

Visualización de datos con la biblioteca Seaborn

ACTIVIDAD 7

Montiel Ramírez B. Karytza

Grupo 1 - Marzo 2019

0.1. Introducción

En esta actividad analizaremos datos de una estación de meteorológica ubicada en un campo de nogal. Utilizando un archivo con datos del año 2009 y parte del 2010. Sin aún saber mucho del significado de las variables presentadas en los archivos, nos interesará saber si hay algún tipo de correlación o dependencia entre ellas. Finalmente se construirá la gráfica o mapa de correlaciones o mejor conocido como "heat map" con ayuda de Matplotlib y con Seaborn y se juzgará cuál gráfica es más fácil de hacer y parece mejor

0.2. Análisis de los archivos

En primera instancia después de agregar las librerías a utilizar, se lee el archivo "meteo-nogal-09.csv" se eliminan las columnas nombradas 'unnamed'.

Creamos una nueva columna de fecha con el motivo de tener más organizados los datos y que se nos facilite al momento de realizar las gráficas. La tabla que se nos presenta contiene un renglón donde se nos muestra las unidades de cada variable a utilizar, al no ser de nuestra relevancia se elimina.

Una vez realizado los pasos anteriores, creamos los mapas de correlación con Seaborn y Matplotlib:

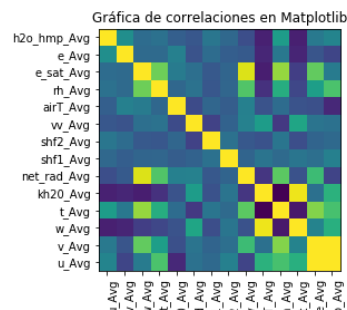


Figura 1: Gráfica de Matplotlib

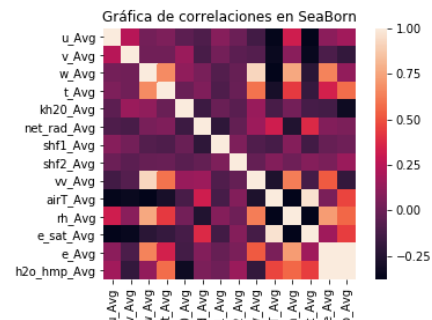


Figura 2: Gráfica de Seaborn

Ambos mapas de correlación muestran lo mismo, pero si nos fijamos en el trasfondo que se llevó a cabo para la realización de los mismos, nos daremos cuenta que es mejor hacerlo con Seaborn debido a la corta longitud del código. Comparando dificultad y extensión en las líneas de código, personalmente me quedo con la gráfica de la Figura 2.

Por otra parte, la actividad señalaba la ejecución de gráficos donde nos marca la correlación de dos variables cuando se encuentran entre -1 y 1. Los valores +1 o -1 sólo se logran cuando hay correlación o anticorrelación total.

A continuación se presentan algunas de ellas:

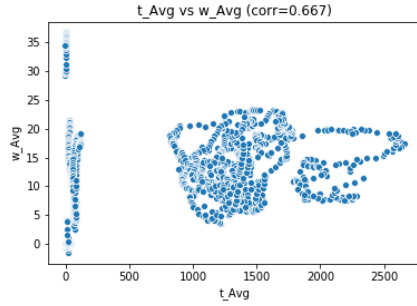


Figura 3: tAvg vs wAvg
(corr=0.667)

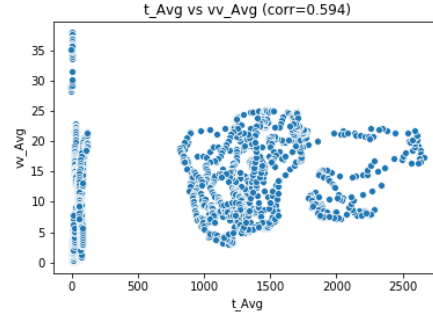


Figura 4: tAvg vs vvAvg
(corr=0.594)

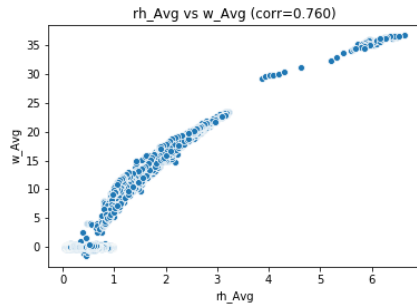


Figura 5: rhAvg vs wAvg
(corr=0.760)

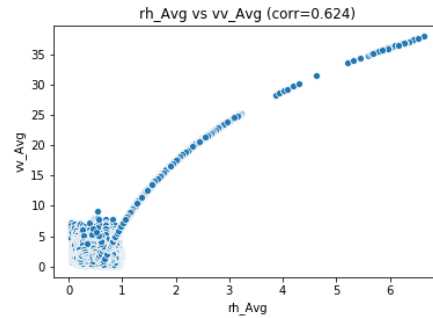


Figura 6: rhAvg vs vvAvg
(corr=0.624)

Como se pudo observar todos los valores de correlación se encuentran en -1 y 1, tal como se planteó en un inicio.

0.3. Conclusión

Finalmente se puede concluir comentando de la facilidad con la que se maneja Seaborn en la creación de mapas de correlaciones. Además, la crea-

ción de gráficos que nos muestran en pares de las variables del archivo de datos que se utilizó.