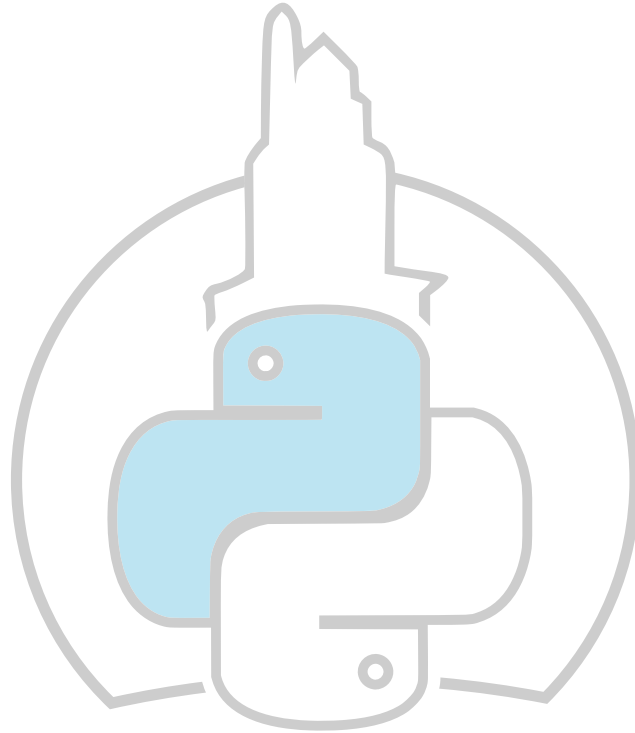


Adán Rivas

2023-06-10



INTRODUCCIÓN A



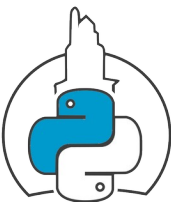
pandas

¿Qué es Pandas?

Es una herramienta de análisis y manipulación de datos de código abierto rápida, potente, flexible y fácil de usar.

Proporciona una estructura diseñada para trabajar con datos **relacionales** y **etiquetados** de una manera intuitiva.

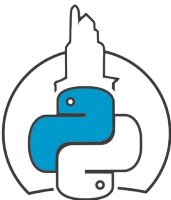
Es una de las librerías más utilizadas en Python y su nombre viene de *panel data*.



¿Qué podemos hacer con Pandas?

Es la navaja suiza de Python:

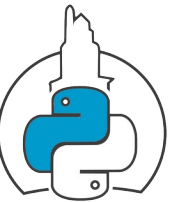
- Lee y escribe en múltiples formatos como: CSV, JSON, HTML, XML, Parquet, Pickle, HDF5, Excel, SQL y un largo etc.
- Puede ejecutar todo tipo de funciones estadísticas (conteos, sumas, máximos, desviación estándar, medias, etc).
- Tiene un montón de funciones y métodos para limpiar datos (borrado de duplicados, nulos, etc).
- Puede concatenar y ejecutar diferentes tipos de JOINS. Puede manejar y crear series temporales.
- Integración con Matplotlib de serie.



¿Dónde utilizarlo?

Con todo lo mencionado antes, Pandas se convirtió en una de las herramientas más utilizadas en el mundo de la ingeniería y análisis de datos. Aquí unos ejemplos de tareas automatizadas con Pandas:

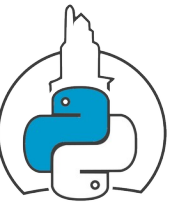
- Data Wrangling.
- ETL o ELT.
- Análisis de datos.
- Creación de informes.



Estructuras de datos

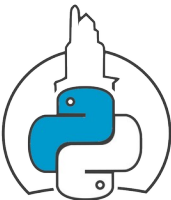
En Pandas manejamos principalmente dos objetos:

- **Series**
- **Dataframe**



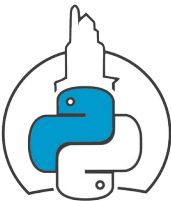
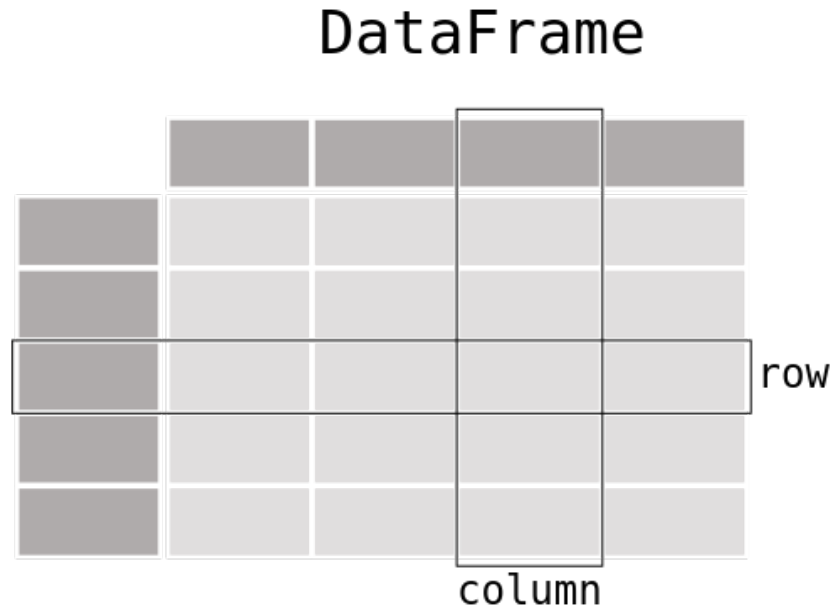
Series

Estructuras unidimensionales que contienen un array de datos soportado por NumPy y una etiqueta o índice. El array de datos tiene que ser del mismo tipo (texto, entero, fecha) y el índice es opcional e inmutable.



Dataframe

Estructuras en formato tabular que contiene uno o más objetos Series para formar filas y columnas. Estos contienen dos índices, uno para las columnas y otro para las filas.



Fuentes

- <https://interactivechaos.com/es/course/tutorial-de-pandas>
- https://pandas.pydata.org/docs/getting_started/index.html
- <https://www.kaggle.com/datasets/>

