Python Programming For Digital Forensics And Security Analysis

opensourceforu.com/2016/11/python-programming-digital-forensics-security-analysis

One of the many uses of the versatile Python programming language is in digital forensics and security analysis. This article covers various aspects like socket programming, port scanning, geo-location and extraction of data from websites like Twitter.

Python is one of the powerful programming languages used in key domains like cloud computing, Big Data analytics, network forensics, mobile app development, Web development and many others. Python has been in use for more than two decades.

Python code follows and provides support for multiple programming paradigms including imperative, functional, procedural and object oriented. Nowadays, Python is widely used for a variety of high performance computing applications by a number of corporate giants including Microsoft, Google, Red Hat, IBM, Amazon and many others. Python is free and open source, and delivers the implementations and interfaces for many other languages and platforms.

Table 1 displays a list of Python implementations, including the support for different platforms and programming models.

Installing Python

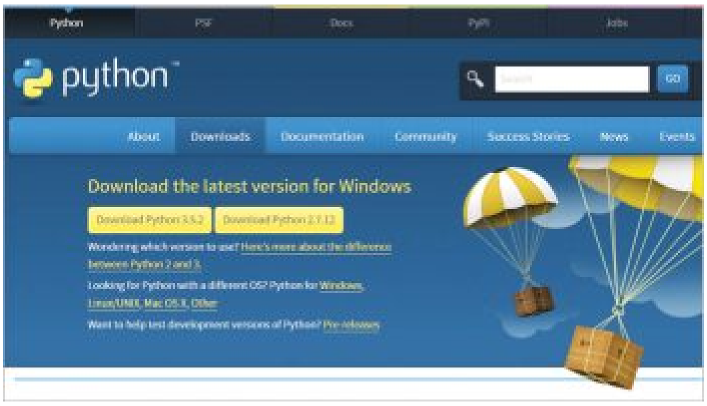
Python is available in two versions — Python 3.5 and Python 2.7 (*https://www.python.org*). Either of these can be downloaded, depending upon your requirements and type of application.

Python programming works with the IDE platform on which coding can be done. The command shell interface or IDE supports a Python program that is saved with a *.py* extension and executed at the command shell interface with the following commands:

```
$ python <filename.py> (For Linux)
DriveLetter:\(Path-To-Python)>python <filename.py> (For Windows)
```

Figure 2 depicts the execution of Python code on a system in which Python 2.7 is installed in Drive E: of the Windows OS.

IDE based programming with Python can include any IDE to write, debug and execute the code.



Download the latest version for Windows

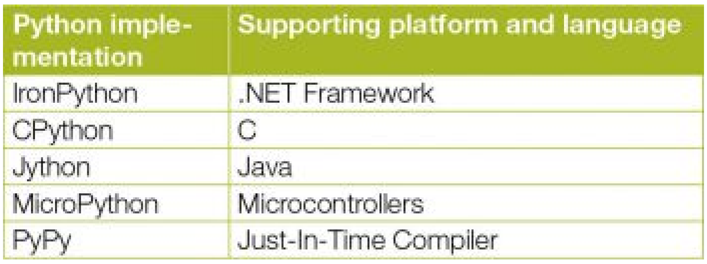
Download Python 3.5.2 | Download Python 2.7.12

Wondering which version to use? Here's more about the difference between Python 2 and 3.

Looking for Python with a different OS? Python for Windows, Linux/UNIX, Mac OS X, Other

Want to help test development versions of Python? Pre-releases

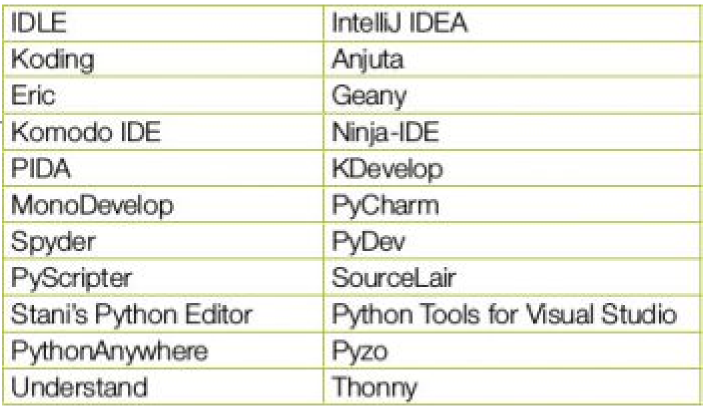
Figure 1: Python download page from the official portal



Python implementation	Supporting platform and language
IronPython	.NET Framework
CPython	C
Jython	Java
MicroPython	Microcontrollers
PyPy	Just-In-Time Compiler

Table 1

Given below is a list of Python IDEs where a graphical user interface is provided for easy programming:



IDLE	IntelliJ IDEA
Koding	Anjuta
Eric	Geany
Komodo IDE	Ninja-IDE
PIDA	KDevelop
MonoDevelop	PyCharm
Spyder	PyDev
PyScripter	SourceLair
Stani's Python Editor	Python Tools for Visual Studio
PythonAnywhere	Pyzo
Understand	Thonny

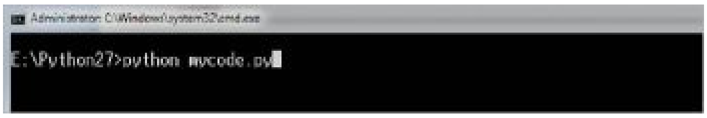
Digital forensics using Python programming

Whenever the topics of digital forensics, cyber security and penetration testing are discussed, professionals generally depend on a number of third party tools and operating systems. Kali Linux, MetaSploit, Parrot Security OS and many other tools are used for digital forensics. These tools come with in-built applications which the users deploy without real knowledge of the internal architecture and algorithmic approach of implementation.

Python is a widely used programming language for cyber security, penetration testing and digital forensic applications. Using the base programming of Python, any of the following can be performed without using any other third party tool:

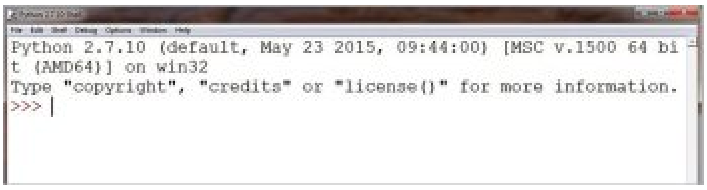
- Web server fingerprinting
- Simulation of attacks
- Port scanning
- Website cloning
- Load generation and testing of a website
- Creating intrusion detection and prevention systems
- Wireless network scanning
- Transmission of traffic in the network
- Accessing mail servers…

…and many other implementations related to digital fingerprinting and security applications



```
E:\Python27>python mycode.py
```

Figure 2: Executing Python code at the Windows command shell interface

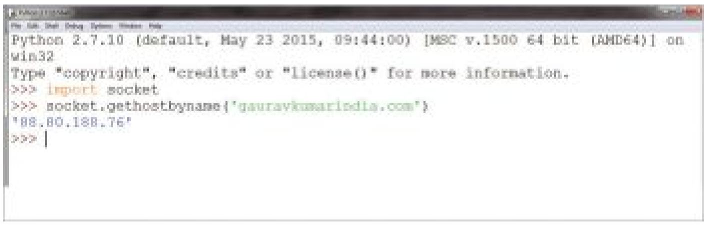


```
Python 2.7.10 (default, May 23 2015, 09:44:00) [MSC v.1500 64 bi
t (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>> |
```

Figure 3: Python IDLE environment

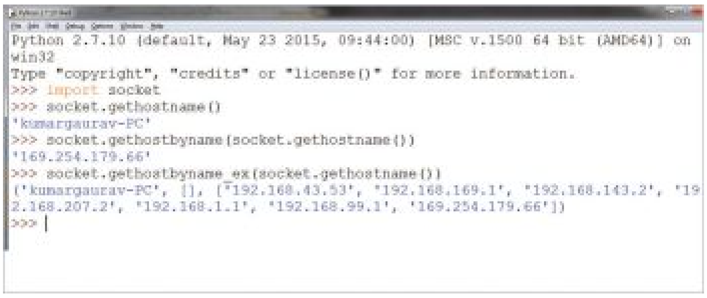
Socket programming

Socket programming is in-built with Python, similar to Java. To work with socket programming, the package socket is first imported and then the related methods can be called. Python installation comes with the in-built IDLE GUI.



```
Python 2.7.10 (default, May 23 2015, 09:44:00) [MSC v.1500 64 bit (AMD64)] on
win32
Type "copyright", "credits" or "license()" for more information.
>>> import socket
>>> socket.gethostbyname('gauravkumarindia.com')
'88.80.188.76'
>>> |
```

Figure 4: Fetching an IP address of a website from a URL



```
Python 2.7.10 (default, May 23 2015, 09:44:00) [MSC v.1500 64 bit (AMD64)] on
win32
Type "copyright", "credits" or "license()" for more information.
>>> import socket
>>> socket.gethostname()
'kumargaurav-PC'
>>> socket.gethostbyname(socket.gethostname())
'169.254.179.66'
>>> socket.gethostbyname_ex(socket.gethostname())
('kumargaurav-PC', [], ['192.168.43.53', '192.168.169.1', '192.168.143.2', '19
2.168.207.2', '192.168.1.1', '192.168.99.1', '169.254.179.66'])
>>> |
```

Figure 5: Fetching IP addresses associated with the local system

Network port scanning

Generally, the *nmap* tool is used for the implementation of network port scanning, but using Python socket programming, it can be implemented without any third party tool. In Kali Linux, there are many tools available for digital forensics related to networks, but many of these implementations can be done using Python programming with just a few lines of instruction.

The code for port scanning of any IP address can be downloaded from here. The code checks which particular ports are open from the PortList [20, 22, 23, 80, 135, 445, 912]. Each value in the PortList specifies a particular service associated with the network.

Geolocation extraction

The real-time location of an IP address can be extracted using Python and Google APIs with the use of the *pygeoip* module. First of all, import the GeoIP database from the URL.

Once the database is loaded and mapped with the Python installation, any IP address can be scanned with global visibility and location.

```
>>> import pygeoip
>>> myGeoIP = pygeoip.GeoIP(‘GeoIPDataSet.dat’)
>>> myGeoIP.country_name_by_addr(‘<IP Address>’)
 ‘United States‘
```

To look up the country, use the following commands:

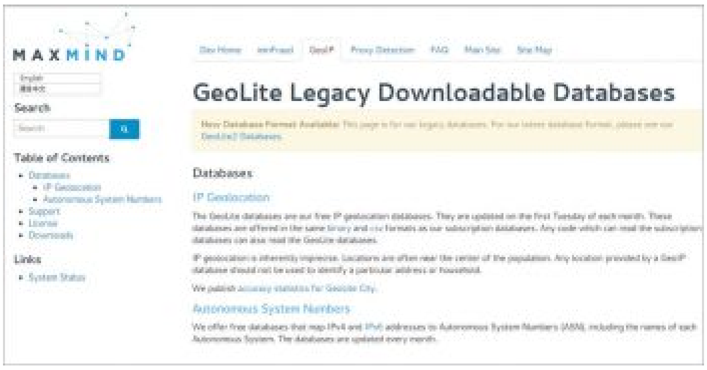
```
>>> myGeoIP = pygeoip.GeoIP('GeoIPDataSet.dat')
>>> myGeoIP.country_code_by_name('google.com')
'US'
>>> myGeoIP.country_code_by_addr('<IP Address>')
'US'
>>> myGeoIP.country_name_by_addr('<IP Address>')
'United States'
```

To look up the city, use the following commands:

```
>>> myGeoIP = pygeoip.GeoIP('GeoIPCity.dat')
>>> myGeoIP.record_by_addr('<IP Address>')
{
    'city': u'Mountain View',
    'region_code': u'CA',
    'area_code': 550,
    'time_zone': 'America/Los_Angeles',
    'dma_code': 807,
    'metro_code': 'San Francisco, CA',
    'country_code3': 'USA',
    'latitude': 38.888222,
    'postal_code': u'94043',
    'longitude': -123.37383,
    'country_code': 'US',
    'country_name': 'United States',
    'continent': 'NA'
}
>>> myGeoIP.time_zone_by_addr('<IP Address>')
'America/Los_Angeles'
```

Real-time extraction from social media

The live and real-time data from social media platforms can be downloaded using Python scripts. In Python, there are many modules and extensions with which the interfacing with WhatsApp, Twitter, Facebook, LinkedIn and many other platforms can be done.



GeoLite Legacy Downloadable Databases

New Database Format Available: This page is for our legacy databases. For our latest database format, please see our GeoLite2 Databases.

Databases

IP Geolocation

The GeoLite databases are our free IP geolocation databases. They are updated on the first Tuesday of each month. These databases are offered in the same binary and csv formats as our subscription databases. Any code which can read the subscription databases can also read the GeoLite databases.

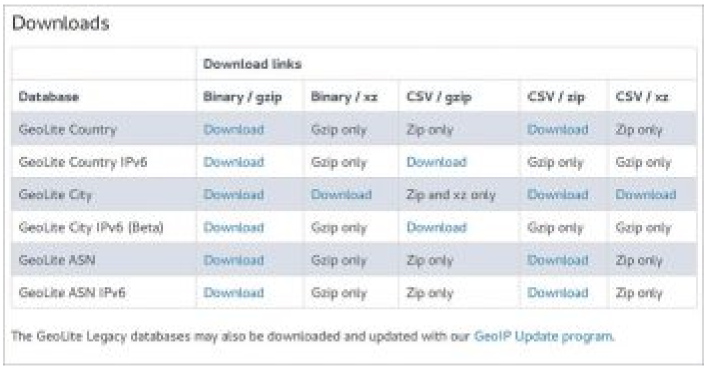
IP geolocation is inherently imprecise. Locations are often near the center of the population. Any location provided by a GeoIP database should not be used to identify a particular address or household.

We publish accuracy statistics for GeoLite City.

Autonomous System Numbers

We offer free databases that map IPv4 and IPv6 addresses to Autonomous System Numbers (ASN), including the names of each Autonomous System. The databases are updated every month.

Figure 6: Downloadable databases for GeoIP mapping



Downloads

	Download links				
Database	Binary / gzip	Binary / xz	CSV / gzip	CSV / zip	CSV / xz
GeoLite Country	Download	Gzip only	Zip only	Download	Zip only
GeoLite Country IPv6	Download	Gzip only	Download	Gzip only	Gzip only
GeoLite City	Download	Download	Zip and xz only	Download	Download
GeoLite City IPv6 (Beta)	Download	Gzip only	Download	Gzip only	Gzip only
GeoLite ASN	Download	Gzip only	Zip only	Download	Zip only
GeoLite ASN IPv6	Download	Gzip only	Zip only	Download	Zip only

The GeoLite Legacy databases may also be downloaded and updated with our GeoIP Update program.

Figure 7: Download links for GeoIP databases with IPv6 compatibility

Python package index (PyPI)

PyPI (*https://pypi.python.org*) is the software repository of enormous Python packages for interfacing with other platforms. PyPI is freely available for Python developers without any licensing or subscriptions.