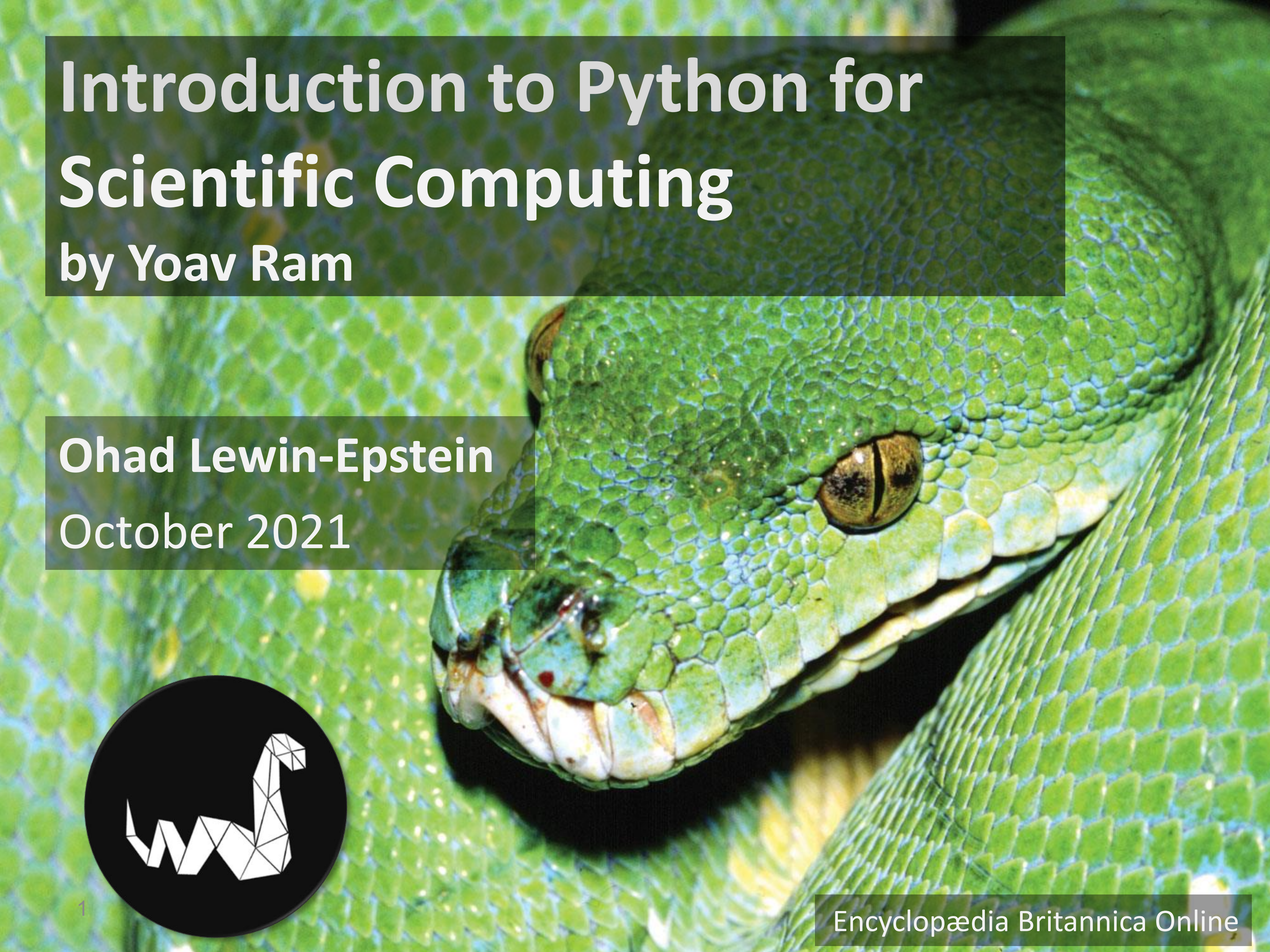
Introduction to Python for Scientific Computing

by Yoav Ram

Ohad Lewin-Epstein

October 2021



Encyclopædia Britannica Online

What is Python?

Python is a

- Widely used
- High-level
- General-purpose
- Interpreted
- Dynamic

Programming language

Wikipedia

Design emphasizes code readability

- Uncluttered **visual layout** (whitespaces…)
- **English keywords** used where other languages use punctuation (and, or, not…)
- Aims for **simplicity** and **generality**

Wikipedia

Example

C

```c
void foo(int x)
{
    if (-1 < x && x < 1) {
        bar();
        baz();
    } else {
        qux(x);
        foo(x - 1);
    }
}
```

Python

```python
def foo(x):
    if -1 < x < 1:
        bar()
        baz()
    else:
        qux(x)
        foo(x - 1)
```

Language features

- Interpreted language
- Dynamic type system (duck-typing)
- Automatic memory management (GC)
- Large and comprehensive standard library

Wikipedia

Multi-platform

- Interpreters available for many operating systems
- Code can be executed on a wide variety of systems
- Code can be packaged into stand-alone executable programs

Wikipedia

Python culture

- Free and open-source software
- Community-based development model
- Managed by the non-profit *Python Software Foundation* (PSF)

Wikipedia

History of Python

- Developed in 1989-91 by **Guido van Rossum** in the Netherlands
- Python 2.0 released Oct 2000
- Many major new features:
  - cycle-detecting garbage collector
  - support for Unicode
  - shift to transparent and community-backed development
- Python 3.0 released Dec 2008
  - major backwards-incompatible release
  - many of major features backported to Python 2.6 and 2.7
- **Python 3.7** released June 2018
- Python 3.8 alpha released; final version October 2019

Version history

Python 1.0 - January 1994

Python 1.5 - December 31, 1997

Python 1.6 - September 5, 2000

Python 2.0 - October 16, 2000

Python 2.1 - April 17, 2001

Python 2.2 - December 21, 2001

Python 2.3 - July 29, 2003

Python 2.4 - November 30, 2004

Python 2.5 - September 19, 2006

Python 2.6 - October 1, 2008

Python 2.7 - July 3, 2010

Python 3.0 - December 3, 2008

Python 3.1 - June 27, 2009

Python 3.2 - February 20, 2011

Python 3.3 - September 29, 2012

Python 3.4 - March 16, 2014

Python 3.5 - September 13, 2015

Python 3.6 - December 16, 2016

Python 3.7 – June 27, 2018

Python 3.8 – October, 2019

Guido van Rossum

- Python's principal author
- Still has a central role in deciding
  the direction of Python development
- Titled by the Python community: ***Benevolent Dictator for Life*** **(BDFL)**
- Employed by Google 2005-2012
- Spent half his time developing Python
- Since 2013 works for Dropbox
- Spends half his time developing Python…

Wikipedia

Homepage

Why Python?

Python is Free

Gratis: Free as in Beer

- MATLAB is **expensive**
  - Individuals: $2,350
  - Academia: $550
  - Personal: $95
  - Student: $29-55
  - Batteries (toolboxes…) not included
- Python is totally **free**
  - Batteries included (NumPy, SciPy…)

MathWorks Pricing

Libre: Free as in Speech

- MATLAB source code is **closed** and proprietary
  - You cannot **see** the code
  - You cannot **change** the code
  - You can participate in the discussion as a **client**
- Python source code is **open**
  - You can **see**, you can **change,** you can **contribute** code and documentation (python, numpy)
  - You can participate in the discussion as a **peer** (python, numpy)

Python is a general-purpose language

Python is used for:

- Scientific computing
- Enterprise software
- Web design
- Back-end
- Front-end
- Everything in between

Python is used at

Google, Rackspace, Microsoft, Intel, Walt Disney, MailChimp, twilio, Bank of America, Facebook, Instagram, HP, Linkedin, Elastic, Mozilla, YouTube, ILM, Thawte, CERN, Yahoo!, NASA, Trac, Civilization IV, reddit, LucasFilms, D-Link, Phillips, AstraZeneca, Applied Materials, KLA-Tencor, Nova, Lam Research, Marvell

https://us.pycon.org/2016/sponsors/
https://www.python.org/about/quotes/
https://en.wikipedia.org/wiki/Python_%28programming_language%29#Use
https://en.wikipedia.org/wiki/List_of_Python_software
https://www.python.org/about/success/

Python is portable

More or less same code runs on Windows, Linux, macOS, and any platform with a Python interpreter

Python for "other" platforms

Python syntax is beautiful

Python syntax is beautiful

Once you get over the use of **meaningful whitespace**, you realize how much it makes sense.

Famous entrepreneur and investor **Paul Graham**:

> ***You spend more time reading code than writing it***.
>
> *You push blobs of source code around the way a sculptor does blobs of clay.*
>
> *So a **language that makes source code ugly is maddening** to an exacting programmer, as clay full of lumps would be to a sculptor.*

Python is inherently object-oriented

Almost everything is an object

- strings, lists, dictionaries, tuples, functions, classes, and more
- The implied usefulness is that these things each have their own members and methods that encapsulate its functionality and information
- Strong **polymorphism**

Python is high level, easy to learn, and fast to develop

NumPy

Numba

Cython

Python is fast enough

Written in C (and some Fortran)

Easy to wrap more C

Easy to parallelize

Benchmark Game | NumFocus Benchmarks

Python is popular and has a great community

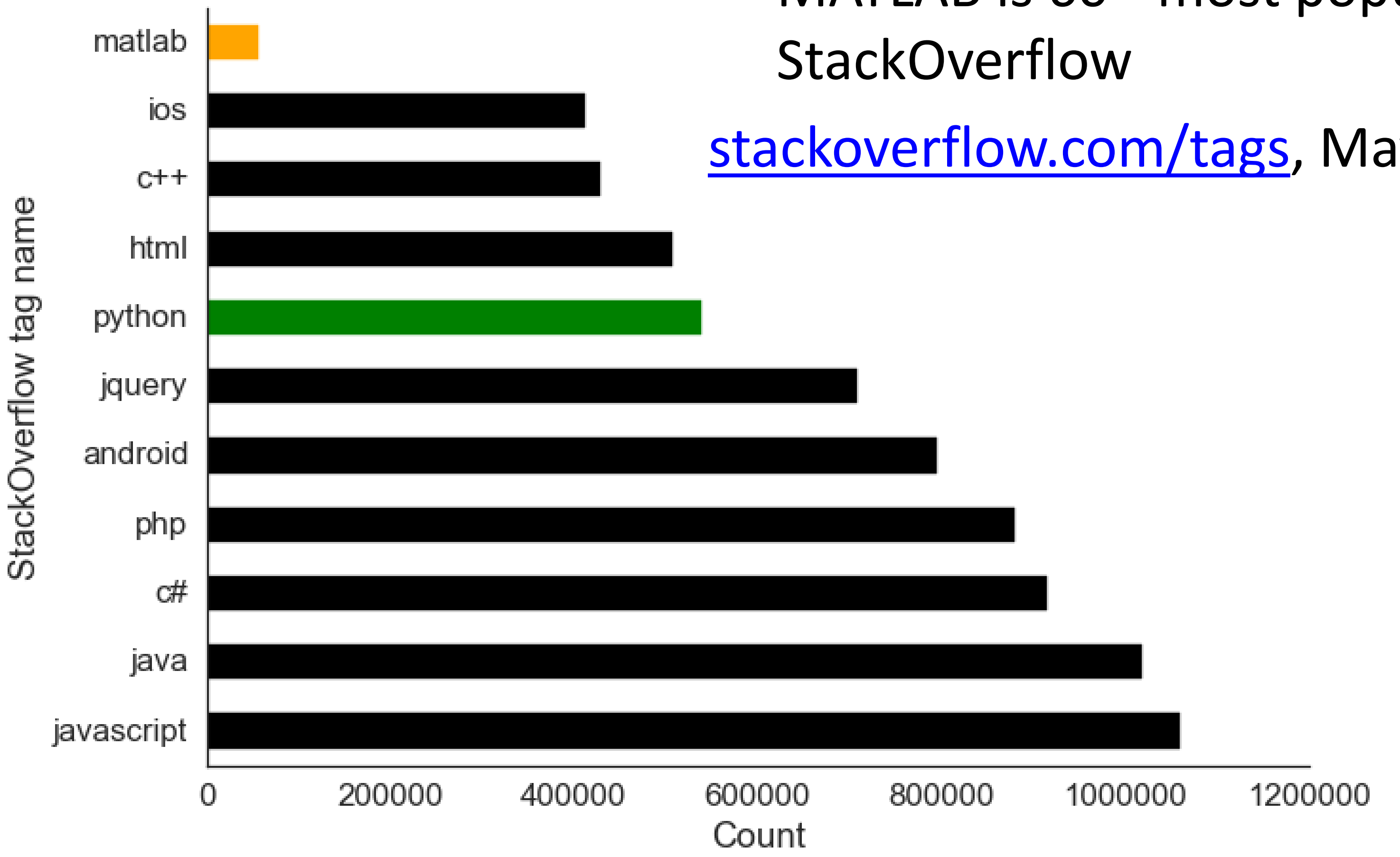
Great community

- Programmers
- Scientists
- Mathematicians
- Engineers

Easy to find help on the Internet

- Python is 6th most popular tag on StackOverflow
- MATLAB is 60th most popular tag on StackOverflow

stackoverflow.com/tags, May 2018

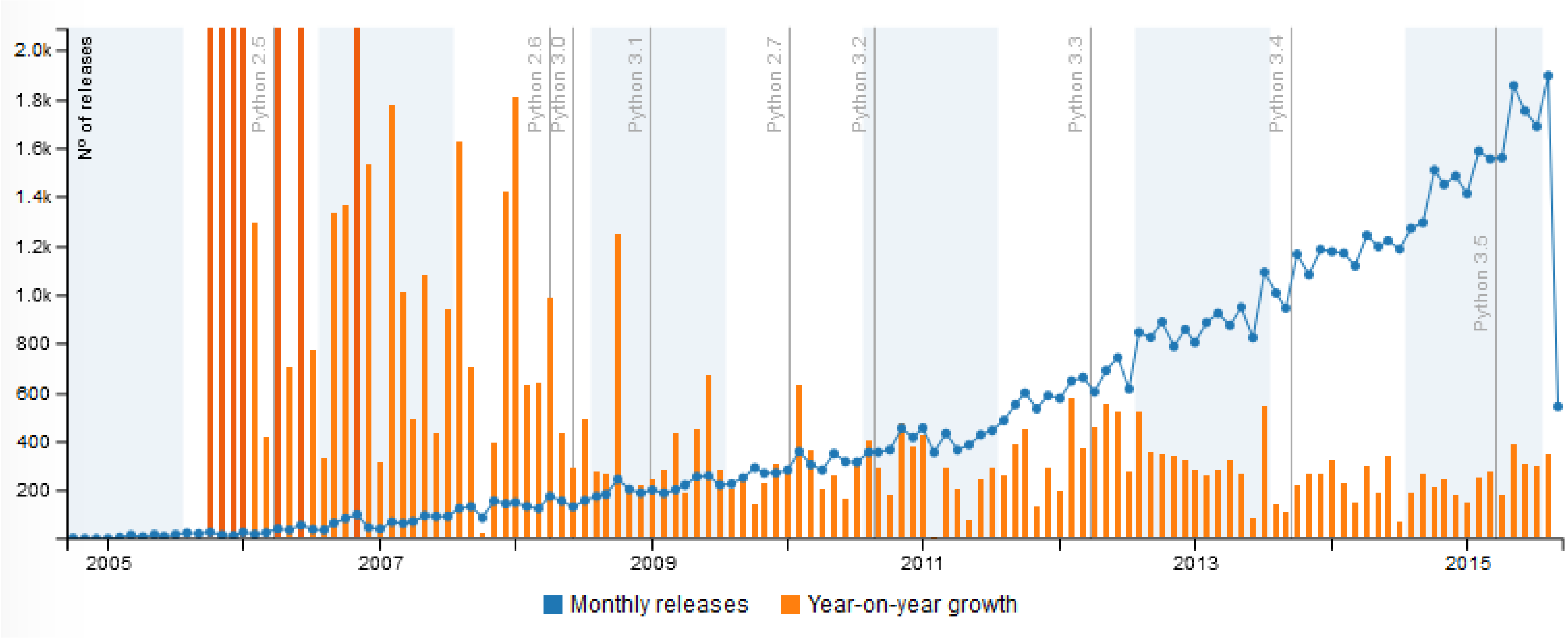


Active community

- 3rd most active repositories on *GitHub* after *Java* (incl. *Android*) and *JavaScript* (incl. *node.js*)
- ~27-fold more than MATLAB
- As of May 2018
- See breakdown at githut

Python has a lot of great libraries

Many new libraries released every month



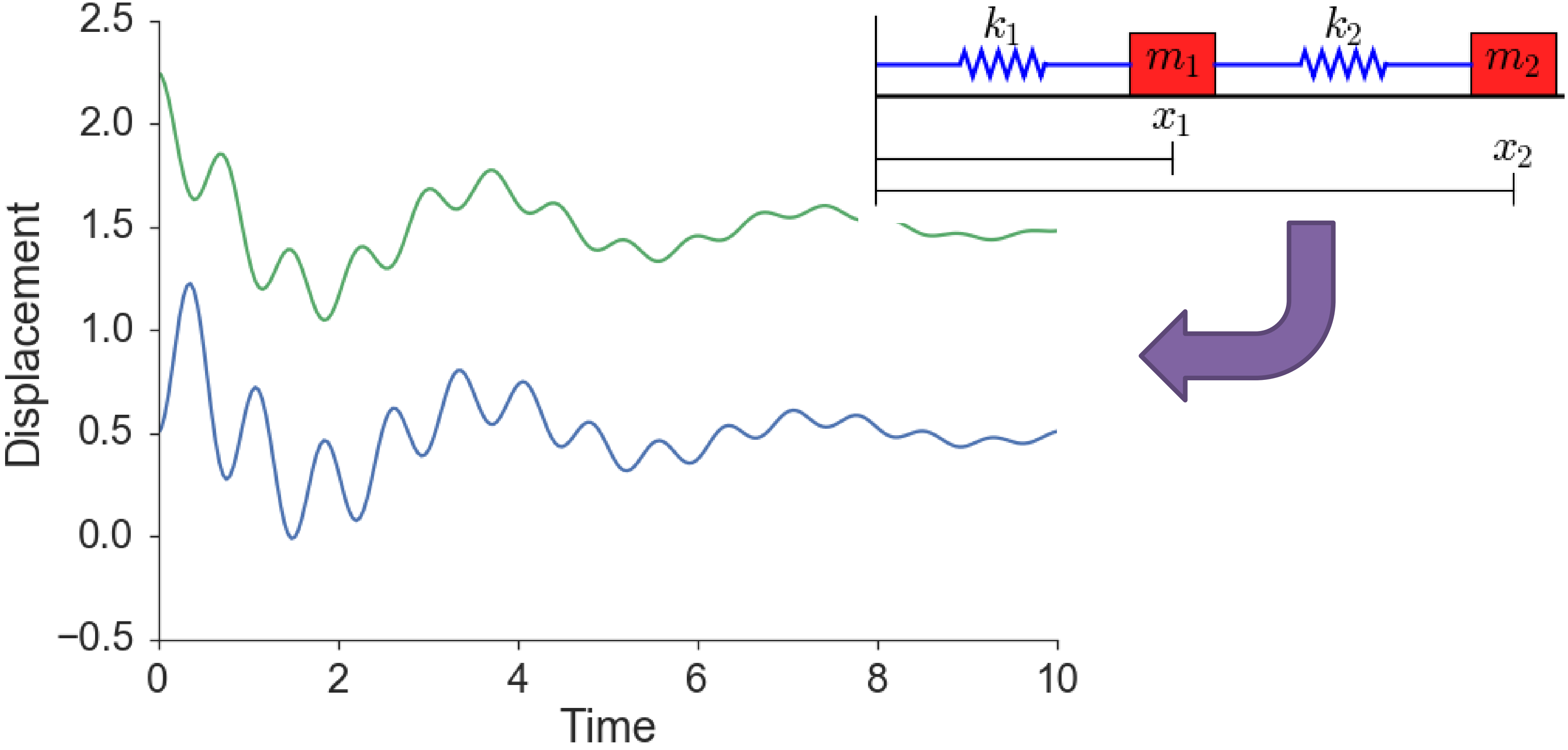
During 2015, > 1,500 new packages released every month to PyPI.
See more stats at PyGarden/stats.

Python can do nearly everything MATLAB can do

With libraries like NumPy, SciPy, Matplotlib, IPython/Jupyter,

Scikit-image, Scikit-learn, and more

Differential equations

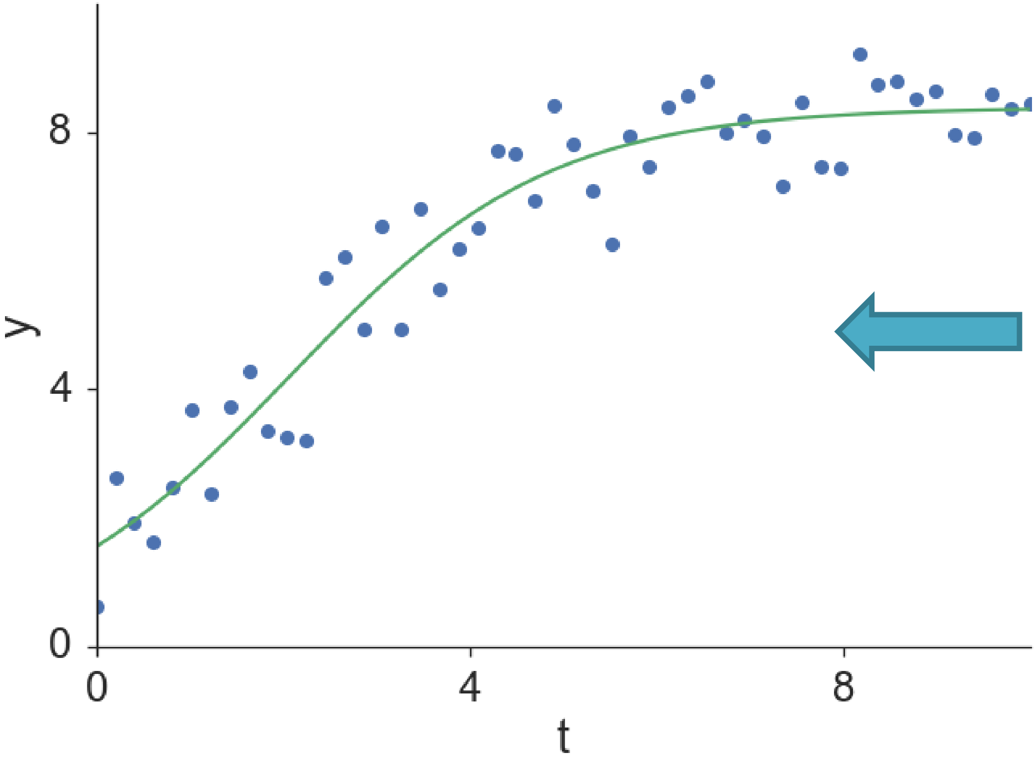


```
x = scipy.integrate.odeint(f, w0, t, …)
plot(t, x[:, 0])
plot(t, x[:, 1])
```

Curve fitting

```
params, cov = scipy.optimize.curve_fit(
        f=logistic, xdata=t, ydata=y, p0=(1, 10, 1))
```

N0=1.512, K=8.462, r=0.758



$$N(t) = \frac{K}{1 - \left(1 - \frac{K}{N(0)}\right)e^{-rt}}$$

Optimization

```
res = scipy.optimize.minimize_scalar(
          f, method="bounded", bounds=[8, 16])
```

```
fun: -0.23330441717143405
message: 'Solution found.'
nfev: 9
status: 0
success: True
x: 11.706005
```

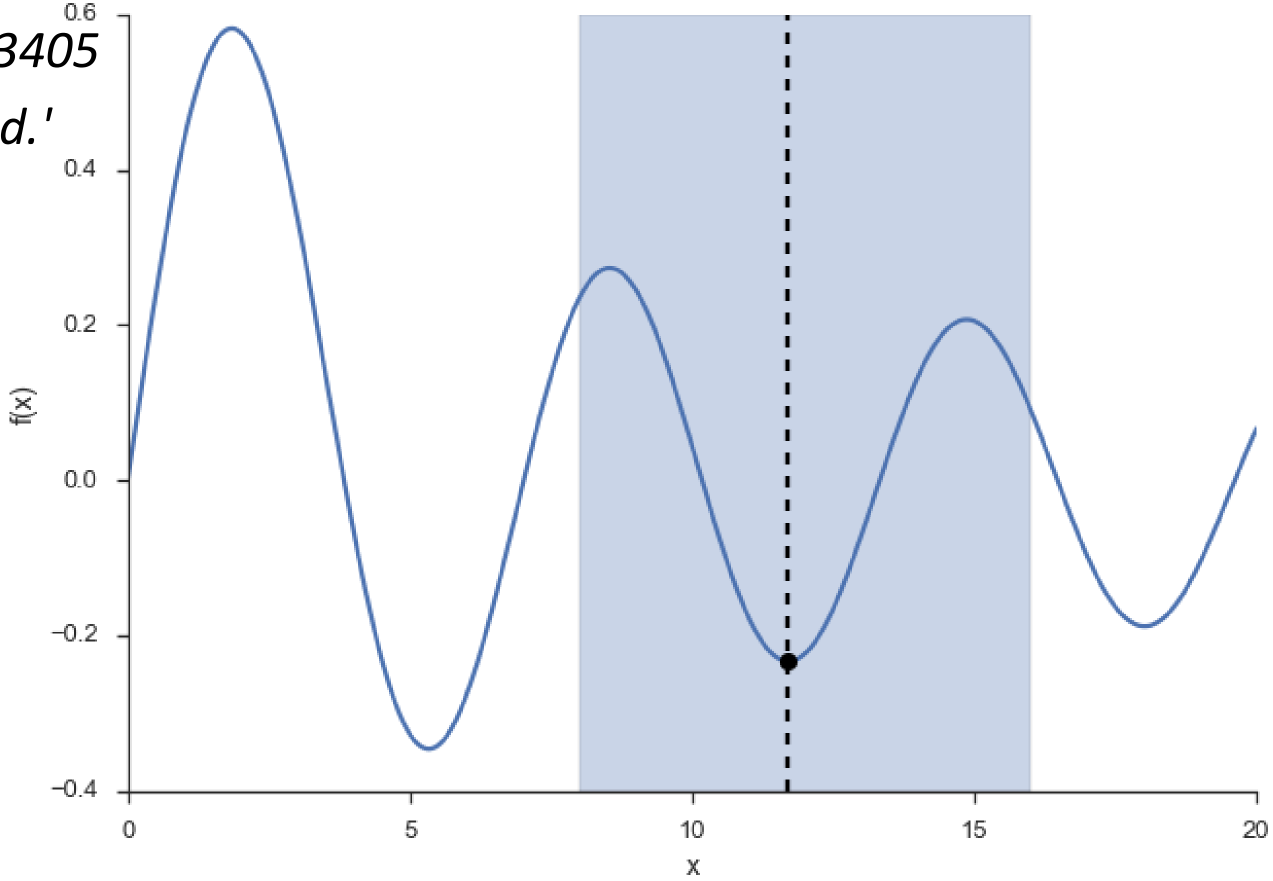
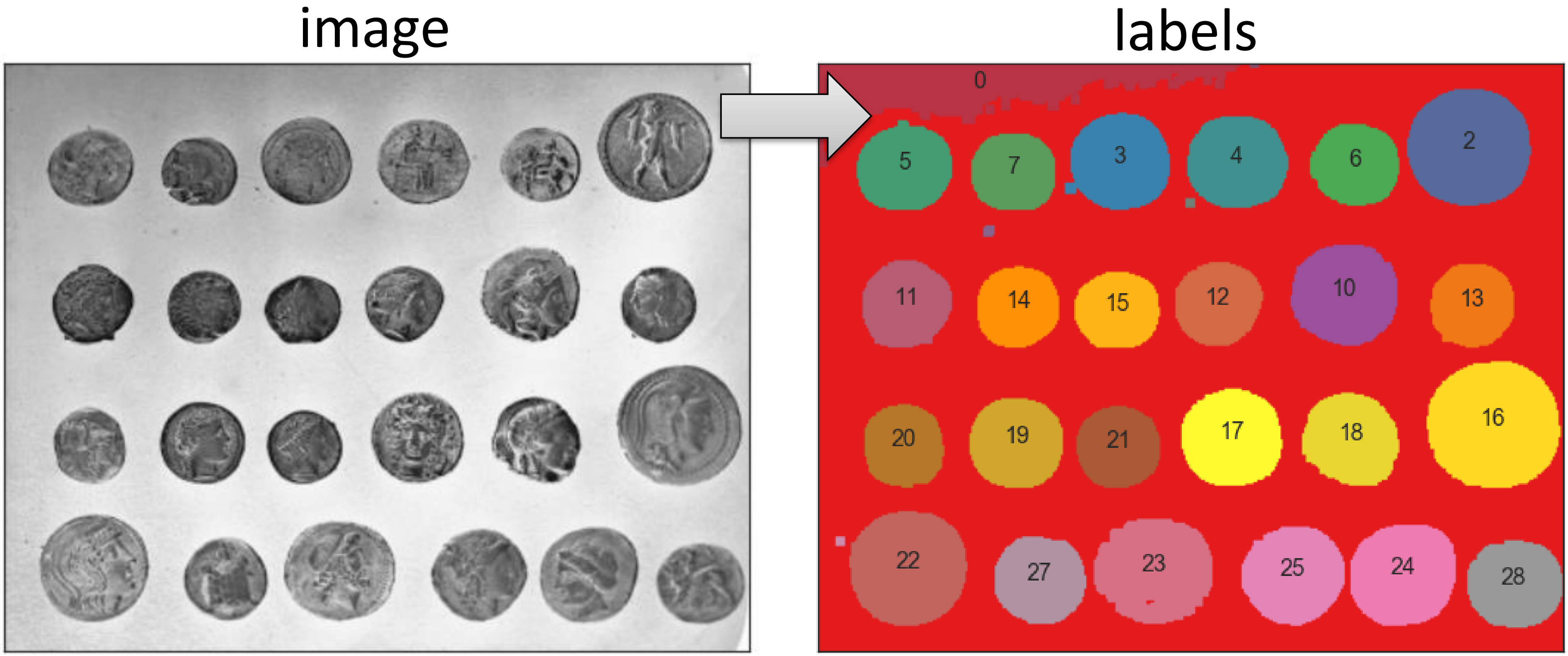


Image analysis

```
segmented = image > threshold
dilated = scipy.ndimage.generic_filter(segmented, max)
labels = skimage.measure.label(dilated)
```



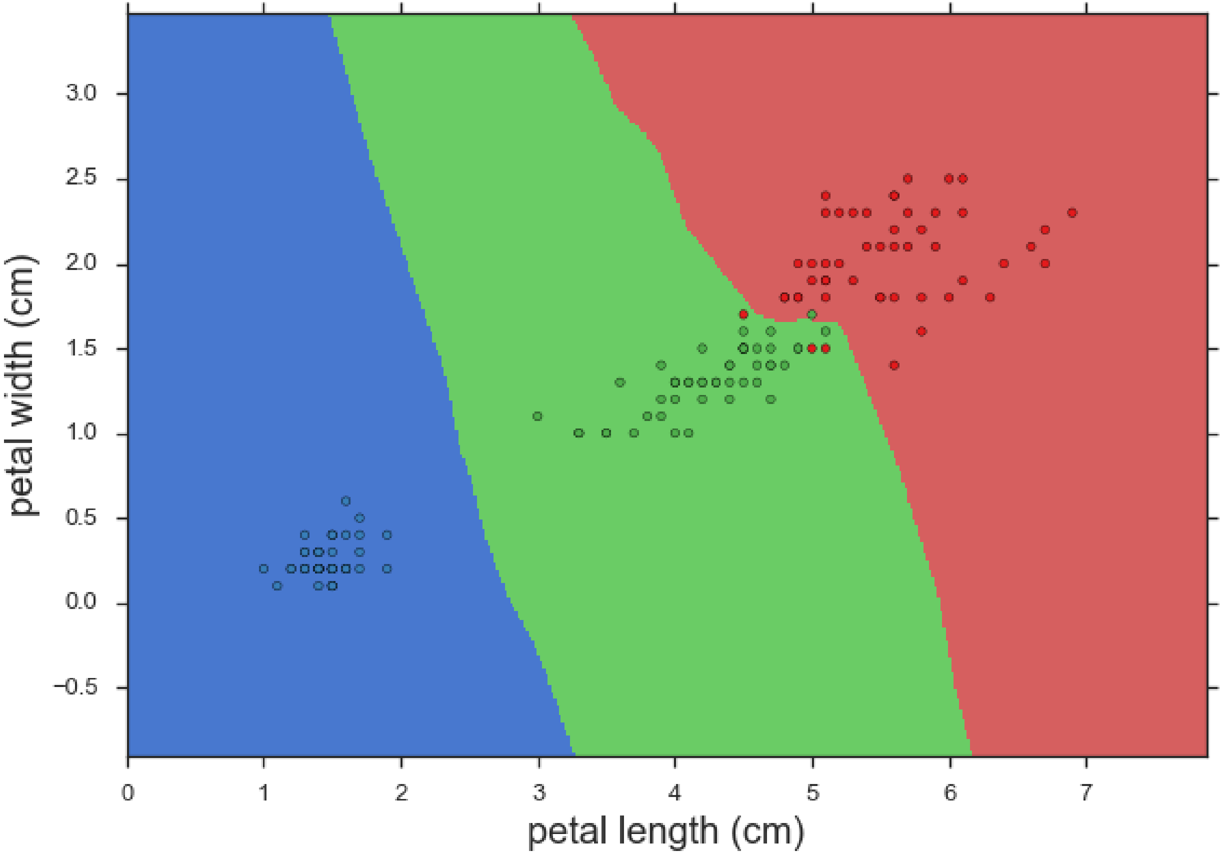
image

labels

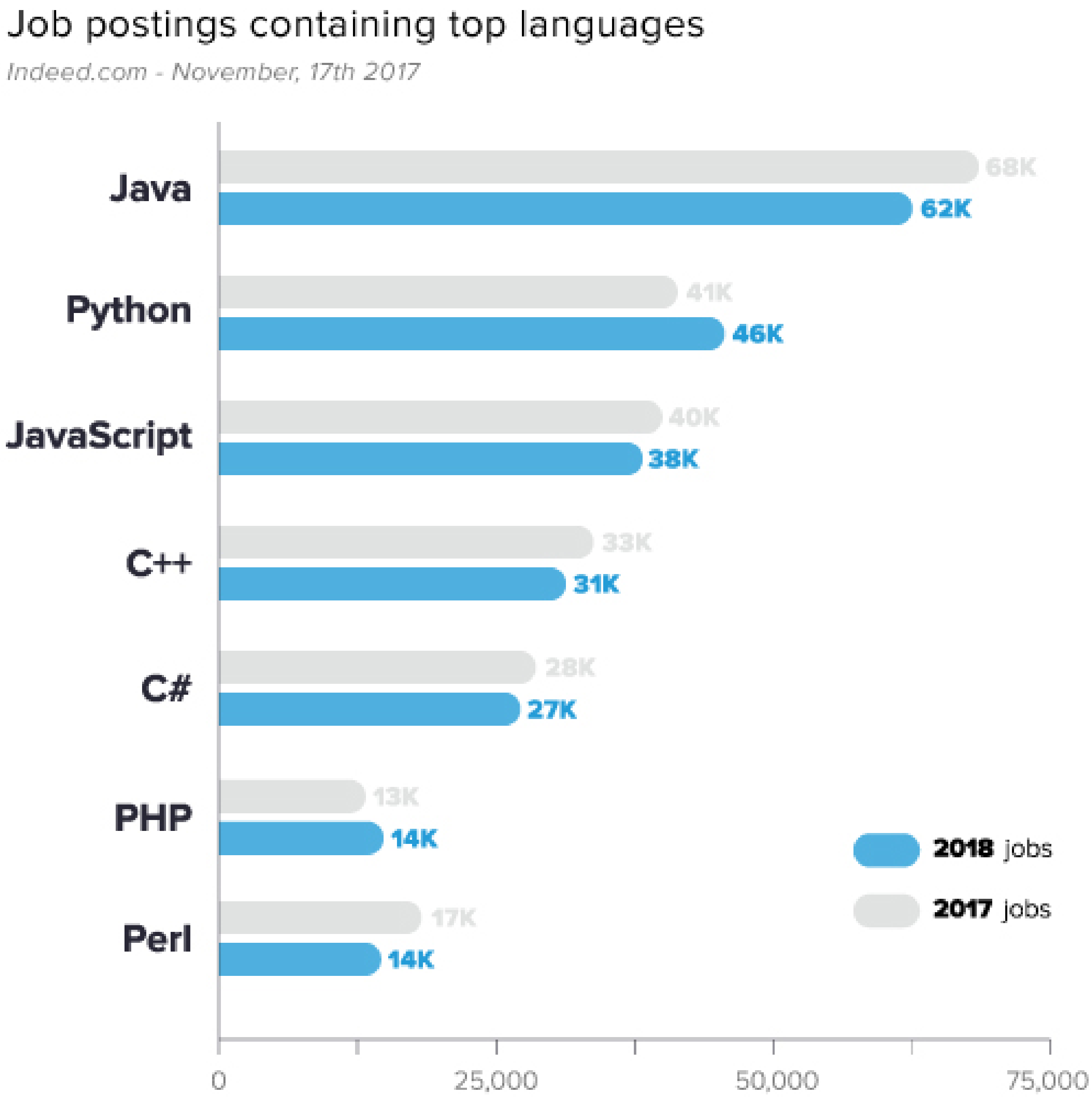
Machine learning

```
knn = sklearn.neighbors.KNeighborsClassifier()
knn.fit(X_train, y_train)
knn.predict(X_test)
```

Accuracy: 0.9



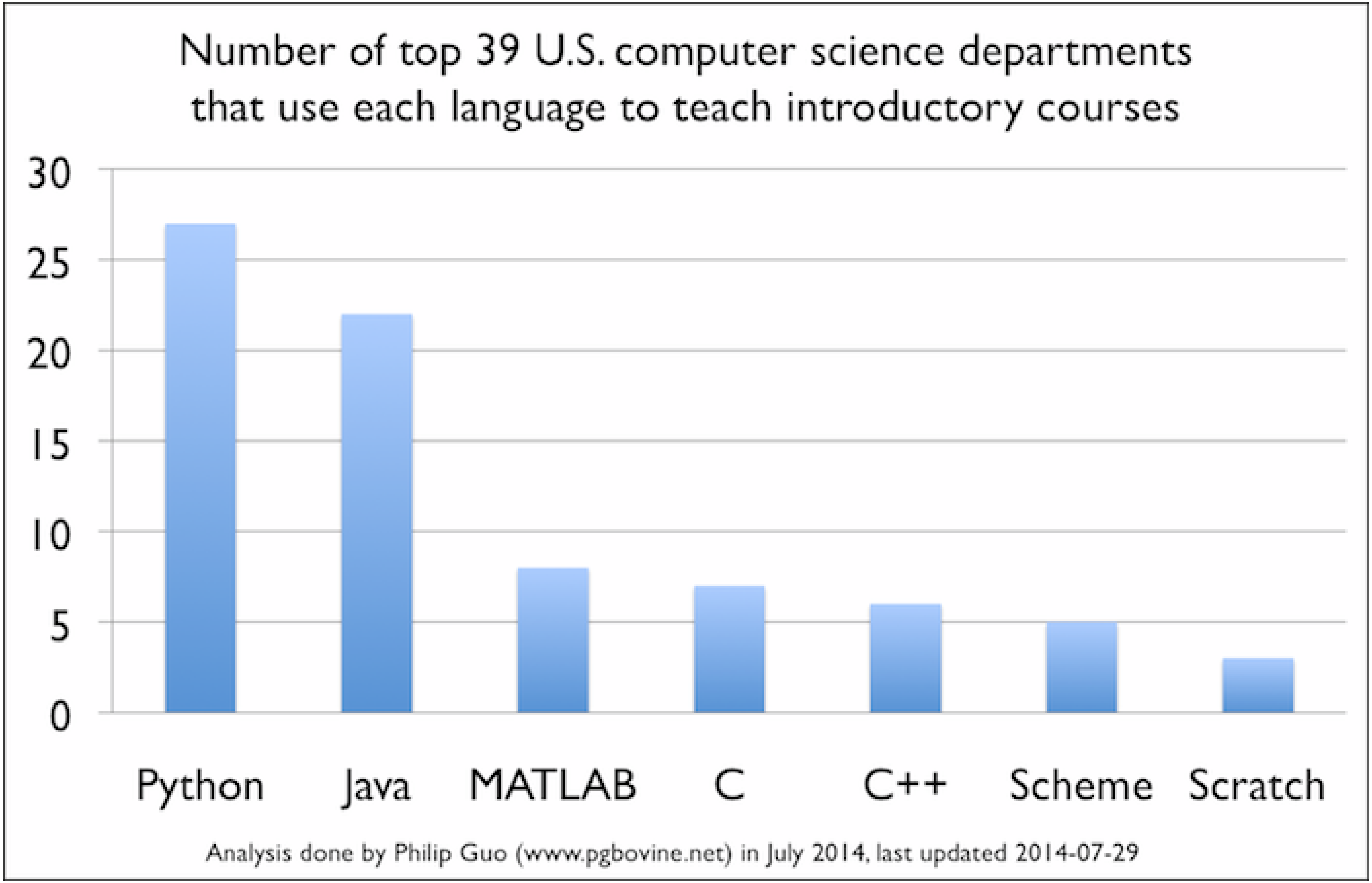
Demand & supply of Python programmers is high



Job postings containing top languages

Indeed.com - November, 17th 2017

Coding Dojo



Number of top 39 U.S. computer science departments that use each language to teach introductory courses

Analysis done by Philip Guo (www.pgbovine.net) in July 2014, last updated 2014-07-29

Phillip Gou @ CACM

Thanks for listening

YOAV RAM
PYTHON TRAINING

✉ ohadletaz@gmail.com

github.com/yoavram
⌂ python.yoavram.com

Death to the Stock P