



**El saber de mis hijos
hará mi grandeza**

UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

DEPARTAMENTO DE FÍSICA

LICENCIATURA EN FÍSICA

FÍSICA COMPUTACIONAL I

Reporte de actividad 5

Ricardo Ruiz Hernández

Profesor del curso
Dr. Carlos Lizárraga Celaya

07 de marzo del 2018

1. Introducción

Esta práctica tuvo como principal objetivo que, de manera intensiva, como se menciona, aprendamos a manejar Emacs en la edición de archivos, con el claro objetivo de hacernos acreedores de la habilidad para limpiar datos, muy importante dentro de esta área del conocimiento.

2. Conceptos

- CAPE: *ES una herramienta que se utiliza para determinar en todo momento y lugar, cuan potenciales son los climas severos. Estas siglas corresponden a la cantidad de trabajo que hace un parcel en el ambiente*
- PW: *Agua precipitable; expresa la cantidad de agua en términos de altura o masa.*

3. Procesamiento y limpieza de los datos

En el sitio de la universidad de Wyoming obtuve los datos de Smolensk, Rusia.

- Se realizó un script, con el que se creó *df2017*.
- El siguiente paso fue filtrar los datos, mediante el comando *grep*, solo dejamos los datos que contuvieran las palabras "CAPE", "PRECIP", la fecha.

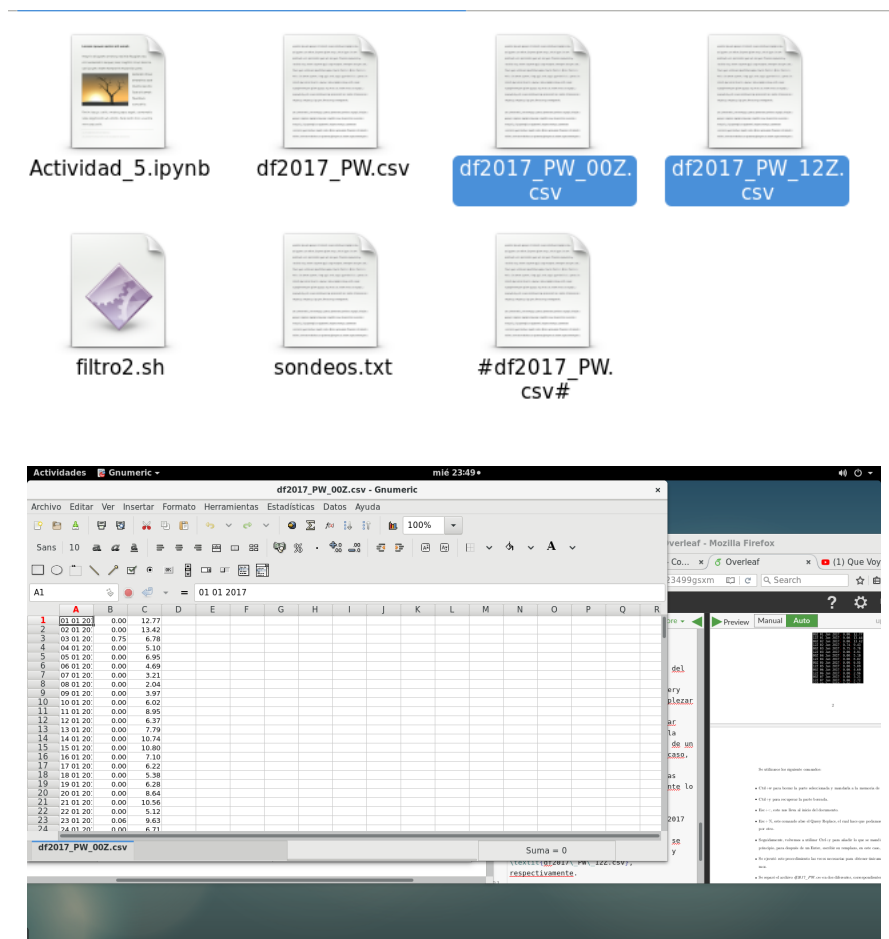
Es así que se creó un archivo al que se le llamó *df2017_PW.csv*

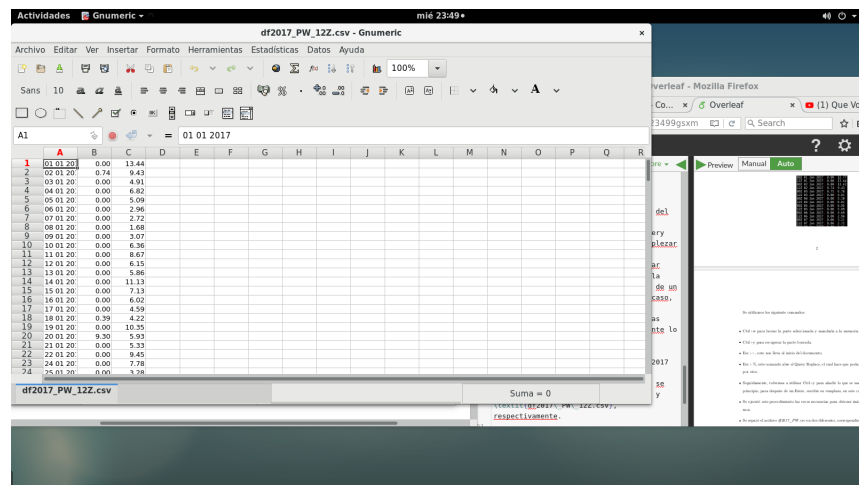
- El archivo antes mencionado contenía palabras, o información que no nos sería útil, para lo que usamos Emacs. Así por medio de distintos comandos, limpiamos, quedándonos únicamente con lo numérico de lo que nos interesaba.

```
00Z 01 Jan 2017, 0.00, 12.77
12Z 01 Jan 2017, 0.00, 13.44
00Z 02 Jan 2017, 0.00, 13.42
12Z 02 Jan 2017, 0.74, 9.43
00Z 03 Jan 2017, 0.75, 6.78
12Z 03 Jan 2017, 0.00, 4.91
00Z 04 Jan 2017, 0.00, 5.10
12Z 04 Jan 2017, 0.00, 6.82
00Z 05 Jan 2017, 0.00, 6.95
12Z 05 Jan 2017, 0.00, 5.09
00Z 06 Jan 2017, 0.00, 4.69
12Z 06 Jan 2017, 0.00, 2.96
00Z 07 Jan 2017, 0.00, 3.21
12Z 07 Jan 2017, 0.00, 2.72
00Z 08 Jan 2017, 0.00, 2.64
```

Se utilizaros los siguiente comandos:

- Ctrl+w para borrar la parte seleccionada y mandarla a la memoria de igual manera.
- Ctrl+y para recuperar la parte borrada.
- Esc+<, esto nos lleva al inicio del documento.
- Esc+%, este comando abre el Query Replace, el cual hace que podamos reemplazar un código por otro.
- Seguidamente, volvemos a utilizar Ctrl+y para añadir lo que se mandó a la memoria en un principio, para después de un Enter, escribir su remplazo, en este caso, nada.
- Se ejecutó este procedimiento las veces necesarias para obtener únicamente lo que requeríamos.
- Se separó el archivo *df2017_PW.csv* en dos diferentes, correspondientes a 00Z y a 12Z; estos se llamaron *df2017_PW_00Z.csv* y *df2017_PW_12Z.csv*, respectivamente.





4. Análisis de datos

Aquí le damos propósito a lo que hicimos anteriormente.

- Después de cargar las bibliotecas, se leyeron los datos y se le asignó un formato a CAPE.

Una vez esto, se creó una columna con formato de fecha, como se muestra a continuación.

```
# Leer archivo de datos
# Convertir la columna CAPE de objeto a número
df = pd.read_csv("df2017_PW_00Z.csv", header=None, names=['Date', 'CAPE', 'PW'])
df.CAPE=pd.to_numeric(df.CAPE, errors='coerce')
df.head()
```

	Date	CAPE	PW
0	01 01 2017	0.00	12.77
1	02 01 2017	0.00	13.42
2	03 01 2017	0.75	6.78
3	04 01 2017	0.00	5.10
4	05 01 2017	0.00	6.95

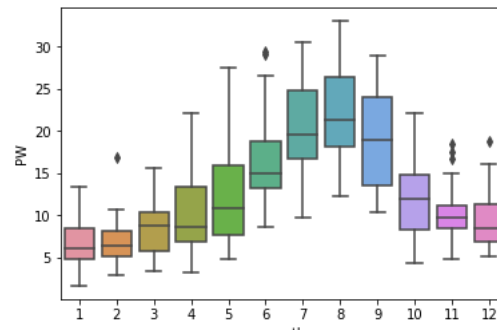
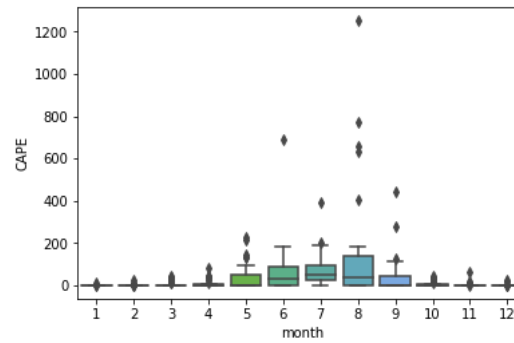
```
# Convertir la cadena de caracteres 'Date' en variable temporal 'NDate'
df['Ndate'] = pd.to_datetime(df['Date'], format='%d %m %Y')
df['month'] = df['Ndate'].dt.month
df.head()
```

	Date	CAPE	PW	Ndate	month
0	01 01 2017	0.00	13.44	2017-01-01	1
1	02 01 2017	0.74	9.43	2017-01-02	1
2	03 01 2017	0.00	4.91	2017-01-03	1
3	04 01 2017	0.00	6.82	2017-01-04	1
4	05 01 2017	0.00	5.09	2017-01-05	1

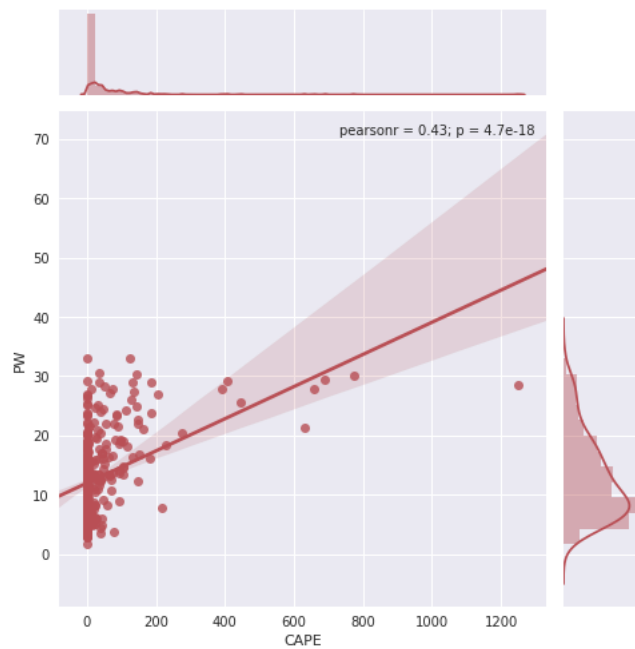
- Se cargaron las bibliotecas seaborn y matplotlib, estas fueron utilizadas para realizar gráficas tipo boxplot, además se compararon las variables (CAPE y PW).

5. Resultados

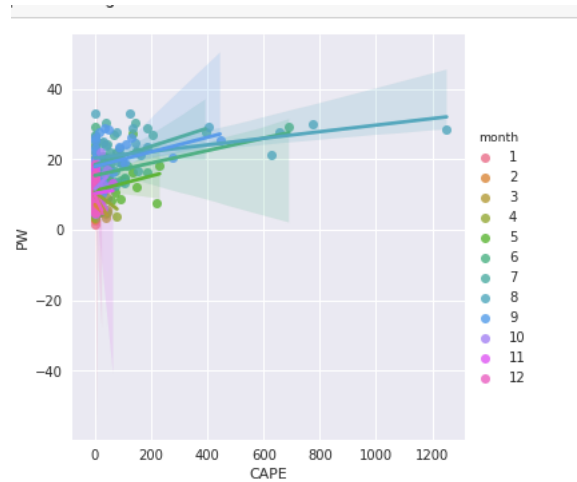
- Gráficas tipo Boxplot de CAPE y PW, respectivamente:



- En la gráfica de Jointplot se expone el coeficiente de correlación de Pearson:



- En Lmplot se compara por cada mes, con el fin de ver como cambia con en cada estación del año:



6. Conclusiones

Mediante esta actividad pude percatarme de la importancia de saber limpiar datos, puesto que en muchas ocasiones, estos se encuentran con información de sobra o que no nos interesa en nuestra investigación. Además, aprendimos a cargar distintas bibliotecas para darle uso a estos datos.

7. Bibliografía

(2008) https://en.wikipedia.org/wiki/Precipitable_water

(2018) https://en.wikipedia.org/wiki/Convective_available_potential_energy

8. Apéndice

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

No me pareció realmente compleja.

2. ¿Qué te llamó más la atención?

Los comandos de Emacs, me pareció algo realmente útil e interesante.

3. ¿Qué parte fue la que menos te interesó hacer?

Francamente, nada.

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

Me pareció muy bien.

5. ¿Hasta este punto, qué te pareció el uso de Jupyter para programar en Python?

Me ha parecido muy ameno, me gusta.