



PROGRAMACIÓN EN PYTHON

YOLANDA GARCÍA RUIZ

# Introducción

---

## El lenguaje Python

Python es uno de los lenguajes líderes a nivel mundial ([tiobe-index](#))

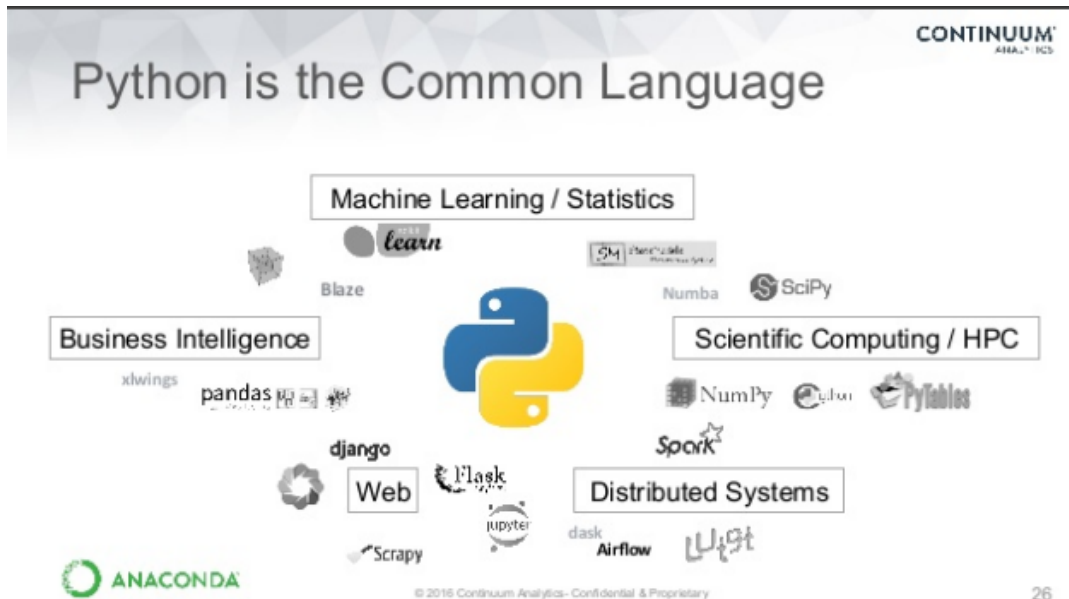
Es fácil de aprender y de leer

- Es de manejo sencillo
- Curva de aprendizaje es muy corta

Python es una herramienta poderosa en el análisis numérico, el cual es ampliamente utilizado en ciencia e ingeniería.

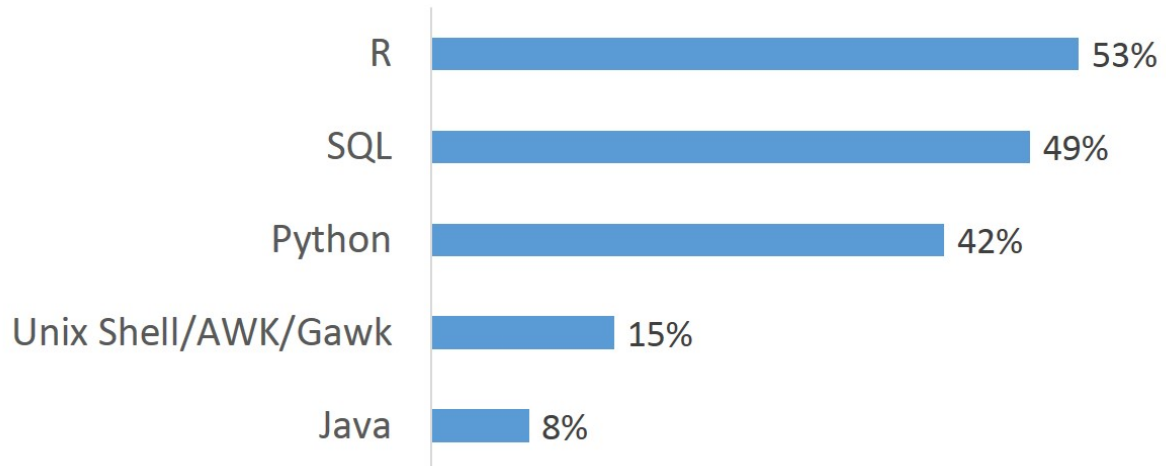
- Posee una gran eficiencia a la hora de analizar datos
  - Se utiliza en casi todas las API's de las aplicaciones relacionadas con el dato
-

Su creador fue Guido Van Rossum (1991) y debido a que es de código libre la comunidad de desarrolladores han creado librerías o módulos para hacer casi cualquier cosa.



- Se trata de un lenguaje que permite desarrollar un proyecto completo o puede ser parte de un sistema.
- Se recomienda aprender a programar Python, por cosas muy sencillas:
  - Es fácil de aprender, fácil de leer y cuenta con una gran cantidad de librerías, tipos de datos y funciones incorporadas en el propio lenguaje, que ayudan a realizar muchas tareas habituales sin necesidad de tener que programarlas desde cero.
  - La curva de aprendizaje es muy corta y en muy poco tiempo uno puede estar programando cosas no tan triviales.
- Aunque desconocido aún por algunos, hoy en día es uno de los lenguajes más populares junto con R, Ruby, Perl y otros. A Python todavía le falta algo de la riqueza de otros lenguajes como R para el análisis de datos, pero se está acercando rápidamente.

## Top 5 of programming languages mostly used for data science's activities



Fuente: <http://www.datasciencecentral.com/profiles/blogs/top-programming-languages-for-data-science-first-partial-survey>.

## Características del lenguaje

**Lenguaje interpretado:** Interpretado quiere decir que no se debe compilar el código antes de su ejecución.

- Esto implica que la mayor parte del código Python se ejecuta más lento que el código compilado de C++ o Java.
- En casos donde se necesita mucha rapidez de proceso es preferible asumir el coste en tiempo de programación en un lenguaje de bajo nivel como C++ para conseguir mayor rapidez de proceso.

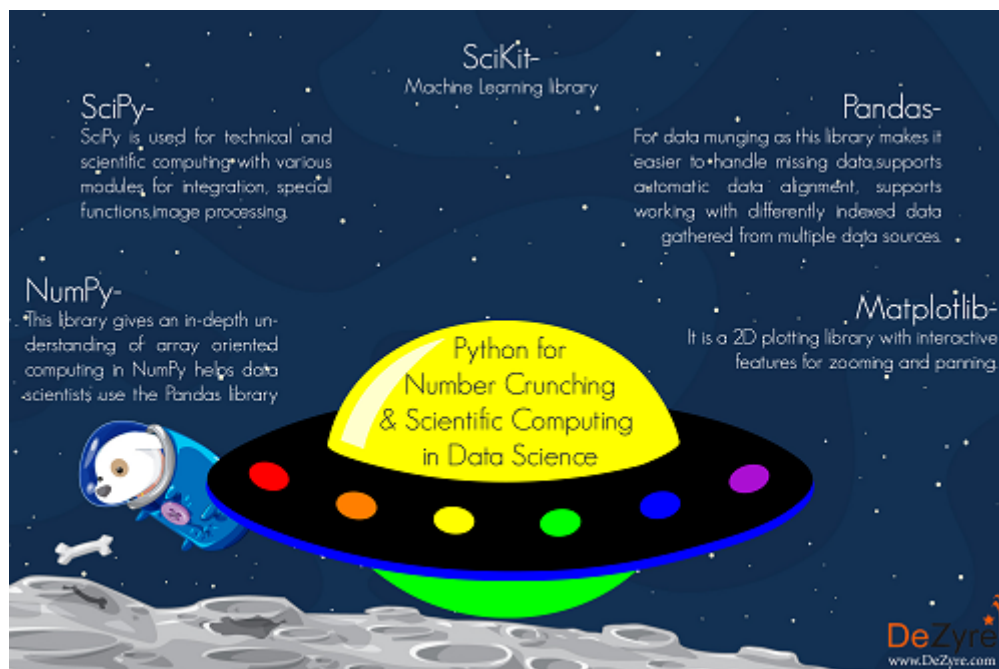
**Multiparadigma:** Es un lenguaje imperativo y orientado a objetos. También posee características de los **lenguajes funcionales**.

**Multiplataforma:** Python es un lenguaje disponible en los principales Sistemas Operativos (Windows, Linux y Mac). El lenguaje permite realizar cualquier tipo de programa, desde aplicaciones Windows a servidores de red o incluso, páginas web.

**Tipado dinámico:** El tipo de los datos es inferido en tiempo de ejecución. Esta es una característica que diferencia a Python de la mayoría de los lenguajes compilados.

## Librerías básicas para el tratamiento de datos

Python tiene disponibles un gran número de librerías y herramientas para el tratamiento de datos.



Entre las librerías denominadas científicas tenemos:

### \* **NumPy (Numerical Python)**

- Es el pilar fundamental de todo el ecosistema numérico en Python.
- Proporciona un objeto de tipo array multidimensional: ndarray
- ndarray permite el almacenamiento de datos homogéneos y heterogéneos.
- Proporciona funciones y métodos para operar eficientemente con este objeto.
- Proporciona herramientas para la escritura y lectura de datos a disco.
- Además: operaciones de álgebra lineal, números aleatorios, funciones financieras, etc.

### **Desventajas:**

- Pensado para conjuntos de datos de tamaño fijo
- No optimizado para datos heterogéneos (números, texto, fechas)
- Manejo de datos textuales engorroso
- Solución: pandas

### \* **pandas (panel data sets)**

- Se trata de una librería construida sobre NumPy.
- Proporciona estructuras de datos de alto nivel y optimizada para análisis de datos: DataFrame.
- Se trata de una estructura que permite el almacenamiento de datos que en su origen se representan de forma tabular (como por ejemplo: hojas de cálculo).
- Incorpora herramientas para leer y escribir datos en diversos formatos: CSV y texto, Excel, bases de datos SQL, documentos XML y JSON, etc.
- Manejo de series temporales.
- Está integrada con otras librerías:
  - Matplotlib y Seaborn : para la representación de los datos de forma gráfica

### \* **SciPy**

- Colección de sublibrerías y funciones que implementan funcionalidad estándar útil en ciencias y finanzas.
- Interpolación e integración numérica.
- Regresión.

### \* **Scikit-learn**

- Para aprendizaje automático.

---

Content on this site is licensed under a Creative Commons Attribution 4.0 International license.

