



LIBRERÍAS DE **PYTHON** PARA EL ANÁLISIS DE DATOS



¿QUÉ SON LAS LIBRERÍAS DE PYTHON?

Python es ampliamente utilizado en la comunidad de ciencia de datos. Las **librerías de Python** son **esenciales para el análisis de datos**, ya que **proporcionan** un conjunto de funciones y **herramientas** que se pueden usar **para manipular y analizar los datos**. Hay diferentes tipos de librerías de Python disponibles que se clasifican según su caso de uso.



Python se utiliza para diversos **fines**, incluida la **visualización de datos**, el **aprendizaje automático** y el **procesamiento del lenguaje**. También se puede usar para **extraer datos** de páginas web, bases de datos u otras fuentes.

LIBRERIAS DE PYTHON PARA EL ANÁLISIS DE DATOS



NumPy

NumPy es una librería de Python para trabajar con **arreglos, matrices y objetos de dimensiones superiores**. Es uno de los paquetes fundamentales para la computación científica con Python.



Pandas

Pandas es una librería de herramientas de análisis de datos escrita en Python. Proporciona estructuras de datos y funciones que facilitan el **trabajo con diferentes tipos de conjuntos de datos**.

Los pandas se pueden utilizar para diversas tareas, como: **manipulación de datos, visualización de datos, modelado estadístico, aprendizaje automático y análisis de los datos**.





SciPy

SciPy es una librería de Python para matemáticas, ciencias e ingeniería. Proporciona **herramientas para el cálculo numérico, la optimización, la integración y otras operaciones matemáticas.**



Matplotlib

Matplotlib es una librería de trazado 2D de Python que **produce gráficos en una variedad de formatos.** Se puede utilizar en otros idiomas a través de una interfaz.



La versión actual de matplotlib utiliza el paradigma de la programación orientada a objetos, pero también brinda acceso a la mayoría de sus funciones a través de una interfaz de procedimiento. Ofrece modos interactivos y no interactivos, en los que los usuarios **pueden** crear gráficos y **guardarlos en formatos de archivo como JPEG o PNG** (PNG solo para el modo no interactivo).





Plotly

Plotly es una librería de Python para la visualización de datos. Se puede usar en aplicaciones web, incrustado en cuadernos Jupyter y como una herramienta de línea de comandos independiente.

Permite crear rápidamente **gráficos sofisticados a partir de diferentes fuentes de datos**, incluidos datos bursátiles en vivo e históricos.



Bokeh



Bokeh es una librería de Python que **proporciona una interfaz de alto nivel para visualizaciones interactivas en navegadores web**. Tiene características como interfaces de usuario ampliables, panorámica, cepillado y vinculación.

También tiene soporte para exportar a formatos de video y puede usarse en notebooks Jupyter. La librería bokeh se basa en la librería JavaScript D3 y el lenguaje Python.

Seaborn

Este libro está **diseñado para enseñar a los programadores principiantes cómo trabajar con el lenguaje Python**. Cubre temas como conceptos básicos de programación, tipos y estructuras de datos, flujo de control, funciones, módulos y paquetes, programación orientada a objetos (OOP), interfaces gráficas de usuario (GUI) y manejo de excepciones.



seaborn

SciKit-Learn

La librería SciKit-Learn es una librería de Python que proporciona un conjunto de **herramientas para el aprendizaje automático**.



La librería **incluye una gama de algoritmos de aprendizaje supervisados y no supervisados**, así como **herramientas para el preprocesamiento de datos**. Los algoritmos incluyen regresión lineal y logística, agrupamiento, reducción de dimensionalidad, clasificación y más. El paquete es compatible con varias plataformas y es una extensión de SciPy.

TensorFlow

TensorFlow es una poderosa librería para el aprendizaje automático, desarrollada y mantenida por Google. Está escrito en C++ y Python. Se puede usar para **crear una amplia variedad de modelos de aprendizaje automático**, incluidas redes neuronales profundas que se usan para tareas como el reconocimiento de imágenes y el procesamiento del lenguaje natural.



NLTK

NLTK

NLTK es una librería que brinda acceso a la funcionalidad de Natural Language Toolkit (NLTK) desde el lenguaje Python. Ofrece **herramientas** poderosas pero fáciles de usar para el **procesamiento del lenguaje natural**, que incluyen tokenización, segmentación de oraciones, etiquetado de partes del discurso, extracción de entidades nombradas, análisis de sentimientos y más.

DatabaseInterface

DatabaseInterface es una librería de Python que proporciona una **API uniforme para acceder a varias bases de datos**.

Se puede utilizar para manipular datos en bases de datos SQL, como MySQL, PostgreSQL y SQLite. DatabaseInterface también admite bases de datos NoSQL, incluidas MongoDB y Redis.



SQLAlchemy

SQLAlchemy es una librería de Python que **permite a los desarrolladores crear y ejecutar consultas SQL**, así como **administrar el esquema de la base de datos**. Es una librería muy popular que se puede encontrar en muchos marcos web como Django o Flask.



Proporciona a los desarrolladores una **capa de abstracción** para acceder a una o más bases de datos relacionales y proporciona un **mapeador relacional de objetos**, que permite al desarrollador acceder a datos relacionales utilizando objetos de Python.





BeautifulSoup

Beautiful Soup es una librería de Python para **analizar documentos HTML y XML**.

Esta librería está diseñada para proyectos de respuesta rápida, como **extracción de pantalla**, y se puede usar para tareas como **extracción de datos, extracción de información, moderación de contenido, extracción de datos web y archivo**.

BeautifulSoup

Lxml



Lxml es una librería de Python para **analizar y manipular archivos XML y HTML**. Ofrece una API fácil de usar para scripts de Python y un conjunto de contenedores para las librerías libxml2 y libxslt, lo que le permite analizar archivos XML y HTML.



Scrapy

Scrapy es un marco rápido de raspado de pantalla y rastreo web de alto nivel, que **se utiliza para rastrear sitios web y extraer datos estructurados de sus páginas**. Se puede utilizar para una amplia gama de propósitos, desde extracción de datos hasta monitoreo y pruebas automatizadas.



Scrapy



pontia

¿Estás interesado en aprender más sobre las librerías de Python?

En Pontia tenemos másters en Data Analytics y en Data Science, donde te enseñamos más sobre estas librerías de este lenguaje tan demandado.

www.pontia.tech

Pontia.Tech

