

NUMPY Y PANDAS

1

QUE ES NUMPY

Numpy es una biblioteca de Python para el cálculo científico. Proporciona funciones para crear y manipular arrays multidimensionales, así como herramientas para realizar operaciones matriciales y estadísticas. Numpy es esencial para la mayoría de las bibliotecas de ciencia de datos de Python, ya que proporciona una base eficiente para el almacenamiento y el cálculo de datos numéricos.



2

QUE SON PANDAS

Pandas es una biblioteca de Python que proporciona estructuras de datos y herramientas para el análisis de datos. Es muy utilizado en el campo de la ciencia de datos y es muy similar a las herramientas de análisis de datos que se encuentran en software estadístico como R o SAS.



3

CARACTERISTICAS DE NUMPY

Los datos, presentados en forma de números, se tratan como arrays (cuadros) para las funciones multidimensionales y las operaciones de reordenación. Se trata de una herramienta muy utilizada en el campo de la Data Science.

Entre las múltiples bibliotecas de Python, NumPy es una de las más utilizadas. Esto se debe a que muchas técnicas de Data Science requieren cuadros y matrices de gran tamaño y cálculos complejos para extraer información valiosa de los datos. NumPy simplifica este proceso con funciones matemáticas diversas y variadas.

Aunque es básica, es una de las bibliotecas de Python más importantes para el cálculo científico. Además, otras bibliotecas dependen en gran medida de los arrays NumPy que utilizan como inputs y outputs (entrada y salida de datos). Por ejemplo, TensorFlow y Scikit aprenden a utilizar arrays NumPy para calcular multiplicaciones de matrices.



4

CARACTERISTICAS PANDAS

Entre las principales características de Pandas se encuentran:

- Lectura y escritura de diferentes formatos de archivos: Pandas puede leer y escribir datos en formatos como CSV, Excel, JSON, entre otros.
- Manipulación de datos: Pandas proporciona una variedad de funciones para manipular datos, como filtrar, agrupar, unir, entre otras.
- Tratamiento de datos faltantes: Pandas proporciona funciones para manejar datos faltantes, como rellenar o eliminar datos faltantes.
- Análisis estadístico: Pandas proporciona funciones para realizar análisis estadísticos simples, como calcular medias, desviaciones estándar, entre otras.
- Visualización: Pandas se integra con otras bibliotecas de visualización como Matplotlib y Seaborn para crear gráficos y visualizaciones de los datos.



5

QUE HACE NUMPY

NumPy (pronunciado /ˈnʌmpaɪ/ (numpai) o, a veces /ˈnʌmpɪ/ (numpi)) es una biblioteca para el lenguaje de programación Python que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.



6

QUE HACEN LOS PANDAS

La biblioteca de software de código abierto Pandas está diseñada específicamente para la manipulación y el análisis de datos en el lenguaje Python. Es potente, flexible y fácil de usar. Gracias a Pandas, por fin se puede utilizar el lenguaje Python para cargar, alinear, manipular o incluso fusionar datos.



7

EN RESUMEN

En resumen, Pandas es una biblioteca de Python para el procesamiento y análisis de datos, mientras que Numpy es una biblioteca de Python para el cálculo científico. Juntas, estas dos bibliotecas proporcionan un gran conjunto de herramientas para facilitar la manipulación, limpieza y análisis de datos en Python.

