

Uniendo dataframes en Pandas

Eduardo Peñuñuri Bolado

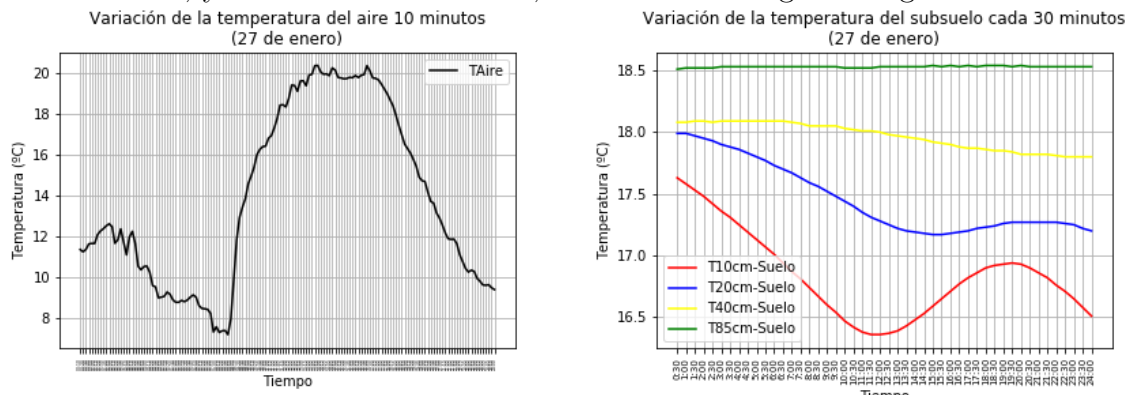
Grupo 1

24 de marzo de 2019

Reporte

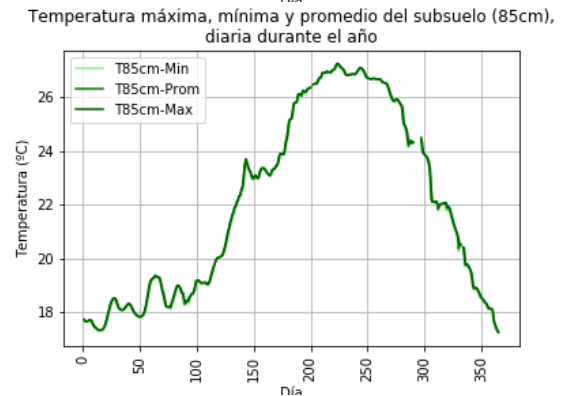
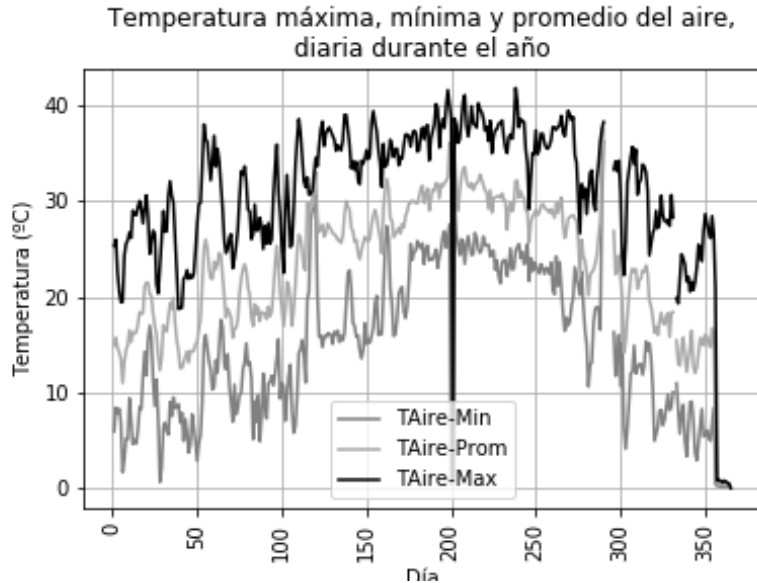
En esta actividad trabajamos los datos meteorológicos de la estación de Nogal, así como con datos del suelo a diferentes profundidades (nosotros usamos las de 10cm, 20cm, 40cm y 85cm), para realizar gráficas de diferentes cosas que se piden utilizando los mismos datos mencionados.

Lo primero que se hizo fue como siempre leer los datos y almacenarlos en Data Frame, para así poder comenzar a trabajar, y separamos los datos que nos interesaban (temperatura del aire y del suelo) para así poder graficar, tomando todo el mes de enero en intervalos de, el aire 30 minutos, y el suelo de 10 minutos, obteniendo las siguientes gráficas:



Aquí podemos ver como la del aire va cambiando según la hora del día, como uno esperaríamos, mientras que las temperaturas se diferencian entre sí, comportándose igual pero entre más profundo, menos variación hay.

Después de eso, se nos pidió realizar gráficas que tomaran todo el año de las temperaturas máxima, mínima y promedio, en intervalos de un día, por lo que primero delimitamos lo necesario (para no tener datos que no fueran de 2009), y luego sacamos promedios, mínimos y máximos, para posteriormente realizar las gráficas, obteniendo lo siguiente:

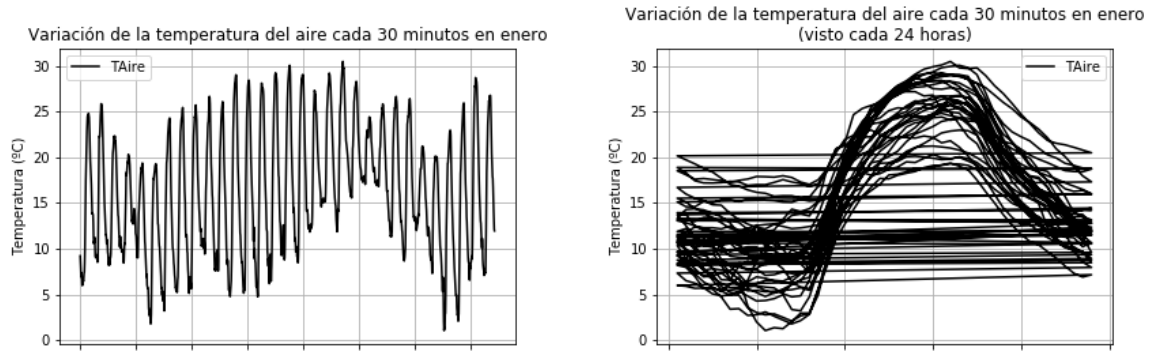


En éste caso, es curioso ver que las líneas se parecen bastante, solo que parecieran estar unas más arriba o abajo, pero en las del suelo, entre más profundo más difícil es ver la diferencia entre máxima, mínima y promedio, al punto de que a 85cm se hace prácticamente imposible ver más de una línea.

Luego, volvimos al mes de enero, para volver a graficar, en intervalos de 30 minutos, la variación en 24 horas de las mismas variables que hemos estado trabajando, para lo cual fue necesario sacar promedios de las del suelo, ya que se encontraban cada 10 minutos, así que se agarraron de tres en tres y se promediaban.

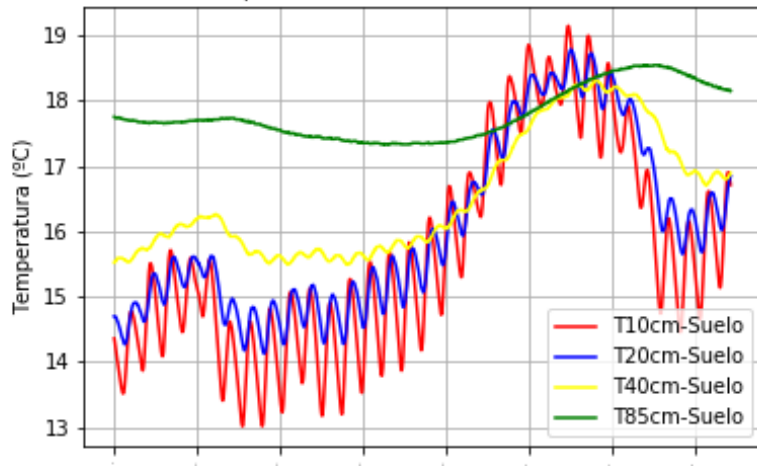
Se hicieron dos tipos de gráfica para cada cosa que se pedía, en la que se muestra la temperatura según va pasando el tiempo, con una sola línea para cada variable a medir temperatura, y en la que se ve un mismo día y cada línea representa un día diferente, y de esta forma se puede ver como en el mes de enero se diferencian los días (a esta forma le llamé “Formato 24”). Los resultados obtenidos se muestran a continuación:

La temperatura del aire

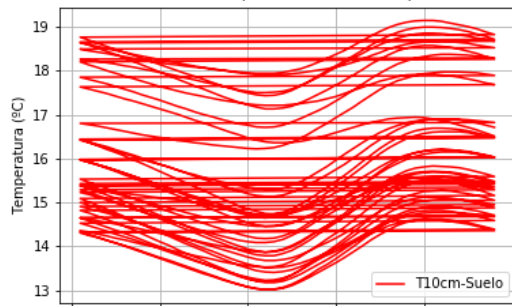


Las temperaturas del suelo

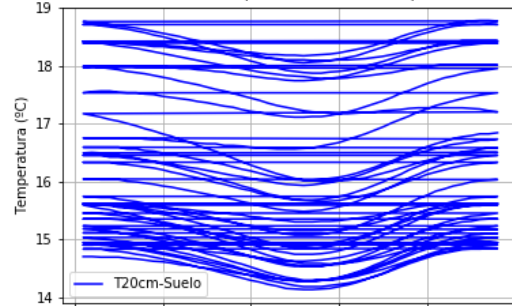
Variación de la temperatura del subsuelo cada 30 minutos en enero



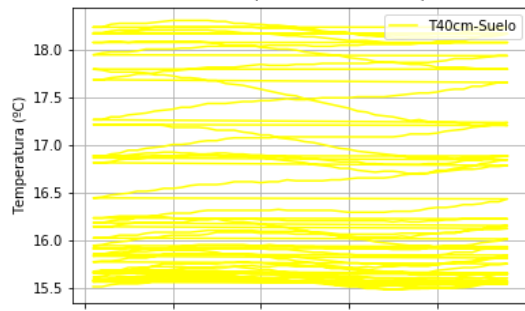
Variación de la temperatura del subsuelo (10cm) cada 30 minutos en enero (visto cada 24 horas)



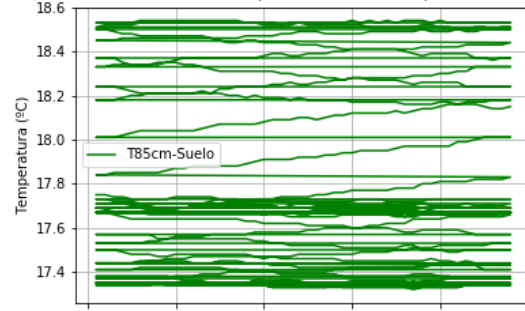
Variación de la temperatura del subsuelo (20cm) cada 30 minutos en enero (visto cada 24 horas)



Variación de la temperatura del subsuelo (40cm) cada 30 minutos en enero (visto cada 24 horas)



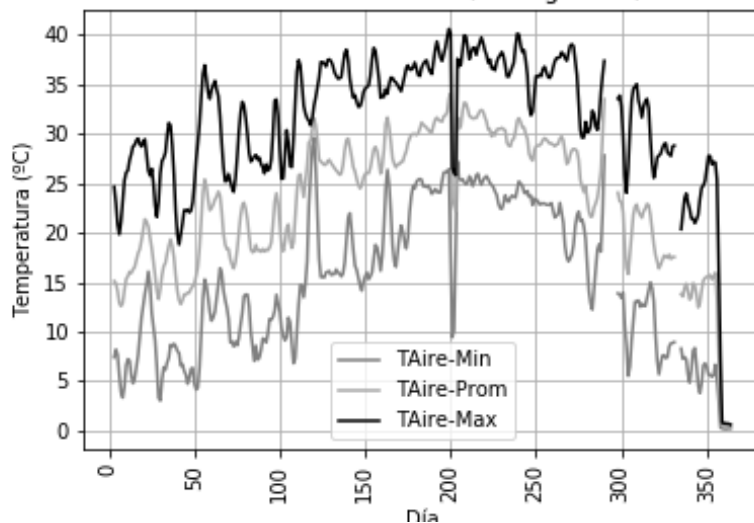
Variación de la temperatura del subsuelo (85cm) cada 30 minutos en enero (visto cada 24 horas)



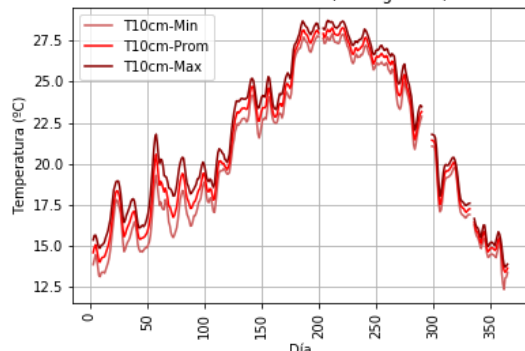
Igual que como se dijo anteriormente, lo que más se puede apreciar es que la temperatura de los días suelen variar bastante durante el mes, cambiando de día en día, pero en el suelo se nota menos, llegando al punto de que en lo más profundo varía por menos de un grado.

Finalmente, se hizo uso del “Rolling mean”, lo cual se aplicó a las gráficas, y a parte de ser más estéticas de esta forma, se puede llegar a tener más claridad al estudiarlas. Se aplicó a las gráficas de temperaturas máxima, mínima y promedios vistas anteriormente, obteniendo estas nuevas:

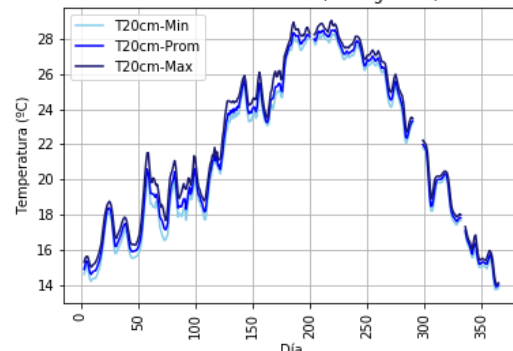
Temperatura máxima, mínima y promedio del aire, diaria durante el año (Rolling Mean)

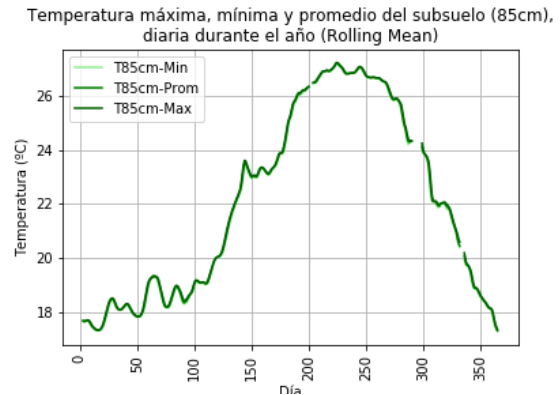
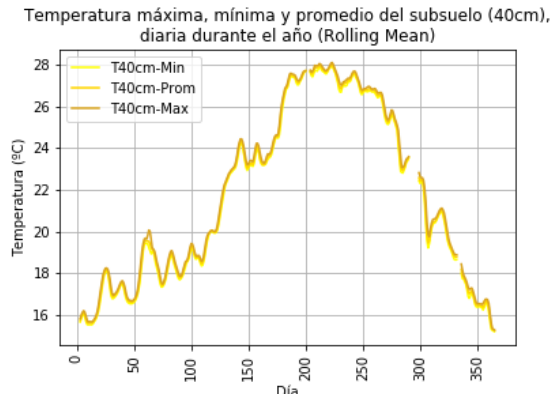


Temperatura máxima, mínima y promedio del subsuelo (10cm), diaria durante el año (Rolling Mean)



Temperatura máxima, mínima y promedio del subsuelo (20cm), diaria durante el año (Rolling Mean)





Aún suavizadas podemos ver lo mismo, por lo que en este caso sirvió más bien de estética, ya que se analiza igual, e incluso puede perjudicar un poco, ya que con las otras había menos redondeo y por lo tanto eran más precisos los datos.

Conclusión

En conclusión, la actividad se trató principalmente de graficar, por lo que no tuvo demasiada dificultad en eso, sin embargo, se ocupó trabajar con bastantes datos de diversas formas, lo cual si complicó algunas cosas, además de que al pedir tanta variedad se tuvieron que hacer bastantes Data Frame, lo cual se volvió extremadamente tedioso, pero fuera de eso no estuvo mal; en cuanto al rolling mean, no estuvo difícil y creo que fue la parte más sencilla y rápida.

Referencias

- [1] Meenakshi, A. (2019). *Python Strings, Functions and Examples*. Recuperado en marzo de 2019 de: <https://www.techbeamers.com/python-strings-functions-and-examples/>
- [2] Pandas Pydata (2019). *Merge, join, and concatenate*. Recuperado en marzo de 2019 de: https://pandas.pydata.org/pandas-docs/stable/user_guide/merging.html