

Actividad 3

Hannah Borrego

29 January 2021

1 Introduction

En esta actividad número tres se utiliza el archivo de la actividad 1 de datos meteorológicos, en mi caso del poblado de Santa Ana. Se hace uso la biblioteca pandas que es una librería de Python especializada en el manejo y análisis de datos. Permite acceder a ellos de principio a fin y da la posibilidad de hacer un análisis a profundidad sobre dicho archivo. En este trabajo eso fue lo que se hizo, profundizar con la biblioteca pandas y con todos los comandos que permiten cambiar columnas, renombrar datos, contar renglones, ver los primeros y últimos renglones, fecha, cambiar el tipo de números a flotante, etc. Se va a explorar con ellos mediante el archivo meteorológico, experimentando así, cada una de sus funciones.

1.1 bibliotecas

Esta semana se nos pidió revisar algunas de las preguntas más comunes sobre la biblioteca pandas en la página de "Stack overflow" y recopilé algunas de las que más se me hicieron interesantes:

Pregunta 1:

¿Cómo puedo pivotar la tabla/DataFrame para que salga justo al revés respetando los valores?

Respuesta: Prueba utilizando `transpose(*args, copy=False)`.

link: <https://es.stackoverflow.com/questions/424393/pivoteo-de-un-dataframe-completo-python-pandas>

Pregunta 2:

Conseguir de un dataframe filas con valores contenidos en otro dataframe

Respuesta:

Si en ambos dataframes `id` son columnas (no índice) debes usar `pandas.Series.isin` para poder seleccionar las filas de `util` que cuyos valores en `id` están en `fin.id`:

link: <https://es.stackoverflow.com/questions/365514/conseguir-de-un-dataframe-filas-con-valores-contenidos-en-otro-dataframe>

Pregunta 3:

¿Cómo tomar fecha de fecha de hoy php?

Respuesta:

Lo que puedes hacer es tomar directamente la fecha con `date()` de php y guardarlo en una variable, hay varios formatos pero un ejemplo rápido puede ser algo así:

Usos Horarios para php:

```
defino antes el uso horario  
date_default_timezone_set('UTC');
```

link= <https://es.stackoverflow.com/questions/241022/c>

Hablando de funciones que ofrece pandas:

Leí varias que me parecen las más importantes como "read csv" para leer el archivo, "df info" para ver que contiene tu archivo, "df to numeric" para convertir a flotantes y "df head" y "df tail" para imprimir el encabezado así como el final del data frame. Por supuesto import numpy as np, import pandas as pd al inicio del código a ejecutar, para poder importar correctamente dichas bibliotecas.

Esas son las funciones que considero más importantes y que más se me quedaren grabadas. Sin olvidar el "raw" de donde copié el archivo .txt de github. Lee hasta donde le indiques del archivo, y hace más rápido un análisis de archivos con grandes cantidades de datos.

1.2 conclusión

En esta actividad me pareció entretenido el hecho de poder optimizar el análisis de datos, aún mas datos con los que ya habíamos trabajado, que conocemos más y sabemos que columnas de información contiene. La carga de trabajo estuvo fácil solo que en mi caso me atoró un par de días para trabajar el hecho de que me daba error al poner mi link de github, hasta que vi un tutorial donde me sugerían usar la función de "raw", volví a ver la clase, vi que tambien lo sugería el profesor y decidí intentarlo y funcionó, apartir de ahí el trabajo consistió en leer para que servía cada función y llevarlas a cabo con los datos de mi archivo. Se me dificultó cuando algo no corría bien encontrar el error y corregirlo. Le asignaría valor de ficitultad bajo, actividad muy amena y sobre todo de mucho aprendizaje.