

# CHEATSHEET PANDAS – ESTRUCTURA DE DATOS

## SET UP

- **Instalación:** `conda install pandas`
- **Importación:** `import pandas as pd`

## ESTRUCTURA DE UN OBJETO SERIES

### Axis 0 / index



## ESTRUCTURA DE UN OBJETO DATAFRAME



# CHEATSHEET PANDAS – ESTRUCTURA DE DATOS

## TIPOS DE DATOS

- **float**: es el float de Numpy, que soporta datos nulos
- **int**: es el int de Numpy, que NO soporta datos nulos
- **Int64**: es un int que introduce Pandas para soportar datos nulos
- **object**: es el object de Numpy, que se usa como texto para meter cualquier cosa
- **category**: es un tipo que introduce Pandas específicamente para variables categóricas
- **bool**: es el booleano de Numpy, que NO soporta datos nulos
- **boolean**: es el booleano que introduce Pandas para soportar datos nulos
- **datetime64**: es el tipo de Numpy para las fechas, que NO soporta datos nulos

## CONVERSIONES DE ESTRUCTURAS DE DATOS

- **Crear un Series a partir de una lista:** `pd.Series(['a','b','c'])`
- **Crear un Series a partir de un diccionario:** `pd.Series({'a': 1, 'b': 2, 'c': 3})`
- **Crear un Series a partir de un array de numpy:** `pd.Series(np.array([1,2,3]))`
- **Pasar de Series a Dataframe:** `serie.to_frame()`
- **Crear un Dataframe a partir de una lista con sus componentes:**  
`pd.DataFrame(data = datos_para_df, columns=columnas_para_df, index = indice_para_df)`
- **Crear un Dataframe a partir de un diccionario:**  
`pd.DataFrame({'letras': ['a','b','c'], 'numeros': [1,2,3]})`
- **Crear un Dataframe a partir de un array de 2 dimensiones:**  
`pd.DataFrame(vector_2d, index = ['fila1','fila2'], columns= ['col1','col2','col3'])`
- **Pasar de Dataframe a Series (notación de punto):** `df.var1`
- **Pasar de Dataframe a Series (notación de corchetes):** `df['var1']`