

Bibliotecas en Python

La popularidad de Python en la ciencia y la industria

Python es uno de los lenguajes de programación más importantes en la actualidad, ya sea para la ciencia de datos, el aprendizaje automático o el desarrollo de software en general; tanto en el ámbito académico como en el industrial. Gigantes tecnológicos como Google, Apple, NASA, Instagram, Pixar y otros, utilizan Python de manera intensiva.



Razones de su popularidad:

The diagram consists of two side-by-side callout boxes. The left box contains an icon of a folder with code snippets and the text: "Su gama de **bibliotecas**; esto, porque posee una nutrida colección de código que puede utilizarse en otros programas para realizar tareas reales." The right box contains an icon of a brain with code snippets and the text: "Su **sintaxis** concisa y sencilla. Las reglas de sintaxis definen las palabras clave, los símbolos y el formato utilizados en los programas. En comparación con otros lenguajes de programación, Python es más conciso y directo."

Bibliotecas para el análisis de datos

Las bibliotecas de código abierto mejoradas de Python, lo han convertido en una opción popular para tareas de análisis de datos. Combinado con la solidez general de Python para la ingeniería de software de propósito general, es una excelente opción como lenguaje principal para el desarrollo de aplicaciones de datos.

NumPy

- NumPy, abreviatura de Numerical Python, ha sido durante mucho tiempo un pilar del cálculo numérico en Python.
- Proporciona las estructuras de datos, los algoritmos y la biblioteca de integración necesarias para la mayoría de las aplicaciones científicas que involucran datos numéricos en Python.



Pandas

- Proporciona estructuras de datos de alto nivel y funciones diseñadas para que, al trabajar con datos estructurados o tabulares, su uso sea intuitivo y flexible.
- Combina las ideas de computación de matrices de NumPy con las capacidades de manipulación de datos que se encuentran en hojas de cálculo y bases de datos relacionales (como SQL).



Matplotlib

- Es la biblioteca de Python más popular para generar gráficos y otras visualizaciones de datos bidimensionales, por ejemplo, gráficos de líneas, de barras, de dispersión, circulares e histogramas.
- Se recomienda utilizarla cuando se necesite el control total sobre cada aspecto del gráfico.



Seaborn

- Es una biblioteca de gráficos estadísticos basada en Matplotlib. Seaborn simplifica la creación de muchos tipos de visualización comunes, además de que permite crear gráficos estadísticos más estéticos y complejos, con menos código.
- Es adecuada para crear rápidamente visualizaciones estadísticas atractivas, por ejemplo: gráficos de violín y mapas de calor, con datos de Pandas.



Referencias:

- McKinney, W. (2022). *Python for Data Analysis* (Third edition). O'Reilly Media, Inc.
- Das, U., Lawson, A., Mayfield, C., Norouzi, N. (2024). *Introduction to Python Programming*. OpenStax.

Elaboró contenido: Dra. Lisbeth Rodríguez Mazahua

