

Preparando Datos Con Emacs

Ramses Pacheco Ortiz

6 De marzo Del 2018

1 Introduccion

Esta sección es un curso intensivo sobre el editor de textos emacs, en donde la actividad anterior lo utilizamos para descargar unos datos de sondeos y para la realización de los scripts, en esta ocasión utilizaremos emacs para editar el archivo de datos descargados en la actividad anterior (Introducción a la programación de los intérpretes de comandos), cambiando las fechas por números de meses, modificando los espacios, etc.

Emacs es un editor de texto con una gran cantidad de funciones, muy popular entre programadores y usuarios técnicos. GNU Emacs es parte del proyecto GNU y la versión más popular de Emacs con una gran actividad en su desarrollo. El Emacs original significa, **E**ditor **MAC**ro**S** para el **TECO**. Fue escrito en 1975 por Richard Stallman junto con Guy Steele. Fue inspirado por las ideas de TECMAC y TMACS, un par de editores TECO-macro escritos por Guy Steele, Dave Moon, Richard Greenblatt y otros. Se han lanzado muchas versiones de EMACS hasta el momento, pero actualmente hay dos que son usadas comúnmente:

GNU Emacs, iniciado por Richard Stallman en 1984, y **XEmacs**, una fork de GNU Emacs, que fue iniciado en 1991. Ambos usan una extensión de lenguaje muy poderosa, Emacs Lisp, que permite manejar tareas distintas, desde escribir y compilar programas hasta navegar en Internet. GNU Emacs es mantenido por el Proyecto GNU Emacs, el cual cuenta entre sus miembros a Richard Stallman.

2 Conceptos CAPE y PW

- CAPE

Uno de los problemas en la estimación de la inestabilidad atmosférica (uno de los ingredientes necesarios para el desarrollo de la convección, junto con la humedad y mecanismo de disparo que inicie el ascenso de la masa de aire) es su cuantificación y su aplicabilidad en la realidad diaria. Para poder llegar a este punto vamos a presentar uno de los índices más populares y usados para tal fin, la CAPE (Convective Available Potential Energy, en inglés) o energía potencial disponible convectiva. La CAPE se

mide en Julios por kilogramos, J/kg, y estrictamente hablando no es un índice de inestabilidad propiamente dicho como otros índices más clásicos.

a CAPE representa la cantidad de energía de flotabilidad disponible de una parcela o burbuja de aire que se eleva vertical y libremente” o “la cantidad de trabajo que una burbuja o parcela de aire hace en el medio ambiente al ascender libremente”.

- PW La distribución y evolución del vapor de agua es determinante en el funcionamiento climático y meteorológico de la atmosfera terrestre. Hasta hace poco tiempo, esta variable atmosférica no estaba bien definida al no disponerse de medidas con la suficiente resolución para representar su alta inestabilidad espacial y temporal. El incremento en los últimos años de estaciones de referencia permanentes GNSS a escala mundial, supone un importante avance para la monitorización del vapor de agua o cantidad de **agua precipitable** (PW). El objetivo de esta Tesis es analizar el comportamiento del contenido de vapor de agua atmosférico determinado a partir de las observaciones GNSS realizadas en las estaciones de referencia localizadas en la Comunidad Valenciana, estableciendo un sistema de observación y monitorización atmosférica en la zona de estudio, donde se han establecido las relaciones del PW con las variables meteorológicas presión atmosférica y precipitación, evaluando la tendencia y variabilidad del vapor de agua atmosférico cuando se han sucedido eventos de lluvias intensas.

3 Descripción Del Proceso De Limpieza Y Preparación De Los Datos

Anteriormente a esta actividad descargamos unos datos de un país de nuestra elección,unas de las clasificaciones que tienen estos datos son la temperatura, presión, altura, humedad, velocidad en nudos,etc.

Con ayuda de emacs modificaremos el archivo antes mencionado, eliminando clasificaciones y dejando solamente 2 las cuales llevan el nombre de CAPE Y PW,los cuales lo guardamos en otro archivo llamado df2017CAPE_PW.csv A continuación se le muestra el código con el cual se llevo la acción mencionada:

```
#!/bin/bash
```

```
# Cambiar la estación
```

```
# La Grande 71823
```

```
#cat sounding?* > sondeos.txt
```

```
egrep -v 'PRES|hPa' sondeos.txt | egrep '71823|CAPE|Precip' > df2017CAPE_PW_00Z.csv
```

Despues creamos otros dos archivos a partir del archivo creado anteriormente, llamandolos a uno de ellos como df2017CAPE_PW_00Z.csv y otro df2017CAPE_PW_12Z.csv, en la siguiente imagen muestra como los creamos y los datos de los archivos:

A	B	C
12Z 01 Jan 2017	0.00	1.26
12Z 02 Jan 2017	0.00	2.16
12Z 03 Jan 2017	0.00	2.27
12Z 04 Jan 2017	0.00	7.75
12Z 05 Jan 2017	0.00	5.32
12Z 06 Jan 2017	0.00	3.34
12Z 07 Jan 2017	0.00	1.48
12Z 08 Jan 2017	0.00	1.59
12Z 09 Jan 2017	0.00	1.63
12Z 10 Jan 2017	0.00	1.77
12Z 11 Jan 2017	0.00	10.20
12Z 12 Jan 2017	0.00	1.14
12Z 13 Jan 2017	0.00	1.53
12Z 14 Jan 2017	0.00	1.31
12Z 15 Jan 2017	0.00	1.51
12Z 16 Jan 2017	0.00	4.69
12Z 17 Jan 2017	0.00	2.25
12Z 18 Jan 2017	0.00	7.59
12Z 19 Jan 2017	0.00	7.41
12Z 20 Jan 2017	0.00	11.96
12Z 21 Jan 2017	0.00	6.98
12Z 22 Jan 2017	0.00	2.28
12Z 23 Jan 2017	0.00	2.22
12Z 24 Jan 2017	0.00	1.85

```

ramses@ltsp70:~/computacional1/Actividad5$ ./act5.sh
ramses@ltsp70:~/computacional1/Actividad5$ ./act5.sh
ramses@ltsp70:~/computacional1/Actividad5$ emacs df2017CAPE_PW.csv
ramses@ltsp70:~/computacional1/Actividad5$ egrep -v '00Z' df2017CAPE_PW.csv > ^C
ramses@ltsp70:~/computacional1/Actividad5$ egrep -v '00Z' df2017CAPE_PW.csv > df2017CAPE_PW_12Z.csv
ramses@ltsp70:~/computacional1/Actividad5$ ls
act5.sh          df2017CAPE_PW.csv  sondeos.txt
df2017CAPE_PW_12Z.csv  df2017CAPE_PW.csv~
ramses@ltsp70:~/computacional1/Actividad5$ egrep -v '12Z' df2017CAPE_PW.csv > df2017CAPE_PW_00Z.csv
ramses@ltsp70:~/computacional1/Actividad5$ ls
act5.sh          df2017CAPE_PW_12Z.csv  df2017CAPE_PW.csv~
df2017CAPE_PW_00Z.csv  df2017CAPE_PW.csv      sondeos.txt
ramses@ltsp70:~/computacional1/Actividad5$ █

```

Ya que obtuvimos el archivo con los datos necesarios, lo abrimos con la aplicacion emacs para hacer las modificaciones, unas de las modificaciones es eliminar el 12z y el jan cambiarlo con numero respecto al mes por ejemplo en este caso del jan seria (01).para llevar acabo esto utilizamos unos comando para seleccionar la palabra comun que queremos cambiar y despues remplazando el numero correspondiente al mes, la imagen siguiente muestra el resultado final.

```

16 01 2017, 0.00, 2.55
17 01 2017, 0.00, 1.64
18 01 2017, 0.00, 2.44
19 01 2017, 0.00, 8.73
20 01 2017, 0.00, 4.68
21 01 2017, 0.00, 10.67
22 01 2017, 0.83, 3.59
23 01 2017, 0.00, 2.86
24 01 2017, 0.00, 2.54
25 01 2017, 0.00, 2.65
26 01 2017, 0.00, 12.01
27 01 2017, 0.00, 9.13
28 01 2017, 0.00, 7.63
29 01 2017, 0.00, 6.54
30 01 2017, 0.00, 3.32
31 01 2017, 0.00, 2.58
01 Feb 2017, 0.00, 1.71
02 Feb 2017, 0.00, 3.25
03 Feb 2017, 0.00, 1.07
04 Feb 2017, 0.00, 1.16
05 Feb 2017, 0.00, 1.43
06 Feb 2017, 0.00, 1.65
07 Feb 2017, 0.00, 1.10
08 Feb 2017, 0.00, 1.50
09 Feb 2017, 0.00, 2.03
10 Feb 2017, 0.00, 1.32
11 Feb 2017, 0.00, 1.09
12 Feb 2017, 0.00, 2.01

```

	A	B	C
01 01 2017	0.00	1.55	
02 01 2017	0.00	4.23	
03 01 2017	0.00	2.15	
04 01 2017	0.00	2.33	
05 01 2017	0.00	7.13	
06 01 2017	0.00	3.37	
07 01 2017	0.00	2.57	
08 01 2017	0.00	1.44	
09 01 2017	0.00	1.47	
10 01 2017	0.00	1.66	
11 01 2017	0.00	5.62	
12 01 2017	0.00	3.54	
13 01 2017	0.00	1.75	
14 01 2017	0.00	1.48	
15 01 2017	0.00	1.90	
16 01 2017	0.00	2.55	
17 01 2017	0.00	1.64	
18 01 2017	0.00	2.44	
19 01 2017	0.00	8.73	
20 01 2017	0.00	4.68	
21 01 2017	0.00	10.67	
22 01 2017	0.83	3.59	
23 01 2017	0.00	2.86	
24 01 2017	0.00	2.54	

2017CAPE_PW_00Z.csv

ahora que obtuvimos los dos archivos separados, sin mes por escrito y con las modificaciones realizadas anteriormente procederemos a pasarlas a la librería pandas para posteriormente leer los archivos, en la siguiente imagen se muestra el código utilizado.

```
In [1]: import pandas as pd
import numpy as np
from datetime import datetime
```

```
In [2]: # Leer archivo de datos
# Convertir la columna CAPE de objeto a número
df = pd.read_csv("df2017CAPE_PW_00Z.csv", header=None, names=['Date', 'CAPE', 'PW'])
df.CAPE = pd.to_numeric(df.CAPE, errors='coerce')
df.head()
```

```
Out[2]:
```

	Date	CAPE	PW
0	01 01 2017	0.0	1.55
1	02 01 2017	0.0	4.23
2	03 01 2017	0.0	2.15
3	04 01 2017	0.0	2.33
4	05 01 2017	0.0	7.13

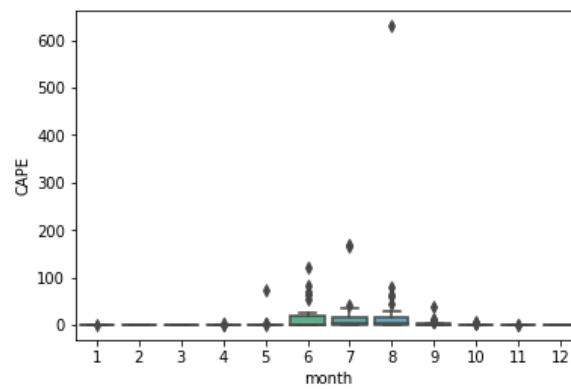
```
In [3]: # Convertir la cadena de caracteres 'Date' en variable temporal 'NDate'
df['NDate'] = pd.to_datetime(df['Date'], format='%d %m %Y')
df['month'] = df['NDate'].dt.month
df.head()
```

```
Out[3]:
```

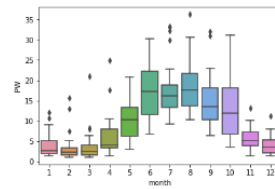
	Date	CAPE	PW	NDate	month
0	01 01 2017	0.0	1.55	2017-01-01	1
1	02 01 2017	0.0	4.23	2017-01-02	1
2	03 01 2017	0.0	2.15	2017-01-03	1
3	04 01 2017	0.0	2.33	2017-01-04	1
4	05 01 2017	0.0	7.13	2017-01-05	1

Ya que leemos los datos que hibamos a utilizar y los caracteres necesarios,proseguiremos a realizar una boxplot con los datos, acontinuacion se muestra la grafica de caja de los dos meses mostrados junio y diciembre:

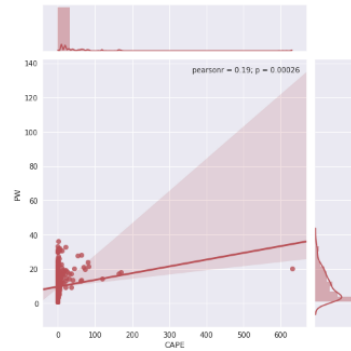
```
In [4]: # graficar Boxplots por mes
# Biblioteca Seaborn
import seaborn as sns
import matplotlib.pyplot as plt
ax = sns.boxplot(x="month", y="CAPE", data=df)
plt.show()
```



```
# graficar Boxplots por mes
# Biblioteca Seaborn
import seaborn as sns
import matplotlib.pyplot as plt
ax = sns.boxplot(x="month", y="PW", data=df)
plt.show()
```

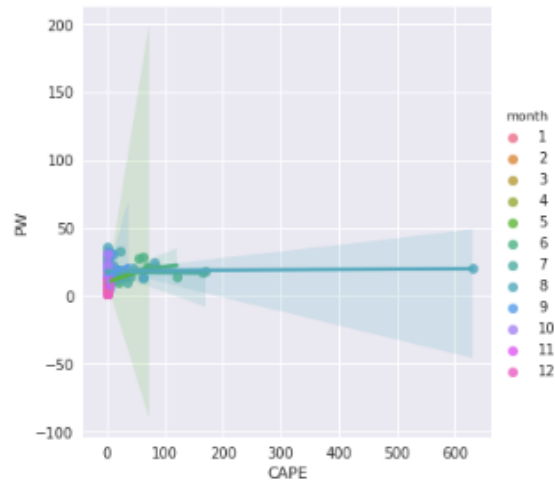


```
In [6]: import seaborn as sns
sns.set(style="darkgrid", color_codes=True)
g = sns.jointplot("CAPE", "PW", data=df, kind="reg",
                  color="r", size=7)
plt.show(g)
```



por ultimo mostramos una grafica mas gradual del mes:

```
In [7]: g = sns.lmplot(x="CAPE", y="PW", hue="month",
                      truncate=True, size=5, data=df)
plt.show(g)
```



4 Conclusiones

Me paracio mu interesante esta practica ya que aprendimos mas sobre emacs y como editar textos, nos mostro como es una herramienta muy util y conocimos ciertos comando que nos facilitara el uso de este editor cuando tengamos muchos datos.

por otro lado aprendimos algo nuevo en python, algo que nunca habiamos visto que son las graficas de caja, me paracio muy interesante aunque, no busque informacion mas a fondo sobre estas graficas.

5 Bibliografia

- Curso sobre interpretación de mapas meteorológicos: la CAPE (Convective Available Potencial Energy) - Revista del Aficionado a la Meteorología. (2018). Revista del Aficionado a la Meteorología.<https://www.tiempo.com/ram/800/curso-sobre-interpretacin-de-mapas-meteorolgicos-el-cape-convective-available-potencial-energy/>
- Emacs. (2018). Es.wikipedia.org. <https://es.wikipedia.org/wiki/Emacs>
- antos, J. (2018). Análisis y predicción de lluvias intensas en la Comunidad Valenciana basados en la estimación del contenido de vapor de agua atmosférico obtenido con técnicas GNSS. Dialnet.unirioja.es.<https://dialnet.unirioja.es/servlet/tesis>

6 Apéndice

¿Cómo se te hizo esta actividad? ¿Compleja, Difícil, Sencilla?

Un poco difícil y confusa , ya que estudiamos algo nuevo que nunca habíamos realizado prácticamente

¿Qué te llamó más la atención?

Me llamo mucho la atencion las graficas de caja y el editor de emacs, como nos facilitara la vida. **¿Qué parte fue la que menos te interesó hacer?**

ninguna

¿Cómo mejorarías esta actividad? ¿Qué le faltó? ¿Qué sobró?

Talvez con un poco mas de instrucciones en las actividades a realizar.

¿Hasta este punto, que te parece el uso de Jupyter para programar en Python?

Bien, la plataforma es como para trabajar y facil de comprender.