



Universidad de Sonora  
Departamento de Física



---

# ANÁLISIS DE ÍNDICES ATMOSFÉRICOS USANDO PYTHON

J. Ernesto Torres Burruel

15 de febrero de 2017

Física Computacional  
Carlos Lizárraga Celaya

---

## Resumen

esta práctica se hace uso de Python para realizar tablas y gráficos sobre los datos obtenidos de sondeos atmosféricos realizados en El Paso, Texas. Se realizará un análisis estadístico de los datos obtenidos sobre la Energía Potencial Convectiva Disponible (CAPE, por sus siglas en inglés) y la cantidad de agua precipitable en la atmósfera.

Los datos se obtuvieron desde la página del Departamento de Ciencias Atmosféricas de la *Universidad de Wyoming* en <http://weather.uwyo.edu/upperair/sounding.html>.

## Índice

<b>1. Estadística y datos</b>	<b>2</b>
<b>2. Índices y parámetros atmosféricos</b>	<b>2</b>
2.1. CAPE . . . . .	2
2.2. Cantidad de agua precipitable . . . . .	6
<b>3. Conclusiones sobre le manejo de los datos</b>	<b>9</b>
<b>4. Referencias</b>	<b>10</b>
<b>5. Apéndice A: Cantidad de datos por mes según hora de sondeo</b>	<b>11</b>

## 1. Estadística y datos

El análisis estadístico nos permite entender, describir y predecir lo que sucede con las características de nuestro interés que posee una muestra o bien toda una población. Para esto se hace uso de diversas herramientas que nos brinda la estadística para describir los datos obtenidos mediante una medición hecha sobre nuestro objeto de estudio.

El manejo de gráficos es una parte importante en la estadística, pues de esta manera podemos representar y encontrar la tendencia que se tiene en los datos y poder realizar aproximaciones para lograr explicar lo que sucede con el objeto. Las gráficas de barras, los histogramas, las gráficas de cajas, etc. son de las herramientas más utilizadas y que brindan una gran cantidad de información sobre los datos. [1]

Para el análisis de los datos obtenidos a partir de los sondeos atmosféricos realizados en El Paso, Texas, se realizará el análisis de dos índices atmosféricos.

## 2. Índices y parámetros atmosféricos

Para explicar el estado en el que se encuentra la atmósfera se hace uso de índices, que son valores únicos e indicativos sobre una situación presente en la atmósfera. Los índices son una idea relativa sobre una condición, en el estudio atmosférico se usan diferentes índices para describirla, es común que estos se relacionen con energías, temperaturas e incluso con la posición en la atmósfera en la que se realiza el sondeo. En esta ocasión el interés es sobre dos valores: la Energía Potencial Convectiva Disponible (CAPE) y la cantidad de agua precipitable. [2]

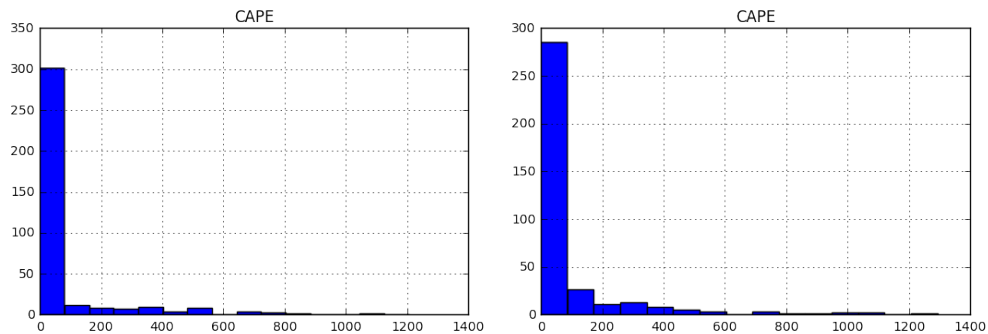
### 2.1. CAPE

Con los datos de los sondeos realizados durante el año 2016 en El Paso, Texas, se obtendrán tablas que muestran datos estadísticos como la media, la desviación estándar, los percentiles y valores máximos y mínimos. Ver Cuadro 1.

Cuadro 1: *Tabla con datos estadísticos de CAPE por hora de sondeo*

	<b>12Z</b>	<b>00Z</b>
<b>Cantidad de datos</b>	363.000000	362.000000
<b>Media</b>	72.564628	75.943122
<b>Desviación Estándar</b>	185.010323	184.319741
<b>Valor mínimo</b>	0.000000	0.000000
<b>25 %</b>	0.000000	0.000000
<b>50 %</b>	0.000000	0.000000
<b>75 %</b>	8.830000	47.070000
<b>Valor máximo</b>	1205.310000	1292.680000

Para analizar más cómo es que se comporta este índice a través de las mediciones hechas en el año se realiza un histograma, el cual secciona en intervalos la muestra, el número de intervalos se aproxima a  $\sqrt{\#datos} \approx 20$ . Con lo que se obtiene el histograma de la Figura 1 Como se ve en las gráficas

Figura 1: *Histograma de mediciones de CAPE a las 12Z (izquierda) y 00Z (derecha) durante un año*

de la Figura 1 hubo muchos valores que estuvieron alrededor del cero, incluso al tabular datos de CAPE se obtiene una gran cantidad de mediciones con valor 0.00 en los datos y hubo pocas mediciones que presentaron un valor diferente para CAPE.

Debido a que los valores que presentan los datos para la Energía Potencial Convectiva Disponible presenta poca cantidad de datos fuera del 0, un diagrama de caja para este índice sería poco apropiado para los datos de un año. Ver Figura 2.

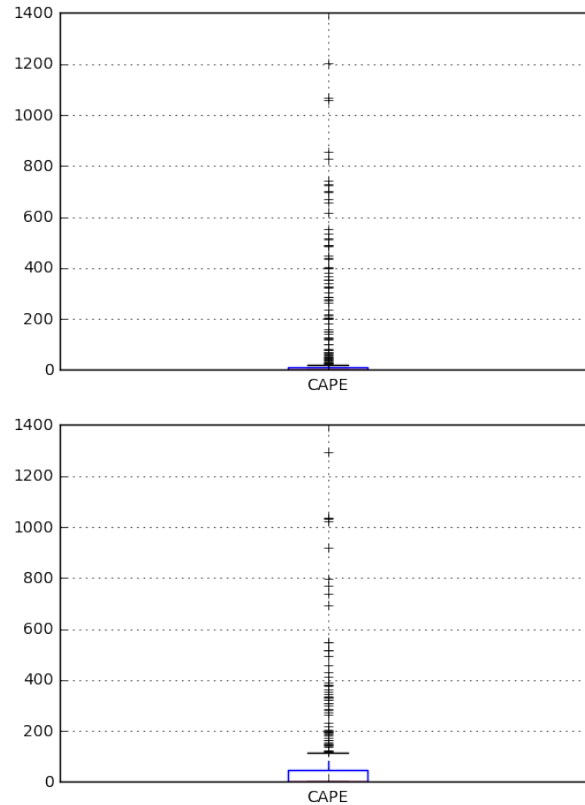


Figura 2: *Diagrama de caja sobre los datos de CAPE a las 12Z (gráfica superior) y 00Z (gráfica inferior) en un año.*

Como las gráficas de caja para el año completo (Figura 2) contiene muchos valores atípicos y es poco informativa se puede separar los datos por los meses del año, de manera que se tiene una mejor visualización para los datos adquiridos del CAPE en nuestro conjunto de datos de un año (ver Figura 3)

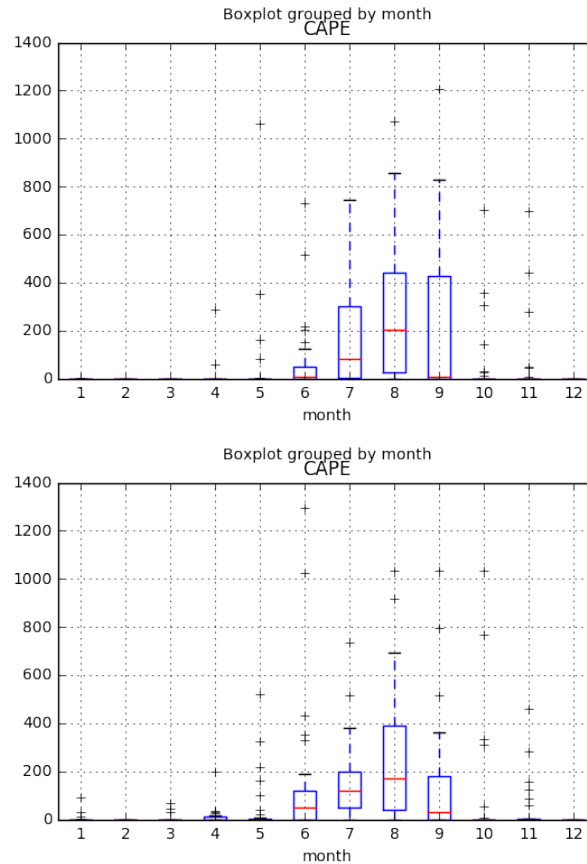


Figura 3: *Diagramas de cajas por mes para los índices CAPE medidos en los sondeos de 12Z y 00Z (superior e inferior, respectivamente)*

Las gráficas de caja ahora muestra mejor la información, pero siguen presentando bastantes datos fuera de la norma, probablemente es porque las mediciones de este índice están centradas en lugares cercanos a 0, pero también contienen datos con valores muy distintos.

## 2.2. Cantidad de agua precipitable

La cantidad de agua precipitable se define como la cantidad de vapor contenida en una columna de aire de la atmósfera [3]. En los sondeos realizados en El Paso, TX. se obtienen diferentes mediciones de este indicador a lo largo del año de 2016, con este se trabaja de la misma manera que para el CAPE. Los datos estadísticos importantes que presenta el conjunto de datos se expresan en la siguiente tabla (Cuadro 2).

Cuadro 2: *Datos estadísticos para cantidad de agua precipitable durante el 2016*

Precipitable	12Z	00Z
Cantidad de datos	363.000000	362.000000
Media	15.830606	14.311823
Desviación estándar	10.131919	9.164505
Valor mínimo	1.940000	1.540000
25 %	7.670000	6.665000
50 %	12.370000	11.430000
75 %	23.735000	21.655000
Valor máximo	43.780000	41.210000

Podemos utilizar histogramas para ver cómo se comportan los datos obtenidos a través del año, separando en los dos horarios (00Z y 12Z), obteniendo así los gráficos en la Figura 4.

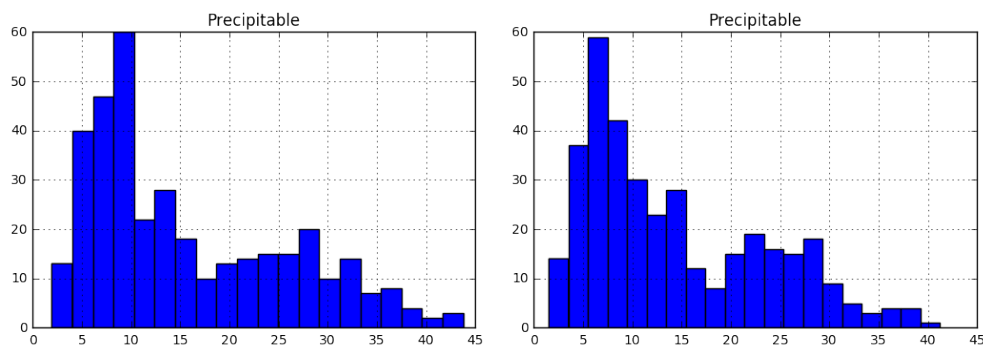


Figura 4: *Histogramas de cantidad de agua precipitable a las 12Z (izquierda) y 00Z (derecha).*

Para el indicador de agua precipitable existe una mejor visualización de los datos que para el CAPE. Si analizamos la cantidad de agua precipitable en el año en un diagrama de caja, veremos que los datos obtenidos se pueden contener muy bien en este tipo de representación y nos da una indicación de los valores y la dispersión de estos en todo el año (Figura 5).

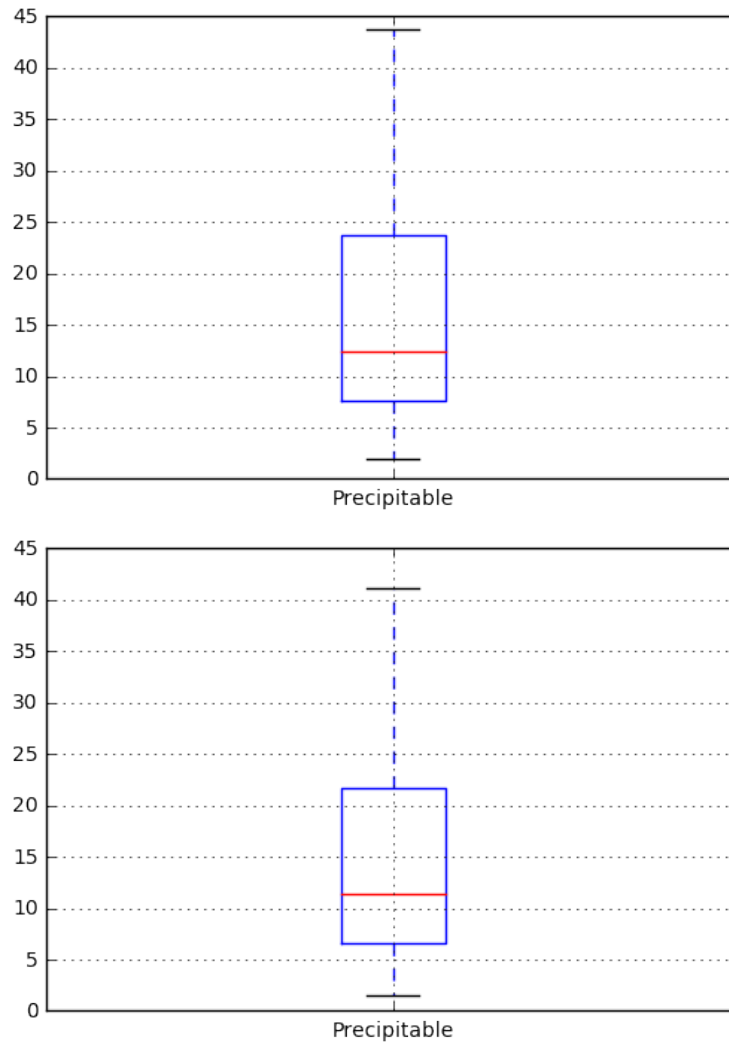


Figura 5: Diagramas de caja de cantidad de agua precipitable observada en los datos de 2016 a las 12Z (superior) y 00Z (inferior)



Para tener una mejor visualización de los datos y la dispersión de la cantidad de agua precipitable en la atmósfera es recomendable analizar por meses del año, pues la cantidad de vapor de agua en la atmósfera varía dependiendo de la época del año y en la temporada de lluvias de la región. Realizando esto se obtiene como resultado la Figura 6

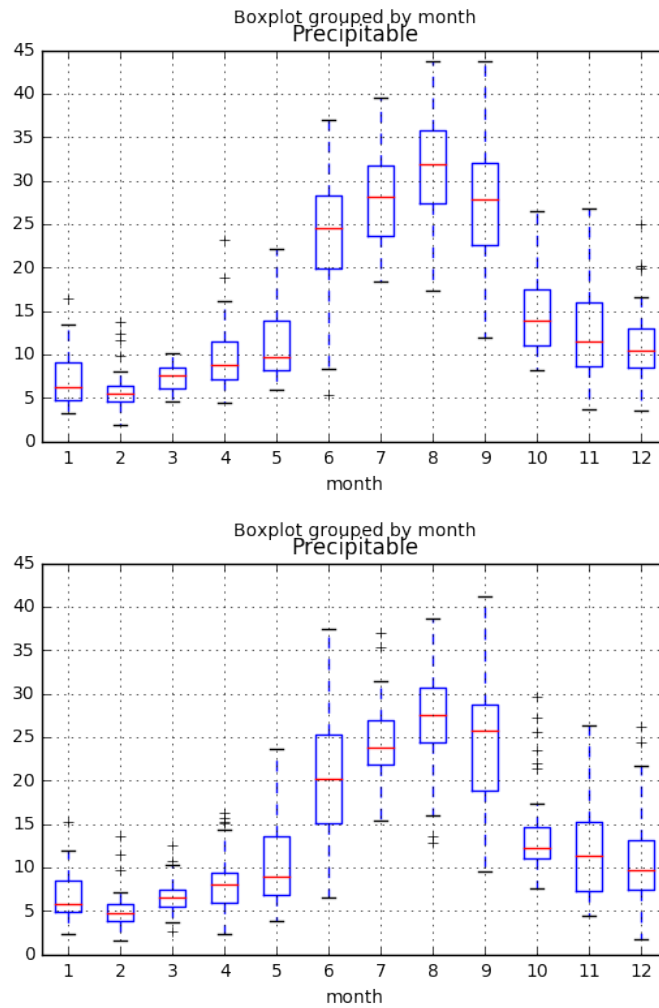


Figura 6: Diagramas de cajas de cantidad de agua precipitable por mes a las 12Z (superior) y 00Z (inferior).

Como se ven en las gráficas los datos para la cantidad de agua precipitable preenta diferentes valores y desviaciones a lo largo del año. Se observa fácilmente que en ciertos meses el agua precipitable es muy poca y tiene una desviación pequeña, en cambio otros meses tiene valores mucho más altos y gran distribución de los datos.

Mezclar los datos en un solo gráfico aumenta la desviación de los datos y sobre todo su dispersión, es mejor seccionar los datos para disminuir los errores y analizar de mejor manera cómo varía este índice atmosférico a lo largo de cada época del año. Gracias a esto se entiende que la cantidad de vapor de agua en la atmósfera es menor en los primeros 5 meses del año, aumenta en los siguientes 4 y vuelve a disminuir en los 3 restantes, por lo que podemos decir que durante el verano aumenta el agua precipitable y disminuye durante el invierno.

### 3. Conclusiones sobre le manejo de los datos

El uso de las herramientas de programación que nos brinda la tecnología actual es bastante práctica al momento de analizar una gran cantidad de datos, aunque en esta ocasión fueron datos relativamente pocos, se puede incluso estudiar datos de varios años y de esta forma manejar una gran cantidad de datos.

Los paquetes que existen para el trabajo en python nos ofrecen una gran gama de herramientas que facilitan la programación con fines científicos para el análisis de datos, en este caso se recurrió a *pandas* para obtener resultados estadísticos de manera fácil, solo programando la lectura de datos y entradas para que el paquete trabaje de forma correcta.

El manejo de python presenta dificultades al inicio, pues quienes no estamos acostumbrados al manejo de este lenguaje de programación y a su forma interpretada resulta bastante incomprensible, pero existen manuales y bastante apoyo en la red que ayuda a sacar el máximo potencial de este lenguaje.

## 4. Referencias

- [1] Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. 8ª edición. Cengage Learning. 2012. Pp. 3-9.
- [2] Blanchard, David O. *Assessing the vertical distribution of convective available potential energy* Journal. Weather and Forecasting. Vol. 13 # 3. 1998. Pp. 870-877.
- [3] American Meteorological Society. *Meteorology Glossary*. Precipitable Water [página web]. Visitado el 16 de Febrero de 2017. [http://glossary.ametsoc.org/wiki/Precipitable\\_water](http://glossary.ametsoc.org/wiki/Precipitable_water)

## 5. Apéndice A: Cantidad de datos por mes según hora de sondeo

Cuadro 3: *Sondeos realizados por mes y hora*

Mes	Días 00Z	Días 12Z
Enero	31	31
Febrero	27	28
Marzo	31	31
Abril	30	30
Mayo	31	31
Junio	29	29
Julio	31	31
Agosto	30	31
Septiembre	30	30
Octubre	31	31
Noviembre	30	30
Diciembre	31	31