

bibliotecas NumPy y Pandas

¿ Que son ?

pandas

pandas es un paquete de manipulación de datos en Python para datos tabulares. Es decir, datos en forma de filas y columnas, también conocidos como DataFrames.

NumPy

NumPy es una biblioteca fundamental para la computación científica en Python. Proporciona soporte para trabajar con grandes matrices y matrices multidimensionales, así como una colección de funciones matemáticas de alto rendimiento

¿ Para que sirven ?

- Manipulación de datos estructurados: Permite trabajar con tablas de datos, como si fueran hojas de cálculo o bases de datos. Los dos objetos clave son:
- DataFrame: Una tabla bidimensional (como una hoja de Excel o una tabla SQL).
- Series: Una estructura unidimensional (similar a una columna en una tabla).

- Permite crear y operar con vectores y matrices multidimensionales
- Ofrece funciones matemáticas de alto nivel
- Permite resolver sistemas de ecuaciones lineales
- Ayuda a ignorar algunos elementos de un array según ciertas condiciones

¿ Que tipo de datos trabajan

- Números enteros (int):
- Números decimales (float)
- Cadenas de texto (string o object):
- Booleanos (bool)
- Fechas y tiempos (datetime)
- Categorías (category)

Enteros (int)

Booleanos

Objetos genéricos

Fechas y horas

Estructuras de datos compuestas

¿ Que beneficios ofrecen ?

Pandas proporciona dos estructuras principales, DataFrame (tabla bidimensional) y Series (estructura unidimensional), que hacen que la manipulación de datos sea más sencilla y eficiente
Filtrado y Selección: Puedes seleccionar, filtrar y acceder fácilmente a datos utilizando condiciones o etiquetas, lo que hace que las consultas sean rápidas y fáciles de entender..

NumPy proporciona una gran variedad de funciones matemáticas (suma, resta, multiplicación, etc.), estadísticas (media, desviación estándar, etc.) y de álgebra lineal (multiplicación de matrices, determinantes, etc.) que se pueden aplicar directamente a los arreglos sin necesidad de usar bucles explícitos.