



## Módulo II | Clase 4

# Python para Data Science. Introducción a Pandas





**¿Ponemos a grabar  
el taller?**

## ¿Qué vamos a ver hoy?



- Introducción a Pandas
  - Crear un DataFrame
  - Atributos y Funciones Básicas
  - Seleccionar y Filtrar Columnas y Filas
- Operaciones y Manipulación de Dataframes
  - Práctica Numpy y Pandas



# Pandas

## PANel DAta

Pandas o *Python Data Analysis* es una librería de Python especializada en el **manejo y análisis de estructuras de datos**.

Incorpora una nueva estructura de datos, los **DataFrames**, basadas en los arrays de la librería NumPy pero incorpora nuevas funcionalidades. Permite acceder y manipular datos fácilmente así como leer archivos.

# ¿Por qué PANDAS?

Pandas ofrece una manera de realizar operaciones eficientes sobre **DataFrames**. Esta define un conjunto de datos estructurado en forma de tabla donde cada **columna** contiene datos del mismo tipo, y las **filas** son registros que pueden tener datos de distintos tipos.

		columns					
		state	color	food	age	height	score
	Jane	NY	blue	Steak	30	165	4.6
	Niko	TX	green	Lamb	2	70	8.3
	Aaron	FL	red	Mango	12	120	9.0
	Penelope	AL	white	Apple	4	80	3.3
	Dean	AK	gray	Cheese	32	180	1.8
	Christina	TX	black	Melon	33	172	9.5
	Cornelia	TX	red	Beans	69	150	2.2

Filas

Índice

## ¿Por qué PANDAS?

Un DataFrame contiene dos índices, uno para las filas y otro para las columnas. Se puede acceder a los datos mediante **índices** o **nombres** para filas y columnas.

Permite **leer** y **escribir** archivos en formato CSV, Excel y bases de datos SQL de manera sencilla y rápida.

Ofrece **métodos** que permiten reordenar, separar, combinar y reformatear los datos.





# Sección práctica

**Aprendemos cómo funciona Pandas con  
la notebook 6**





¿Alguna consulta?



# Descanso

Nos vemos en 10 minutos



# Sección práctica

Resolvemos ejercicios de Numpy y Pandas de las  
Notebooks 5 y 7



# Trabajamos en salas

## Trabajamos en salas de Zoom

### Práctica Numpy y Pandas

En los grupos establecidos, comentamos dudas y soluciones a los ejercicios planteados en las Notebooks 5 y 7

*45 minutos de actividad*



# Descanso

Nos vemos en 10 minutos

## Numpy y Pandas

Trabajamos con la Notebook 5 y 7

Repasamos la funcionalidad y aplicaciones de Numpy y Pandas.

## Primera pre-entrega

¡En la próxima clase, deberán entregar los desafíos resueltos en la notebook 5 y 7. Los mismos serán subidos como repositorio a su cuenta de Github.

**Presentarán lo trabajado entregando el link al de Github en el foro del aula virtual**







¿Alguna consulta?

FUNDACIÓN  
**YPF**

¡Muchas gracias!

