

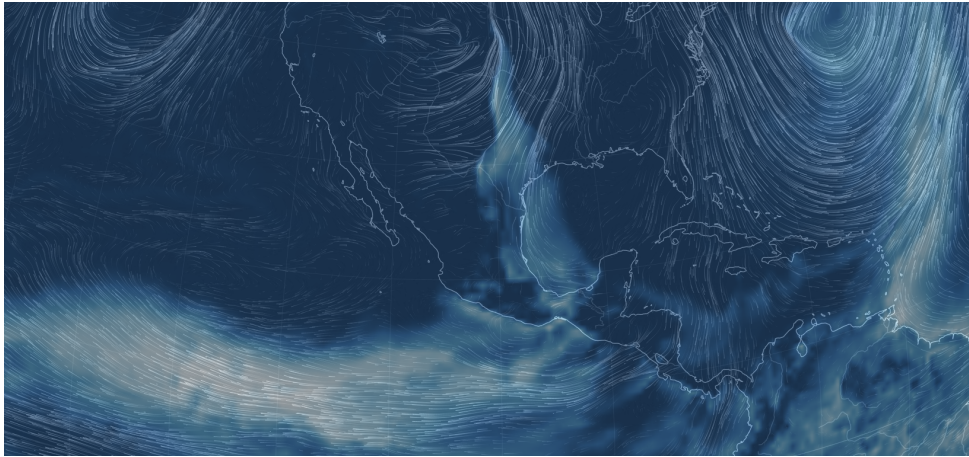
Reporte de Actividad 5

Valenzuela Terán Jonás

March 7, 2018

1 Introducción a la Actividad

Uno de los pasos más importantes en el análisis de datos es la limpieza de ellos, y crear un formato que sea legible por las herramientas que se utilizarán, en este caso, pandas. Para la limpieza se hace uso de emacs, ya que ofrece comandos que permiten automatizar ciertos procesos de edición de texto plano, lo cuál permite cambiar el formato de un archivo extenso con pocas instrucciones, y poco tiempo.



Fuente: <https://earth.nullschool.net/#current/wind/surface/level/orthographic=-93.80,17.09,671>

Los datos tomados describen sondeos de la Atmósfera de la Universidad de Wyoming, se enfocará específicamente en CAPE y PW, con el interés de encontrar una relación entre ambos parámetros.

2 Convective Available Potential Energy y Precipitable Water

2.1 CAPE (Convective Available Potential Energy)

Es la capacidad de energía que una muy pequeña cantidad de aire poseería si fuera elevado por la atmósfera verticalmente. Es útil para indicar inestabilidad atmosférica, que pronostica climas peligrosos.

Es medido en la troposfera, su unidad es Joules por kilogramo (J/Kg), cualquier valor mayor a 0 J/Kg indica inestabilidad y mayor posibilidad de tormentas (Debe existir vapor de agua).

2.2 PW (Precipitable Water)

Comúnmente abreviado como "TPW" = Total Precipitable Water, es la profundidad de agua en una columna de atmósfera, si toda el agua en la columna fuera precipitada en lluvia. Es medida en milímetros.

3 Procesamiento y limpieza de datos

Primero se cargaron los datos de la Universidad de Wyoming, de sondeos atmosféricos de Empalme, Sonora, a las 00:00 y 12:00 horas, usando un script construido en actividades anteriores, una vez obtenidos, en la terminal, se recolecta toda la información en un solo archivo csv, utilizando el comando cat.

Se sigue con la limpieza de datos en emacs, borrando newlines, información que no corresponde con la fecha, CAPE y PW que son de interés, se agregan comas para separar estas categorías y newlines para cada medición, creando un formato legible para pandas usando python, esto es posible gracias a la automatización de acciones que ofrece emacs sobre un patrón indicado, el procedimiento consistió en:

- Seleccionar patrón (Ctrl + SPACE)
- Cortar y pegar información para guardar en memoria (Ctrl + w) y (Ctrl + y)
- Ir al inicio del archivo (Esc + <)
- Entrar al modo del comando (Esc + SHIFT + %)
- Insertar el patrón (Esc + y) (Enter)
- Insertar remplazo del patrón (Insertar reemplazo) (Enter)
- Aplicar (!)

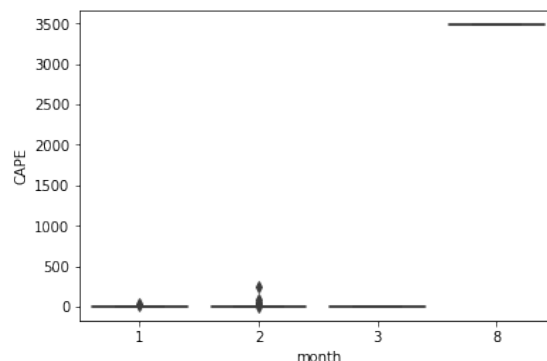
4 Análisis de datos

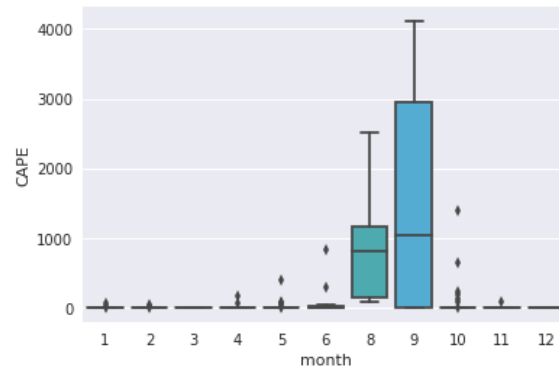
Una vez preparados los archivos, uno para mediciones en 00:00 y otro para 12:00 horas, se leyeron y asignaron a un dataframe en python, usando el entorno de programación jupyter notebook e importando librerías como pandas y numpy para herramientas estadísticas y numéricas, y datetime para detectar correctamente el formato de la fecha.

Esto hace posible la representación gráfica de los datos usando las herramientas disponibles de las bibliotecas, y ver claramente si existe una relación entre CAPE y PW. Realizamos 3 tipos de representación: boxplot, PW contra CAPE, PW contra CAPE con colores indicando mes.

5 Resultados

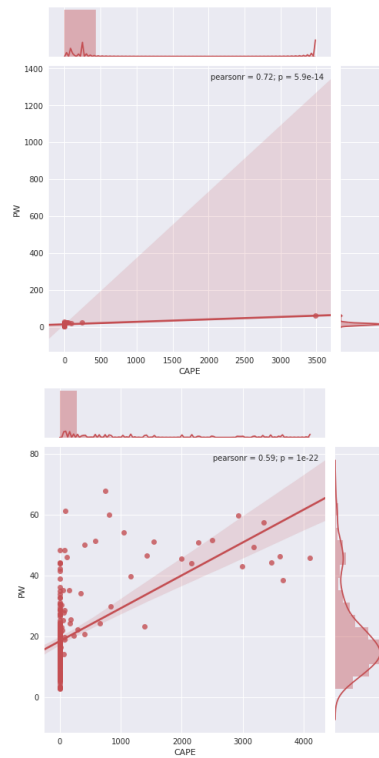
A continuación se muestran las gráficas obtenidas por el programa, la primera imagen corresponde a las mediciones de las 00:00 horas, y la segunda a las 12:00 horas:





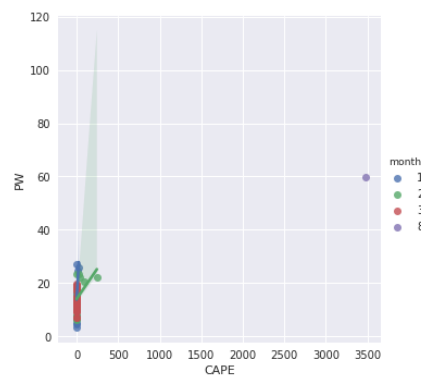
Boxplots de CAPE en cada mes

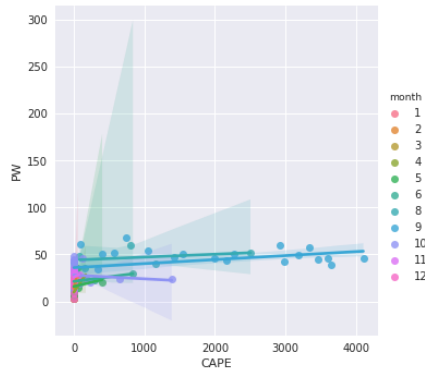
Se observa la ausencia de mediciones a las 00:00 horas. A las 12:00 horas, CAPE se mantiene reducido durante todo el año, a excepción de agosto y septiembre, donde existe una fuerte elevación.



Gráfica PW contra CAPE

Vemos que a las 12:00 horas, CAPE se mantiene relativamente reducido, hasta que PW se eleva por encima de los 40 mm.





Gráfica PW contra CAPE, color distingue cada mes

Al igual que anteriormente, CAPE se eleva cuando PW supera aproximadamente 40 mm, pero muestra información adicional, mostrando que esto ocurre en los meses de agosto y septiembre.

6 Conclusiones

Se enfrentó con la limitación de los datos recibidos, en muchos lugares incompleto, extenso y formatos no compatibles para leer con programas, sin embargo, fue posible obtener información de los datos, realizando una limpieza y adaptación de ellos, y creando gráficas de lo obtenido para analizar claramente la información. Se encontró que CAPE se eleva Empalme, Sonora, a las 12:00 horas, en los meses de agosto y septiembre, Y además que PW se encuentre por encima de 40 mm.

7 Bibliografía

- (2018). Convective available potential energy. 4 de Marzo, 2018, de Wikipedia Sitio web: https://en.wikipedia.org/wiki/Convective_available_potential_energy
- (2016). Precipitable water. 4 de Marzo, 2018, de Wikipedia Sitio web: https://en.wikipedia.org/wiki/Precipitable_water
- (2016). Fluid parcel. 4 de Marzo, 2018, de Wikipedia Sitio web: https://en.wikipedia.org/wiki/Fluid_parcel

8 Apéndice

- ¿Cómo se te hizo esta actividad? ¿Compleja, Difícil, Sencilla?
Al inicio no era evidente que acción realizaba exactamente cada paso de la limpieza de datos, pero una vez comprendido y utilizado, se muestra la utilidad que trae.
- ¿Qué te llamó más la atención?
La automatización de acciones manuales por emacs
- ¿Qué parte fue la que menos te interesó hacer?
La actividad fue bastante interesante y útil, ninguna parte específica fue de menos interesante
- ¿Cómo mejorarías esta actividad? ¿Qué le faltó? ¿Qué sobró?
Brindando una explicación breve en pizarrón, para poder localizar y tomar nota sobre como funcionan las acciones realizadas en emacs que permiten automatizar. No faltó ni sobró información/aprendizaje.
- ¿Hasta este punto, que te parece el uso de Jupyter para programar en Python?
Cómodo y rápido, hace buen trabajo en separar acciones por bloques, ejecutarlos y analizar errores en ellos en el momento.