



UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

FÍSICA COMPUTACIONAL

14 DE FEBRERO DEL 2017

Act 3: Iniciándose en Python

Alumna: Chávez Gutiérrez Yanneth Tzitzin

Profesor: Carlos Lizárraga Celaya.

1. Breve resumen

En este reporte se introdujo lo que es Python, el cual es un lenguaje de programación interpretado. Se nos explico cómo hacer uso de él mediante jupyter notebooks y se busco como utilizarlo para el analisis de datos en la práctica.

2. Introducción

Con base a los datos descargados del 2016 en analisis de la atmósfera en la ciudad de Chihuahua, se utilizaron comandos de la terminal para extraer sólo los valores 00Z y 12Z del archivo anual, de las siguientes variables: CAPE, cantidad de Agua Precipitable, y temperatura. Despues de hacer uso de esto, se utilizo Emacs para limpiar nuestro archivo y tenerlo listo para editarlo con jupyter notebooks.

3. Resultados

3.1. Tablas de datos

Al tener los datos listos para abrirlos en python, abrimos la terminal para correr jupyter notebook. De aqui nos abrirá una pestaña en el navegador y tenemos que correr los sig comandos para importar las bibliotecas que vamos a utilizar.

```
import pandas as pd
import numpy as np
import matplotlib as plt
df = pd.read_csv("/home/01010100/TZITZIN/Computacional/Act_3/12zanual.csv",
names=['Fecha', 'CAPE', 'PW'])
```

De esta forma anunciamos el archivo a utilizar y lo que representa cada columna en el archivo .csv

Y aplicando las siguientes funciones básicas (data frame) se describen las salidas a la tabla:

```
df.head(20),
df.describe()
```

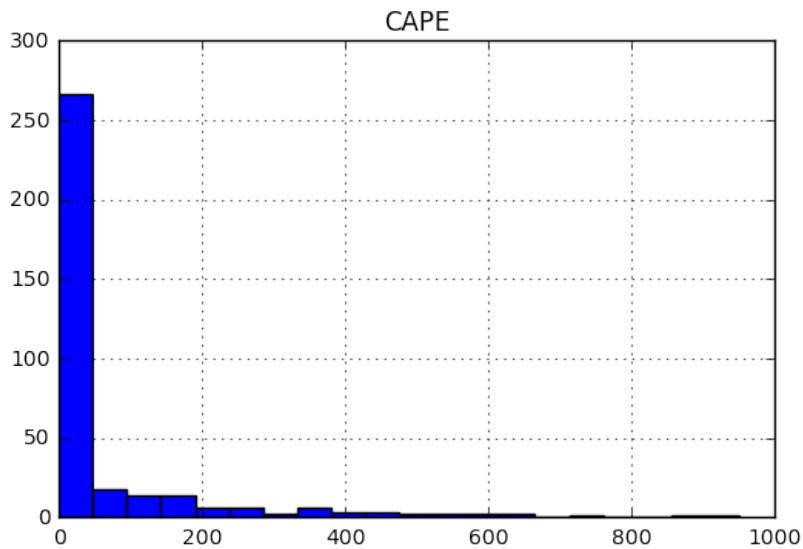
Y usando `df.apply(lambda x: sum(x.isnull()),axis=0)` en Python, encontraremos los valores faltantes en el archivo.

Generandose las siguientes tablas:

	Fecha	CAPE	PW	month
0	01 Jan 2016	0.0	13.76	1
1	02 Jan 2016	0.0	16.36	1
2	04 Jan 2016	0.0	12.39	1
3	05 Jan 2016	0.0	13.42	1
4	06 Jan 2016	0.0	7.56	1
5	07 Jan 2016	0.0	4.40	1
6	08 Jan 2016	0.0	4.94	1
7	09 Jan 2016	0.0	5.73	1
8	11 Jan 2016	0.0	6.90	1
9	12 Jan 2016	0.0	5.01	1
10	13 Jan 2016	0.0	7.49	1

	CAPE	PW
count	349.000000	349.000000
mean	60.128711	16.874756
std	142.369995	10.690086
min	0.000000	2.080000
25%	0.000000	7.020000
50%	0.000000	14.900000
75%	30.930000	26.240000
max	950.050000	41.120000

4. Variables CAPE

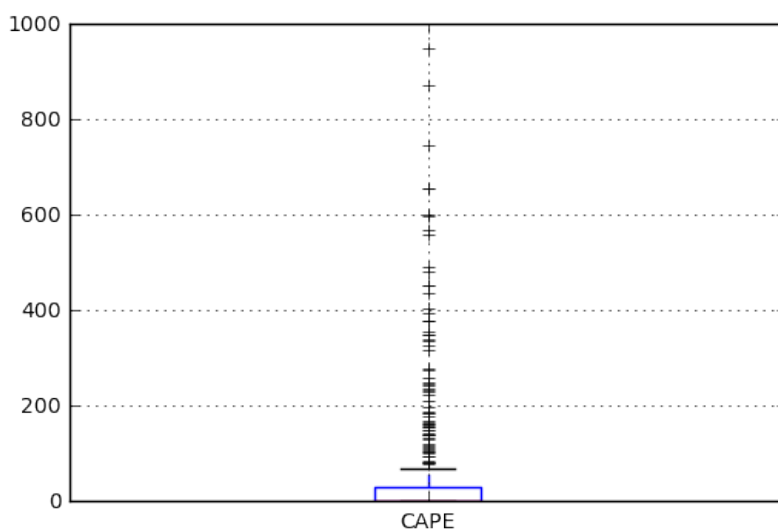


Se creo la gráfica con los sig comandos:

```
df.columns
matplotlib inline
df.hist(u'CAPE',bins=20)
```

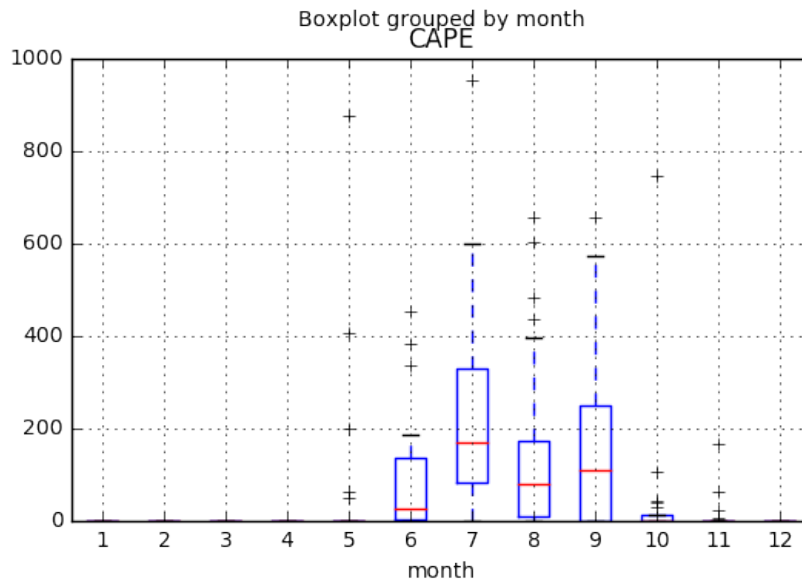
Para generar el diagrama de caja, se utilizo el sig comando:

```
df.boxplot(column=u'CAPE')
```



Graficando el diagrama de caja para cada mes con el sig comando: obtenemos

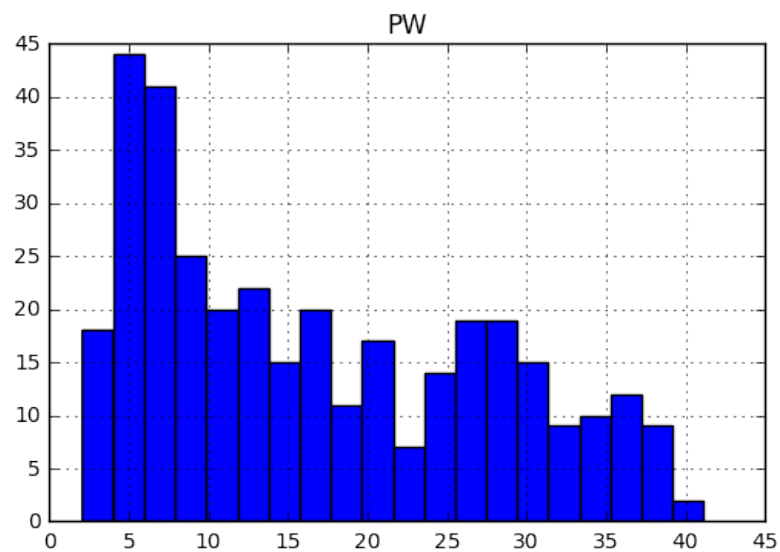
```
df.boxplot(column=u'CAPE',by=u'month')
```



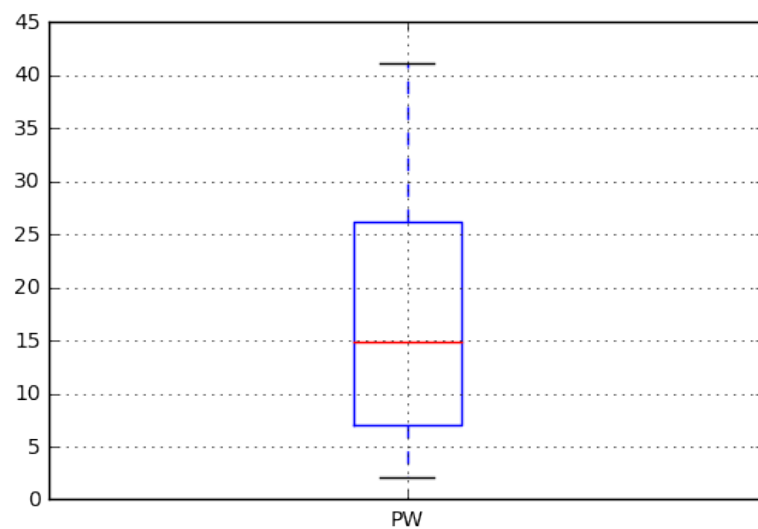
Mediante esta grafica nos podemos dar cuenta de por qué los datos de caja del año salen tan feos, los primeros meses tienen datos de 0 y hasta Junio se presentan datos CAPE.

5. Precipitable water

Gráfico de la variable Precipitable Water (PW)

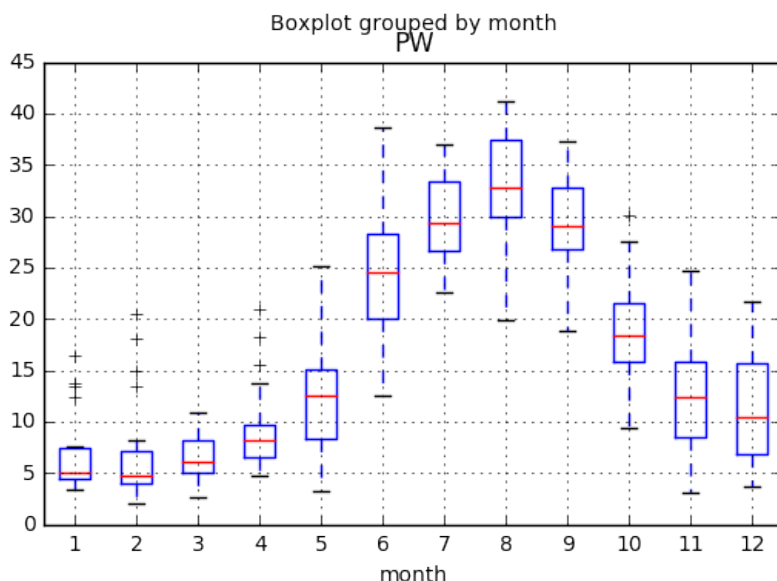


y graficando el diagrama de caja para PW tenemos:



Graficando el diagrama de caja para cada mes con el sig comando: obtenemos

```
df.boxplot(column=u'PW',by=u'month')
```



Para PW se obtuvieron más datos, fue constante y por eso podemos observar esta distribución y apreciar el comportamiento de los datos y como por los meses de junio a septiembre hubo más actividad.

6. Conclusión

En esta práctica se analizaron los datos del archivo 12anual.csv y con ayuda de Python y series de comandos, poder interpretar esos datos mediante representaciones gráficas. Utilizando el comando sed para organizar nuestros datos en la terminal, fue más fácil. Ya que sed se encarga de buscar cierta palabra clave, y con esas palabras identificadas en el archivo, sería más fácil eliminar los datos de esas secciones. Una vez eliminadas, se abrió el archivo en Emacs para por una serie de comandos reemplazar las partes del documento que nos eran innecesarias para solo dejar las columnas con los datos de los días de CAPE y PW.

Al terminar, se utilizó Python para el manejo de tablas y gráficas de estos datos. Utilizando los comandos que se describen en la práctica, se generaron. Una observación, es que al hacer el diagrama de caja de los datos CAPE anuales, se observó mucha dispersión y es por esto que se muestran en su mayoría puntos negros que representan ceros en nuestra muestra.

7. Referencias

- (1) <http://weather.uwyo.edu/upperair/sounding.html>
- (2) [http://computacional1.pbworks.com/w/page/115266988/Actividad %203 %20\(2017-1\)](http://computacional1.pbworks.com/w/page/115266988/Actividad%203%20(2017-1))
- (3) <http://stackoverflow.com/questions/4396974/sed-or-awk-delete-n-lines-following-a-pattern>