

Formulario 6 de Python

Importar pandas, matplotlib.pyplot y numpy

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

Crear base de datos a partir de un directorio

```
base = pd.DataFrame({'Columna1':[0,1,2,3],
                    'Columna2':[4,5,6,7],
                    'Columna3':['a','b','c','d']})
```

Cargar una base de datos en formato csv desde Github

```
url =
'https://raw.githubusercontent.com/RobertoCabal/Bundesliga/master/BUNDESLIGA_19-20.csv'
base = pd.read_csv(url,encoding='unicode_escape')

#El encoding='unico_escape' no siempre se usa, en este caso
#es por temas de codificación de nombres.
#Si se quita marca un error.
```

Cargar base de datos en formato csv desde nuestra computadora

```
from google.colab import files
files.upload()

#Después de ejecutar el código debes seleccionar el archivo
#Cuando se cargue el archivo ya lo puedes abrir

base = pd.read_csv('BUNDESLIGA_19-20.CSV',encoding='unicode_escape')
```

Ver primeras y últimas filas

```
base.head()
base.tail()
len(base) #número de filas
base.shape #número de filas y columnas
```

Valores únicos en una columna

```
base['Columna1'].unique()  
len(datos['Columna1'].unique()) #número de valore únicos
```

Estadísticas simples

```
base.describe()
```

Convertir a lista

```
col_list = base['Columna1'].tolist()
```

Índices de la base

```
index = base.index.tolist()
```

Agregar columna llena de None's

```
base['Nueva_columna'] = None
```

Agregar otra base o fila

```
fila_nueva = pd.DataFrame({'Columna1':[0],  
                           'Columna2':[3],  
                           'Columna3':['d'],  
                           'Nueva_columna':[0]})  
  
base_nueva = base.append(fila_nueva,ignore_index=True)
```

Índices

```
#Columna 1  
base.Columna1  
base['Columna1']  
  
#Elemento i, j  
i = 0  
j = 2  
base.iloc[i,j]
```

```
#Elemento i de la columna 'Columna1'
base.loc[i,'Columna1']
```

Sub-base

```
#Base cuyas filas tienen cero en la columna 'Columna1'
base[base['Columna1']==0]

#Base cuyas filas son mayores a cero en la 'Columna1'
#e igual a 'c' en la 'Columna3'
filtro = (base['Columna1']>0)&(base['Columna3']=='c')
base[filtro]
```

Ordenar

```
#Ordenar en orden descendiente con respecto a la columna 2
base.sort_values('Columna2', ascending=False)
```

Agrupar

[illegible]