NumPy Introduction

Autor: Hermes Antonio Terán Vélez

*IS&C, Universidad Tecnológica de Pereira, Pereira, Colombia*

Correo-e: h.teran@utp.edu.co

Conceptos básicos de NumPy

**import NumPy as np**

Lo primero que se hace es importer la libreria NumPy y se le asigna el sufijo np para abreviar su llamado.

**a = np.array([0,1,2,3,4,5])**

Se procede a crear un vector de 6 posiciones con la función array que hace parte de la librería NumPy

**b = a.reshape((3,2)).copy()**

Luego se copia ese mismo vector en una nueva variable y con una estructura diferente, ya que el nuevo vector generado tendrá 3 filas y 2 columnas.

**c = a.reshape((3,2)).copy()**

Se crea una nueva copia del vector original a y se guarda ese mismo vector en una nueva variable y con una estructura diferente, ya que el nuevo vector generado tendrá 3 filas y 2 columnas. Esto se hace para no alterar los datos del vector original y poder trabajar sobre copias de los mismos datos.

**c[0][0] = -99**

Se modifica un elemento del vector c y se imprime el vector original y la copia para verificar que los datos del vector original no han sido alterados.

**c = np.array([1, 2, np.NAN, 3, 4])**

el vector c se modifica y se le asignan nuevos valores en los cuales se incluye un error. Este tipo de errores se detectan con la función de NumPy np.isnan, la cual retorna valores booleanos, es decir, True y False.

Para realizar cálculos sobre un vector que tiene errores de tipo NAN, primero se debe detectar los NAN en el vector y después eliminarlos con la función c[~np.isnan(c)].

La función np.mean(c[~np.isnan(c)]) se encarga de calcular el promedio de todos los datos que están en el vector c, eliminando los NAN.

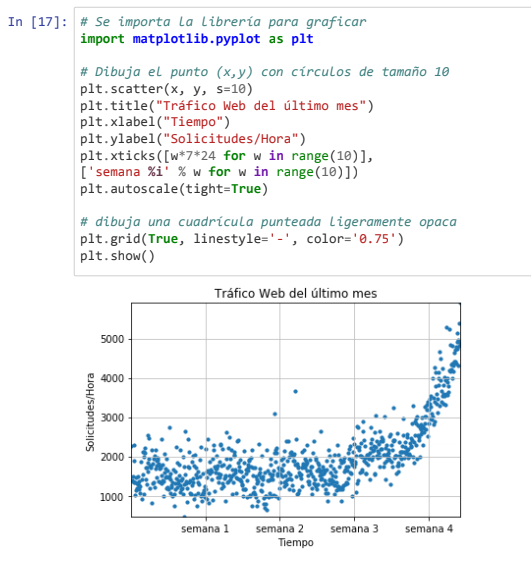
**Data = np.genfromtxt(“web\_traffic.tsv, delimiter=”\t”)**

Esta función se encarga de guardar en una variable llamada data, todos los datos que se encuentran en el archivo tsv y además los separa por un espacio.

**X = x[~np.isnan(y)]**

**Y = y[~np.isnan(y)]**

Las funciones anteriores se encargan de eliminar los errores NAN que se encuentran en los vector X y los datos correspondientes a los errores de X en Y, en los cuales anteriormente se guardó la columna 1 y 2 respectivamente del vector data.



Se procede a importar la librería matplotlib.pyplot como plt para graficar los datos que se encuentran en el vector x y, se les asigna un nombre a la grafica y a los ejes y se divide la información por semanas para observar el comportamiento del gráfico.