



Capacitación Semillero DISW

Python Clase 2

Docentes: Lautaro Delgado, Idc0295@gmail.com Magdalena Bouza, bouza.magdalena@gmail.com

Librerías

Paquetes, Librerías y Frameworks

- Módulo: es la pieza de software más pequeña. Contiene métodos y funciones que pueden ser importados desde otro archivo.
- Paquete: es una colección de módulos. Debe contener un archivo init.py para que el interpretador de Python lo considere como un paquete
- Librería: Es una colección de paquetes
- Framework: es una colección de librerías. Define la arquitectura del programa.

Numpy

Numpy

Una de las librerías que más vamos a usar es Numpy, que significa Numerical Python

Da soporte para arrays multidimensionales y matrices, junto con un conjunto de funciones matemáticas de alto nivel. Numpy introduce la idea de vectorización a Python, y esto es lo que hace a Numpy relativamente rápido.

El objeto central de Numpy son los *ndarrays* (n dimensional arrays)

¿Por qué usar Numpy?

- Posibilidad de trabajar con vectores (arrays)
- Funciones para trabajar en el dominio de álgebra lineal, transformada de Fourier y matrices
- NumPy provee un objeto, el array, que es hasta 50x veces más rápido que las listas de Python.

Visualización

Matplotlib y Seaborn

Matplotlib

Matplotlib es una librería usada para visualización.

Permite crear visualizaciones estáticas, animadas e interactivas usando Python.

La mayor parte de las utilidades de Matplotlib se encuentran dentro del submódulo Pyplot.



API

Quick start

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

X = np.linspace(0, 2*np.pi, 100)Y = np.cos(X)

fig, ax = plt.subplots() ax.plot(X,Y,color='C1')

fig.savefig("figure.pdf") fig.show()

Anatomy of a figure



Subplots layout subplot[s](cols,rows,...) fig, axs = plt.subplots(3,3) G = gridspec(cols,rows,...) [22] ax = G[0,:]ax.inset_axes(extent) d=make_axes_locatable(ax) [33] ax=d.new_horizontal('10%')

Getting help

- matplotlib.org
- github.com/matplotlib/matplotlib/issues
- O discourse.matplotlib.org
- ₫ stackoverflow.com/matplotlib
- ₩ gitter.im/matplotlib
- Matplotlib users mailing list





Tick locators

API

API

Event handling

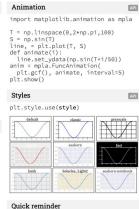
print(event)

fig, ax = plt.subplots()

fig.canvas.mpl_connect(

'button_press_event', on_click)

def on_click(event):



ax.grid() ax.patch.set alpha(0) ax.set_[xy]lim(vmin, vmax) ax.set_[xy]label(label) ax.set_[xy]ticks(list) ax.set_[xy]ticklabels(list) ax.set_[sup]title(title) ax.tick_params(width=10, ...) ax.set_axis_[on|off]()

ax.tight_layout() plt.gcf(), plt.gca() mpl.rc('axes', linewidth=1, ...) fig.patch.set_alpha(0) text=r'\$\frac{-e^{i\pi}}{2^n}\$'

Keyboard shortcuts

ctrl + s Save ctrl + w Close plot r Reset view f Fullscreen 0/1 f View forward b View back o Zoom to rect p Pan view x X pan/zoom y Y pan/zoom g Minor grid 0/1 G Major grid 0/1 X axis log/linear Y axis log/linear

API

READ

Ten Simple Rules

- 1. Know Your Audience 2. Identify Your Message
- 3. Adapt the Figure
- 4. Captions Are Not Optional 5. Do Not Trust the Defaults
- 6. Use Color Effectively 7. Do Not Mislead the Reader 8. Avoid "Chartiunk"
- 9. Message Trumps Beauty 10. Get the Right Tool

Seaborn

<u>Seaborn</u> es una librería basada en Matplotlib, muy utilizada para visualizaciones de datos estadísticos.

Pandas

Pandas

Pandas es una librería usada para trabajar con datasets.

Incorpora funciones para analizar, limpiar, explorar y manipular datos.

¿Por qué usar Pandas?

- Permite analizar datos y sacar conclusiones basadas en teoría estadística.
- Permite limpiar los datasets, de forma de volverlos m'sa legibles y relevantes.
- Muy usado en Ciencia de Datos.