

trabajo 7

pepe88beto

March 2019

1. Visualización de datos con la biblioteca de seaborn

En esta práctica se nos pide visualizar datos con la biblioteca de seaborn. Usamos los datos recuperados de la estación meteorológica ubicada en un campo de nogal. Con datos del 2009 y 2010.

El archivo contenía muchas columnas sin nombres, lo cual debíamos de quitar, al momento de leer el archivo dije que solo leyera las columnas 1,2,3.. etc... así quitando las columnas vacías.

```
cols=[0,1,4,5,6,7,8,9,10,11,12,13,14,15,16,17]
df=pd.read_csv("datos.csv",engine="python",usecols=cols)
```

Después tiré el renglón donde contenía las unidades de las mediciones, dicho renglón me estorbaba para trabajar. Luego creé una columna fecha la cual me ayudó para seleccionar los datos del 2009, con los que se trabajaría.

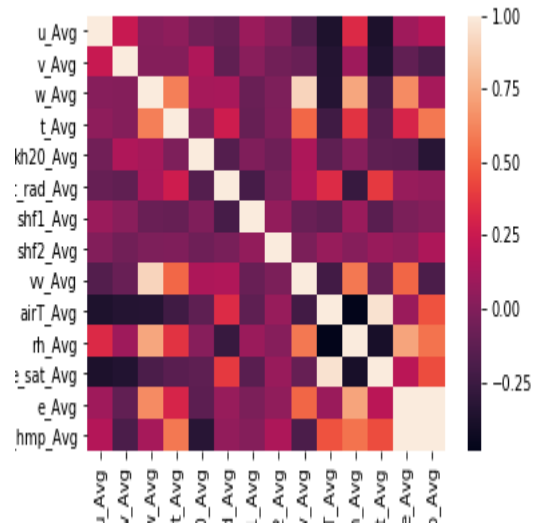
Se nos pide saber si hay algún tipo de correlación o dependencia entre los datos, con la ayuda de la función corr. Dicha función nos regresa una matriz cuadrada con los valores de correlación ($-1 \leq c \leq 1$).

Para encontrar la correlación se tuvo que combiar todas las columnas a numérico python. De la siguiente manera :

```
cols=df1.columns
df1[cols]=df[cols].apply(pd.to_numeric,errors='coerce')
```

Donde en una variable se guardan las columnas de nuestro data frame, luego las pasamos a numérico. Enseguido se usa la función .corr() la cual nos avienta una matriz cuadrada con la correlación.

Para finalizar se grafica esta correlación:



Se le conoce como *heat map*. Es mucho más fácil crear esta gráfica con la biblioteca de seaborn que con matplotlib.