



PANDAS

Utiliza la siguiente convención para importar la librería:

```
> import pandas as pd
```

ESTRUCTURA DE DATOS

Series

Serie de una sola dimensión con datos etiquetados capaz de almacenar cualquier tipo de datos

```
> a = pd.Series( [15, 4, 'Si', 'No'],
                 index = ['huevos', 'manzana', 'leche',
                        'pan'] )
```

huevos	15
manzana	4
leche	Si
pan	No

Dataframe

DataFrame de dos dimensiones con datos etiquetados y estructurados con columnas de diferentes tipos

```
> datos = {'Carlos' : pd.Series(data = [245, 25, 55]),
           'Alicia' : pd.Series(data = [40, 110, 500])}
```

	Carlos	Alicia
0	245	40
1	25	110
2	55	500

```
> df = pd.DataFrame(datos)
```

Buscando ayuda

```
> help(pd.Series.loc)
```

ENTRADAS Y SALIDAS

Cargar datos desde archivos

```
> df = pd.read_csv('train.csv')
```

Guardar datos en archivos

```
> df.to_csv('DataFrame.csv')
```

INSPECCIONAR SERIES Y DATAFRAME

Forma de los datos

Forma de los datos Series > **a.shape**
Forma de los datos DataFrame > **df.shape**

Número de elementos

Número de elementos Series > **a.size**
Número de elementos DataFrame > **df.size**

Número de dimensiones

Número de dimensiones Series > **a.ndim**
Número de dimensiones DataFrame > **df.ndim**

Conocer los datos

Conocer los datos Series > **a.values**
Conocer los datos DataFrame > **df.values**

Conocer las etiquetas de los índices

Conocer las etiquetas de los índices Series > **a.index**
Conocer las etiquetas de los índices DataFrame > **df.index**

Conocer las etiquetas de las columnas

Conocer las etiquetas de las columnas DataFrame > **df.columns**

Conocer la información

Conocer la información DataFrame > **df.info()**

Contar los datos

Contar los datos DataFrame > **df.count()**

Conocer los primeros y los últimos datos

Conocer los primeros datos DataFrame > **df.head()**
Conocer los últimos datos DataFrame > **df.tail()**

SUBCONJUNTO E INDEXACIÓN

Acceder a elementos de una Serie

Acceder a elementos de una Serie > **a['huevos']**
Acceder a elementos de una Serie > **a[0, 1]**
Acceder a elementos de una Serie > **a.loc[['huevos', 'manzana']]**
Acceder a elementos de una Serie > **a.iloc[0, 1]**

Acceder a elementos de un DataFrame

Acceder a elementos de un DataFrame > **df[['Carlos']]**
Acceder a elementos de un DataFrame > **df['Carlos'] [0]**
Acceder a elementos de un DataFrame > **df.loc[0]**

MANIPULACIÓN DE SERIES Y DATAFRAME

Modificar elementos

Modificar elementos de una Serie > **a['huevos'] = 1**
Modificar elementos de un DataFrame > **df['Carlos'] = [20,5,10]**

Añadir elementos

Añadir elementos a un Serie > **a['huevos'] = 1**
Añadir elementos a un DataFrame > **df.insert(2, 'Marcos', [80,50,10])**
Añadir elementos a un DataFrame > **df.append(df2)**

Eliminar elementos

Eliminar elementos a una Serie > **a.drop('manzana')**
Eliminar elementos a un DataFrame > **df.pop('Carlos')**
Eliminar elementos a un DataFrame > **df.drop(['Carlos'], axis = 1)**

Renombrar filas y columnas

Renombrar filas DataFrame > **df = df.rename(index = {0: 'set_1'})**
Renombrar columnas DataFrame > **df = df.rename(columns={'Carlos':'Bob'})**

Añadir un índice

Añadir un índice DataFrame > **df = df.set_index('Carlos')**

ESTADÍSTICAS

Estadísticas descriptivas

Estadísticas descriptivas > **df.describe()**

Valor máximo y mínimo de una columna

Valor máximo de una columna > **df.max()**
Valor mínimo de una columna > **df.min()**

Correlación de datos

Correlación de datos > **df.corr()**

Media de los datos

Media de los datos > **df.mean()**

Mediana de los datos

Mediana de los datos > **df.median()**

