

Páctica 6

Análisis de datos con Pandas y Matplotlib

Jesús Roberto Araujo Sánchez

9 de marzo, 2019

1 Introduction

En esta practica calcularemos los parametros de un conjunto de datos meteorologicos segun el modelo de Utah, con el fin de entender mejor los datos.

El primer parametro utilizado es de la universidad de Utah y se basa en las unidades de frio para poder estudiar las condiciones del sembradio.

El segundo parametro es de la universidad de Gragenda y se basa en el estudio de horas frias, esto es estudiar el sembradio cada hora para ver en que horas fue beneficiado o afectado.

Lo que se busca es observar los resultados de ambos metodos y ver que diferencias presentan entre ellos.

2 Desarrollo

Como ya se menciono antes utilizamos los modelos de Utah y de Gragenda.

El modelo de Utah consiste en asignarle un cierto valor de unidad frio a cada hora. La siguiente tabla muestra los valores:

Temperatura °C	Unidades de frío correspondientes
$< 1,4^{\circ}C$	0
$1,5^{\circ}C$ a $2,4^{\circ}C$	0.5
$2,5^{\circ}C$ a $9,1^{\circ}C$	1.0
$9,2^{\circ}C$ a $12,4^{\circ}C$	0.5
$12,5^{\circ}C$ a $15,9^{\circ}C$	0
$16,0^{\circ}C$ a $18,0^{\circ}C$	-0.5
$> 18,0^{\circ}C$	-1.0

Cuadro 1: Parámetros del método de Utah.

Por medio de estos valores asignados al final del día se puede hacer la suma y verificar si el clima es propicio para el crecimiento de vegetación en dicho sitio.

El método de Grageda consiste en verificar si una hora es propicia o fría, o si es nociva o caliente. Una vez identificadas las horas de cada día se hace la suma de horas frías para ver que tan propicio es el clima en ese día.

Tipo de hora	Temperatura °C
Hora frío	0°C a 10,0°C
Hora calor	> 25,0°C
Horas de frío efectivas	Horas frío - horas calor

Cuadro 2: Parámetros del método de Grageda Grageda.

3 Procedimiento

Primero que nada, importamos las siguientes librerías:

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

A continuación se introdujo el archivo con los datos meteorológicos llamado "vid18 180219" y se convirtió en un dataframe para poder trabajar con él como a continuación se muestra:

```
datos_df = pd.read_csv("vid18_180219.dat", delimiter=',', sep='\s+')
df = pd.DataFrame(datos_df)
```

Y para obtener valores por hora se transformó la variable a fecha mediante los siguientes comandos:

```
df['HORA'] = df['FECHA'].dt.hour
df['DÍA'] = df['FECHA'].dt.day
df['MES'] = df['FECHA'].dt.month
df['AÑO'] = df['FECHA'].dt.year
```

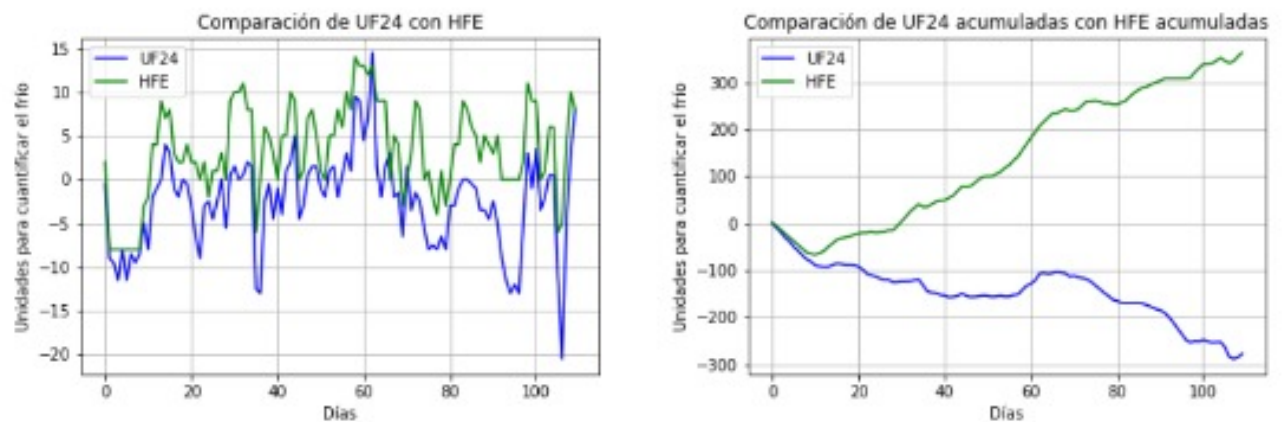
Después aplicamos funciones que nos permitieran obtener los valores por hora y también un promedio acomodados en una tabla mediante los siguientes comandos:

```
uf_df["TPROM"] = uf_df.groupby(['AÑO', 'MES', 'DÍA', 'HORA']).transform(np.mean)
comp_df['HFETOT'] = np.cumsum(comp_df['HFE'])
```

Y una vez que obtuvimos los resultados aplicamos los parametros a nuestro dataframe.

4 Resultados

A continuacion se muestran dos graficas correspondientes al modelo de Utah y al de Graciosa:



Las graficas anteriores muestran la evolución temporal de nuestros datos.

Podemos ver que mientras las unidades frío y las horas frío presentan cierta semejanza, no es si no hasta que las comparamos en forma acumulada que se aprecia una correlacion inversa entre ellas.

5 Conclusions

Ambos metodos tienen lo suyo, y es difícil escoger entre uno de ellos, sin embargo el usarlos en conjunto nos da información valiosa sobre los datos y es ahí cuando podemos hacer mejores predicciones, pues al ser el clima tan caótico no queda más que analizar los datos desde muchos ángulos para poder entender a profundidad su comportamiento.