

Presentación THC

Antonino Equihua Lombera, Maximiliano Proaño Bernal

Facultad de Ciencias, UNAM

10 de diciembre del 2019

Pandas

Presentación
THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

Esta biblioteca nos ayuda a la manipulación de datos de tal manera que se puedan representar en forma de tablas. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales.

Su nombre deriva de un término de econometría: Datos de Panel

Pandas

Presentación
THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

Usa un concepto llamado **DataFrame**, el cual es una estructura de dos dimensiones etiquetada donde podemos almacenar datos de diferentes tipos. Tiene ciertas similitudes con una hoja de cálculo de Excel.

Pandas

Presentación THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

```
Pandas.py - /home/maximilianooproanobernal/THC/OK/Python/Pandas.py (3.7.3) x Python 3.7.3 Shell
File Edit Format Run Options Window Help
import pandas as pd

""" El primer ejemplo que vamos a poner va a ser el de definir una estructura de datos '
es un array de datos unidimensional con indexación. Las 'Series' se definen de la siguiente
Es decir, que en el primer parámetro le indicamos los datos del array y en el segundo pa

Aquí un ejemplo.
"""
index=[1, 15, 3, 5, 11, 14, 16, 8, 18, 6, 7]
JugadoresEspañoles = pd.Series(
    ['Casillas', 'Ramos', 'Pique', 'Puyol', 'Capdevila', 'Xabi Alonso', 'Busquets', 'Xavi',
     'Iniesta', 'Villa'], index)
print("-----")
print ("Jugadores Españoles De Futbol: \n%s" % JugadoresEspañoles)
print("-----")

""" En este siguiente caso; en el que no le indiquemos los índices de forma explícita,
de forma automática empezando desde el valor cero:
"""
JugadoresEspañoles= pd.Series(
    ['Casillas', 'Ramos', 'Pique', 'Puyol', 'Capdevila', 'Xabi Alonso', 'Busquets', 'Xavi',
     'Iniesta', 'Villa'])
print ("Jugadores Españoles De Futbol: \n%s" % JugadoresEspañoles)

"""También podemos crear una estructura de datos 'Series' a partir de una lista o de
si la construimos a partir de una lista nos pondrá los índices por defecto y si lo cream
índices las claves. Vamos en este ejemplo a que me refiero y también, de como crear una
vamos a ver como insertar en esta serie un nuevo elemento:

Aquí el ejemplo.
"""
Jugadores = {'1': 'Casillas', '15': 'Ramos', '3': 'Pique', '5': 'Puyol', '11': 'Capdevila', '14': '
16': 'Busquets', '8': 'Xavi Hernandez', '18': 'Pedrito', '6': 'Iniesta', '7': 'Vil

Jg= pd.Series(Jugadores)
#Insert new player
Jg[10] = 'Isco'
print ("Jugadores Españoles: \n%s" % Jg)

""" Vamos a pasar a continuación a ver un ejemplo con la estructura de datos 'DataFrame'
Como ya se ha comentado es una estructura de datos similar a una tabla de una base de
datos relacionar, una tabla de excel, etc.

Para construir un DataFrame se puede hacer de diferentes formas, como por ejemplo a partir
de una Serie, de otro DataFrame, leyendo una tabla excel, csv, etc. Vamos a ver a contin

Amor al ejemplo.
>>>
```

Figure:

Kivy es una biblioteca de código abierto de Python que se utiliza para crear aplicaciones móviles y también aplicaciones con una interfaz de usuario natural, se puede ejecutar Android, iOS, GNU/Linux, OS X, and Windows.

Histograms:

Este módulo nos permite mostrar la información estadísticamente en forma de histograma, el cual es una representación de la frecuencia de valores utilizando barras. Para esto, el módulo utiliza el concepto de "bin" como el número de divisiones para representar los datos.

Matplotlib

Presentación THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

Multiple Histograms.py - /home/maximilianooproanobernal/Documents/Multiple Histograma... x

File Edit Format Run Options Window Help

```
shape of a histogram, the Astropy docs have a great section on how to
select these parameters:
http://docs.astropy.org/en/stable/visualization/histogram.html
"""
```

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

```
np.random.seed(0)
number_of_bins = 20
```

```
# An example of three data sets to compare
```

```
number_of_data_points = 387
labels = ['A', 'B', 'C']
data_sets = [np.random.normal(0, 1, number_of_data_points),
              np.random.normal(6, 1, number_of_data_points),
              np.random.normal(-3, 1, number_of_data_points)]
```

```
# Computed quantities to aid plotting
```

```
hist_range = (np.min(data_sets), np.max(data_sets))
binned_data_sets = [
    np.histogram(d, range=hist_range, bins=number_of_bins)[0]
    for d in data_sets
]
```

```
binned_maximums = np.max(binned_data_sets, axis=1)
x_locations = np.arange(0, sum(binned_maximums), np.max(binned_maximums))
```

```
# The bin_edges are the same for all of the histograms
```

```
bin_edges = np.linspace(hist_range[0], hist_range[1], number_of_bins + 1)
centers = 0.5 * (bin_edges + np.roll(bin_edges, 1))[:-1]
heights = np.diff(bin_edges)
```

```
# Cycle through and plot each histogram
```

```
fig, ax = plt.subplots()
for x_loc, binned_data in zip(x_locations, binned_data_sets):
    lefts = x_loc - 0.5 * binned_data
    ax.barh(centers, binned_data, height=heights, left=lefts)
```

```
ax.set_xticks(x_locations)
ax.set_xticklabels(labels)
```

```
ax.set_ylabel("Data values")
ax.set_xlabel("Data sets")
```

```
plt.show()
```

Figure 1

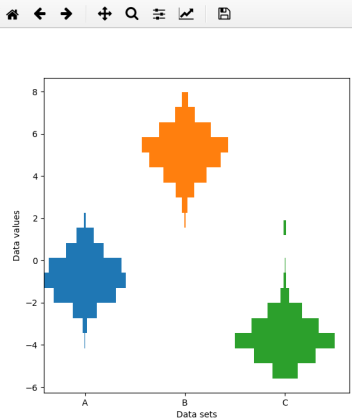


Figure:

Contour:

Con este módulo es posible representar la superficie de un espacio tridimensional en un plano bidimensional. La forma en la que se muestra dicha superficie es a base de colores. Por lo general se usan diferentes gradientes del mismo color para los "niveles" de la superficie, es decir: si el punto más alto se colorea de blanco y el más bajo de azul marino, las alturas que están en medio se colorean con mezclas de blanco y azul

Matplotlib

Presentación THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

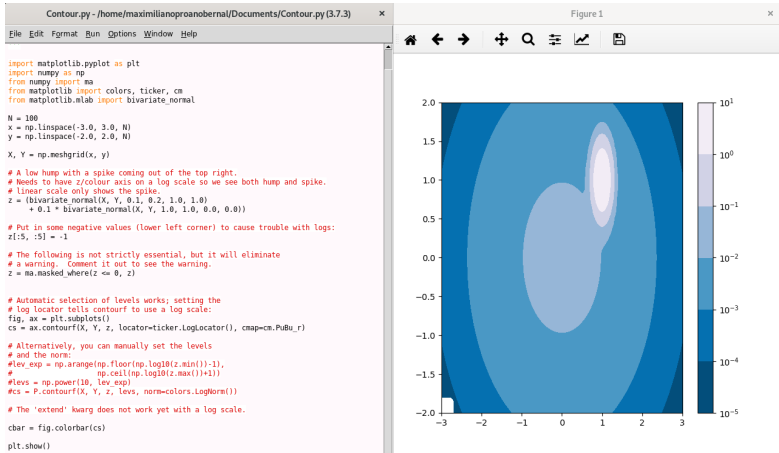


Figure:

Streamplot:

Se usa para representar un campo vectorial en un plano, lo cual, para poder hacerlo, emplea flechas para la dirección de los vectores, modificaciones de colores, grosor de la líneas, entre otras cosas.

Streamplot:

Algunos de los parámetros que utiliza son:

- * x,y : Arreglos unidimensionales. Hace una cuadrícula
- * u,v : Arreglos bidimensionales.
- *Densidad: Controla la cercanía entre los vectores
- *Grosor de la línea: Modifica el grosor de dichos vectores

Matplotlib

Presentación THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

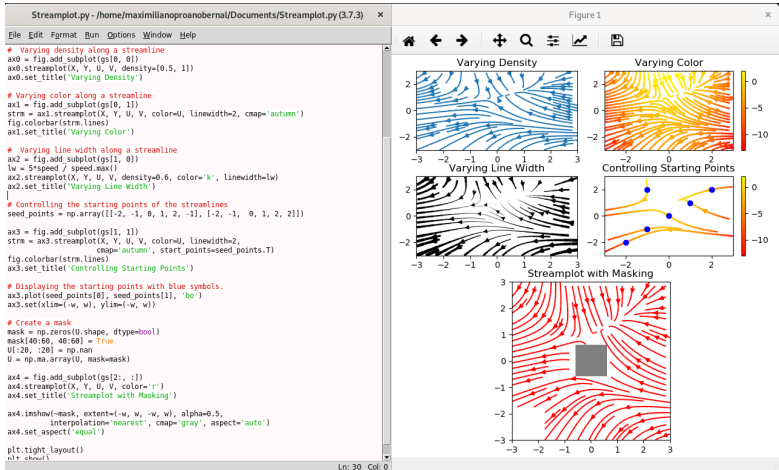


Figure:

Interpolation:

La interpolación es un tipo de estimación estadística que, a partir de valores relacionados y conocidos, se puede llegar a un resultado aproximado de valores desconocidos. Algunos de los métodos posibles son: “nearest”, “none”, “gaussian”, entre otros. Cada uno de ellos muestra una interpolación distinta

Matplotlib

Presentación THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

```
Interpolation.py - /home/maximilianooproanobernal/THC/OK/Python/Interpolation.py (3.7.3) x
File Edit Format Run Options Window Help
'''
Show all different interpolation methods for imshow
'''

import matplotlib.pyplot as plt
import numpy as np

# from the docs:
|
# If interpolation is None, default to rc image.interpolation. See also
# the filternorm and filterrad parameters. If interpolation is 'none', then
# no interpolation is performed on the Agg, ps and pdf backends. Other
# backends will fall back to 'nearest'.
#
# http://matplotlib.org/api/figure_api.html#matplotlib.figure.Figure.imshow

methods = [None, 'nearest', 'bilinear', 'bicubic', 'spline16',
           'spline36', 'hanning', 'hamming', 'hermite', 'kaiser', 'quadric',
           'catrom', 'gaussian', 'bessel', 'mitchell', 'sinc', 'lanczos']

np.random.seed(0)
grid = np.random.rand(4, 4)

fig, axes = plt.subplots(3, 6, figsize=(12, 6),
                        subplot_kw={'xticks': [], 'yticks': []})

fig.subplots_adjust(hspace=0.3, wspace=0.05)

for ax, interp_method in zip(axes.flat, methods):
    ax.imshow(grid, interpolation=interp_method, cmap='viridis')
    ax.set_title(interp_method)

plt.show()
```

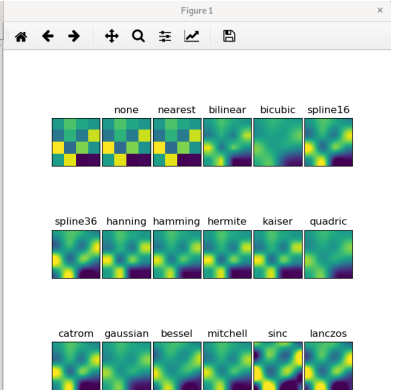


Figure:

Images:

Se puede deducir del nombre que este módulo es capaz de mostrar una imagen guardada en los archivos, sin embargo, también es capaz de cambiar tonalidades de colores, escalas, entre otras cosas.

Matplotlib

Presentación
THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

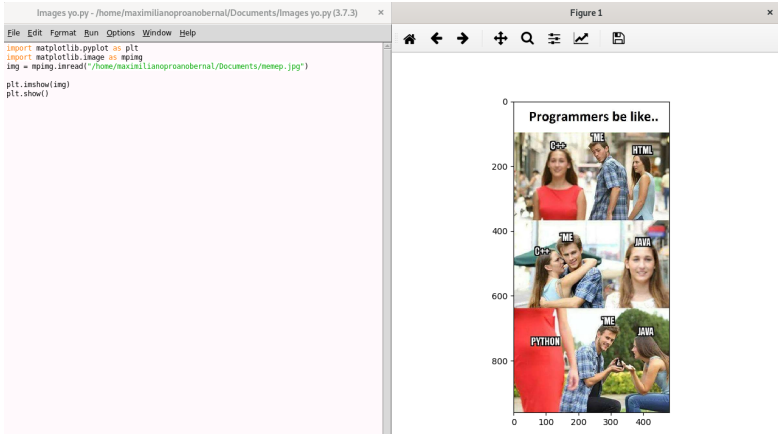


Figure:

Bibliografía

Presentación
THC

Antonino
Equihua
Lombera,
Maximiliano
Proaño Bernal

<https://realpython.com/python-zip-function/>

https://matplotlib.org/examples/images_contours_and_fields/interpolation_methods.html

https://matplotlib.org/examples/images_contours_and_fields/contourf_log.html

<https://pandas.pydata.org/>