

UNIVERSIDAD DE SONORA

Primera vez en Python

Autor:
Raúl MONTES

Profesor:
Dr. Carlos LIZÁRRAGA

7 de febrero de 2017



Resumen

El trabajo consistió en la preparación y limpieza de datos recolectados de la red de sondeos de norteamérica para después crear tablas y gráficas como histogramas y diagramas de cajas. Para la actividad usamos el documento de texto con los datos anuales que habíamos hecho en la práctica pasada y con ayuda de los comandos de Emacs lo limpiamos para solo dejar los datos referentes a la fecha, CAPE y PW. Después usando jupyter procedimos a crear los gráficos para lo cual se necesito de tablas, limpieza, datos estadísticos y algunos otros recursos de python.

Introducción

Para la práctica fue necesario conocimiento del entorno de Emacs para poder filtrar bien la información, además se necesitó de un breve conocimiento de jupyter así como el manejo de gráficas, es por eso que en esta práctica fueron usados los conocimientos adquiridos en las dos prácticas pasadas.

Preparación de datos

Para preparar los datos se partio del documento de texto creado en la práctica no.2 para posteriormente ser filtrado usando la terminal y los comandos "egrep" dejando solo los archivos de las 00Z y 12Z horas según el caso. Después se usó Emacs para limpiar el archivo y dejar solo los datos de la fecha, CAPE y PW separados por comas en un archivo csv.

Jupyter

Después se abrió una sesión en jupyter para leer los datos obtenidos anteriormente con la ayuda de la biblioteca "pandas". Se generaron dos tablas con los datos de las 00Z y 12Z horas. Lo siguiente fue limpiar los datos de números faltantes, cabe destacar que en este caso no habían numeros faltantes por lo que las tablas quedaron igual. Ya limpios los datos se usó el comando "describe" para obtener los valores estadísticos de los datos, los cuales se usaron después para hacer diagramas de cajas.

Histogramas

Para crear los histogramas de los datos de CAPE y PW no se necesitó más que seleccionar el número de columnas y usar el comando "histo" indicando en cada gráfico si se trataba de los datos CAPE o los de PW.

A continuación se muestran los histogramas de CAPE y PW respectivamente.

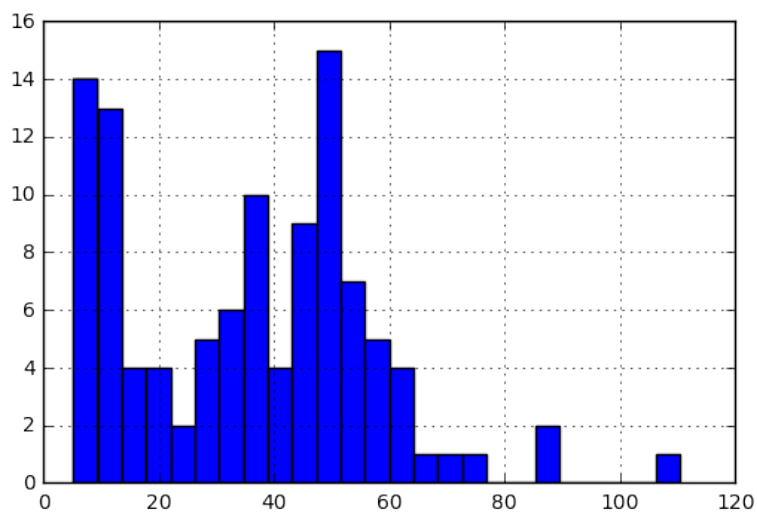
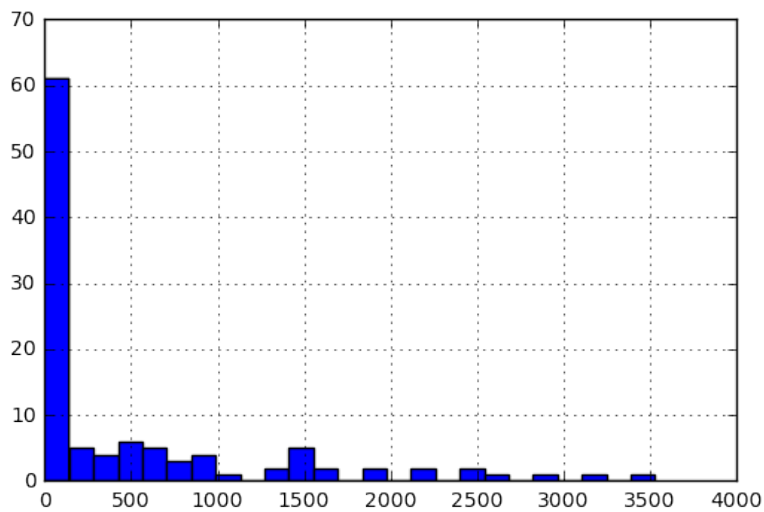
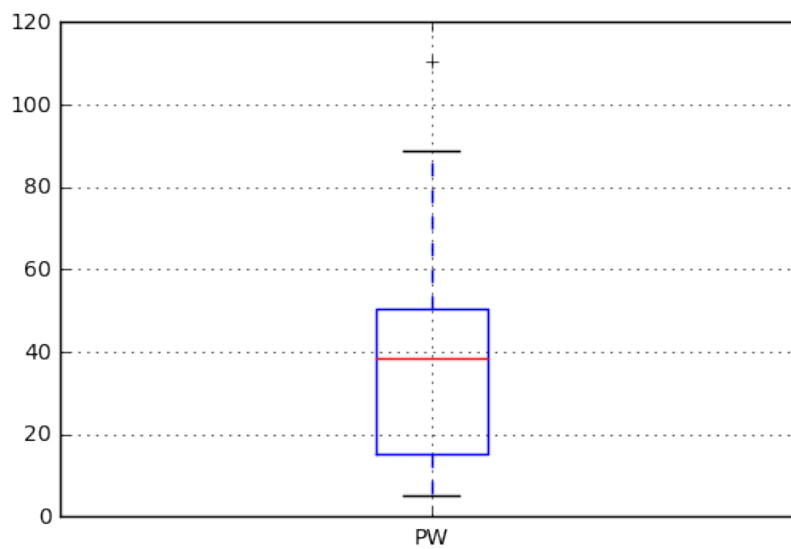
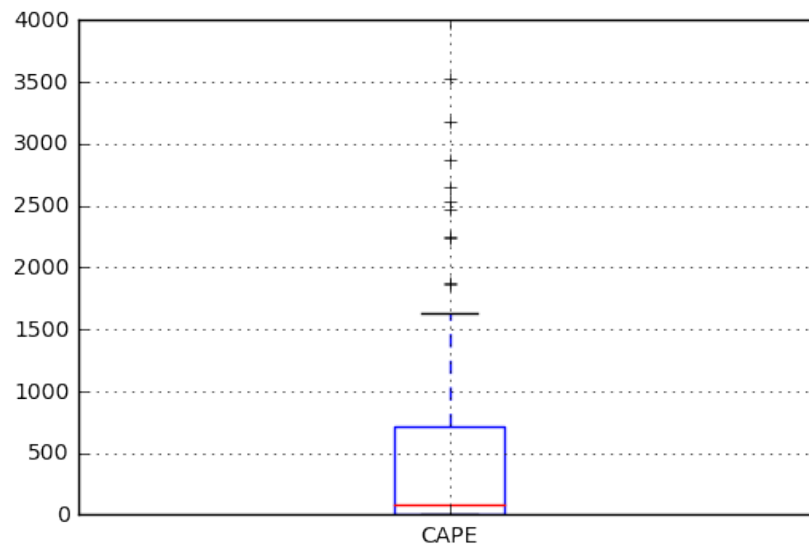


Diagrama de caja

Para los diagramas de cajas se usó el comando "boxplot", se seleccionó el archivo de los datos estadísticos y la columna de CAPE y PW según el caso.



Datos por mes

Por último se construyó una tabla con los datos obtenidos de cada mes en el año 2016 de el archivo original anual. Para hacer la tabla fue necesario el uso de la terminal y los comandos "egrep".

mes	00Z	12Z
Enero	0	27
Febrero	0	3
Marzo	0	0
Abril	0	0
Mayo	0	0
Junio	0	0
Julio	0	11
Agosto	0	29
Septiembre	1	23
Octubre	1	7
Noviembre	0	0
Diciembre	7	8

Table 1: observaciones realizadas a las 00Z y 12Z horas.

Conclusión

Fue una práctica entretenida, aunque al principio no entendí muy bien el objetivo y lo que teníamos que hacer, le terminé entendiendo bastante bien, pienso que contiene información muy útil, además me di cuenta de que Jupyter parece ser útil y no tan complicado, pues comparado con otros medios que he usado, la parte estadística sale bastante fácil usando Jupyter.

Bibliografía

- [1] , WWW.WEATHER.UWYO.EDU *Empalme, Son. Observations*, 2017.
- [2] WWW.MINISCONLATEX.BLOGSPOT.MX *Tablas con latex*, 2017.