it will be dropped on or after December 31, 2020. If you want to learn about the differences between Beautiful Soup 3 and Beautiful Soup 4, see Porting code to BS4.

OpenCV


OpenCV
4.5.4-dev
Open Source Computer Vision
Main Page
Related Pages
Modules
Namespaces
Classes
Files
Examples
Java documentation
OpenCV-Python Tutorials
Introduction to OpenCV
Learn how to setup OpenCV-Python on your computer!
Gui Features in OpenCV
Here you will learn how to display and save images and videos, control mouse events and create trackbar.
Core Operations
In this section you will learn basic operations on image like pixel editing, geometric transformations, code optimization, some mathematical tools etc.
Image Processing in OpenCV
In this section you will learn different image processing functions inside OpenCV.
Feature Detection and Description
In this section you will learn about feature detectors and descriptors
Video analysis (video module)
In this section you will learn different techniques to work with videos like object tracking etc.

Streamlit


Cloud
Gallery
Components
Community
Docs
Blog
Sign in
Try for free
A faster way to build and share data apps
Streamlit turns data scripts into shareable web apps in minutes.
All in pure Python. No front-end experience required.
Try Streamlit now
Sign up for Streamlit Cloud

À chaque fois :

- Identifier la bibliothèque pertinente (savoir antérieur ou recherche sur internet - attention, les choses peuvent changer)
- Lire la documentation / installer
- Tester sur un petit exemple
- Intégrer les fonctions dont on a besoin
- Éventuellement : développer de la virtuosité

Par exemple : on veut vérifier la présence d'un mot sur la page d'un quotidien ? Comment on fait ? (par ex : `https://www.journalducameroun.com/`)