



Programación Avanzada 2025

Lab. 4.3. Introducción a Pandas.

Septiembre 13, 2025

Cree una carpeta (folder) en el disco D, nómbrela con su apellido paterno seguido de su código. Ejemplo: LOPEZ12345

Instalando Pandas

Primer caso: Si Python fue agregado al PATH de Windows durante la instalación, es decir, usted colocó check a la opción (☒ Add Python 3.13.7 to PATH).

Ingresa a la aplicación Símbolo del Sistema (línea de comandos del Sistema Operativo) de Windows **como Administrador**. Debe mostrar algo similar a lo siguiente:

C:\Users\tu_nombre_de_usuario>

Para instalar Pandas coloque el siguiente comando:

C:\Users\tu_nombre_de_usuario> pip install pandas

Con este último comando se habrá instalado Pandas.

Para probar si Pandas está instalado, colocamos el siguiente comando en el interpretador de Python (**IDLE**):

```
>>>import pandas
```

Si no devuelve ningún error es que el programa está instalado correctamente.

Si logró instalar Pandas, entonces puedes instalar también Matplotlib:

C:\Users\tu_nombre_de_usuario> pip install matplotlib

Ejemplo de uso de Pandas:

Coloque en el interpretador de Python, IDLE:

```
import pandas as pd

data = {"apples": [3, 2, 0, 1], "oranges": [0, 3, 7, 2]}

purchases = pd.DataFrame(data)

print(purchases)
```

Segundo caso: Si Python NO fue agregado al PATH de Windows durante la instalación.

- a) Ingresar a la aplicación Símbolo del Sistema (línea de comandos del Sistema Operativo) de Windows **como Administrador**. Debe mostrar algo similar a lo siguiente:

```
C:\Users\tu_nombre_de_usuario>
```

- b) Ubicar la carpeta donde se encuentra instalado Python, esto depende de la instalación. La carpeta podría estar en la siguiente ruta **C:\Program Files\Python313** (la carpeta *Program Files* podría estar con el nombre *Archivos de programas*).

Si la ruta fuera **C:\Program Files\Python313**, ejecutar el siguiente comando:

```
C:\Users\tu_nombre_de_usuario>CD C:\Program Files\Python313
```

Ahora el cursor nos dice que estamos en la carpeta de Python:

```
C:\Program Files\Python313>
```

- c) Colocar el siguiente comando (fíjese que al parámetro **user** le antecede **dos guiones**):

```
C:\Archivos de programas\Python313>python -m pip install pandas --user
```

Con el último comando Pandas debió instalarse.

- d) Ya que logró instalar Pandas, aproveche en instalar Matplotlib, paquetes que usaremos más tarde. En el último comando solo hay que cambiar el nombre numpy por matplotlib:

```
C:\Archivos de programas\Python313>python -m pip install matplotlib --user
```

Para probar si Pandas y Matplotlib están instalados, colocamos el siguiente comando en el interpretador de Python (IDLE):

```
>>>import pandas (import matplotlib para Matplotlib)
```

Si no devuelve ningún error es que los programas están instalados correctamente.

Ejemplo de uso de Pandas:

Coloque en el interpretador de Python, IDLE:

```
import pandas as pd

data = {"apples": [3, 2, 0, 1], "oranges": [0, 3, 7, 2]}

purchases = pd.DataFrame(data)

print(purchases)
```

Para los siguientes ejercicios, elabore un programa en Python. Verifique los programas ejecutando y probando con distintas entradas. Use las librerías NumPy, Pandas.

Pandas. Series

1. Dada la siguiente lista ["carlos","jose","maria","pedro","monica","julia","pavel","carmen"], genere dos series pandas A y B con N valores aleatorios de la lista dada.
 - a) Muestre los valores de A que no estén presentes en B.
 - b) Muestre los valores no comunes de A y B
 - c) Determine la frecuencia de los valores únicos de A
 - d) Modifique la serie pandas B, reemplazando todos los valores con "nuevo", excepto los dos más frecuentes.
 - e) Cambie a mayúscula la primera letra de cada palabra de A
 - f) Determine la longitud de cada palabra de A
 - g) Ordene las palabras de A en forma descendente.

Pandas. Dataframe

2. Dado el siguiente diccionario:

```
exam_data = {  
'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']  
}
```

- a) Genere el dataframe pandas A en base al diccionario exam_data.
- b) Muestre la información básica de A
- c) Extraer los elementos de A de las filas, 1, 3, 5 y 6
- d) Extraer los elementos de A de las filas 1, 3, 5 y 6 pero solo de las columnas 1 y 3.
- e) Extraer las filas donde la columna `score` tenga valores perdidos.
- f) Extraer las filas donde la columna `score` tenga valores entre 15 y 20 ambos inclusive
- g) Determine el promedio de las notas de los alumnos
- h) Muestre los nombres de los alumnos que lograron calificar

3. Leer con pandas el archivo csv desde la siguiente dirección:

https://raw.githubusercontent.com/selva86/datasets/master/Cars93_miss.csv

- a) ¿Cuál es el fabricante, modelo y tipo que tiene el mayor precio?
- b) Renombre la columna `Type` como `CarType`. Algunos nombres de columnas contienen el carácter punto ("."), reemplácelos con el carácter "_".
- c) Cuente la cantidad de valores perdidos (NaN) en cada columna. ¿Qué columna tiene el máximo número de valores perdidos?
- d) Reemplace los valores perdidos en las columnas `Min.Price` y `Max.Price` con sus respectivos promedios.
- e) Ordene el dataframe de acuerdo a la columna `Weight` en orden descendente.
- f) Intercambie la primera y la segunda fila del dataframe.
- g) Extraiga todas las filas correspondientes al fabricante Toyota.

Guarde todos vuestros programas en una carpeta con el nombre su **Apellido** paterno seguido de vuestro **DNI**, luego comprima esta carpeta. Envíe este archivo a Katherine Navarro katherine.navarro@upch.pe especificando como asunto **Lab4.3**.