

Actividad 3

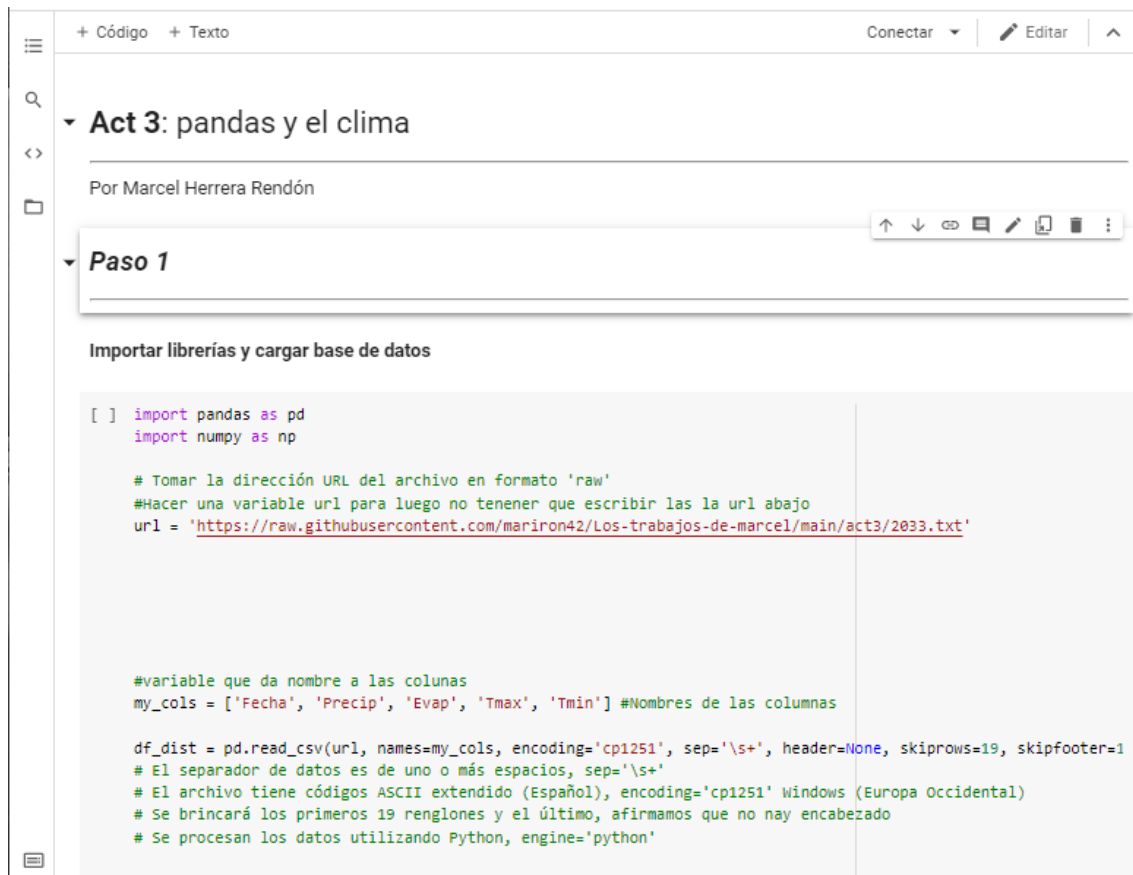
Física Computacional

Marcel Herrera Rendón 219221966

29 de Enero de 2021

1. Introducción

Esta actividad es el primer acercamiento con la librería pandas de Python que sirve para analizar bases de datos, es usada en trabajos ciencia de datos. Mi primera experiencia con este trabajo fue TERROR, todo se veía tan complicado y novedoso hasta que me di cuenta de que realmente no lo era, basta con imaginar que estás usando Excel con comandos o unas listas y diccionarios con esteroides, la verdad la única dificultad está en encontrar la función correcta, además de que perdí varias tardes apenas en echarlo a andar, pero luego fijándome en el GitHub del profe y vi cómo se le hacía. Siendo sinceros, siento que no inventé nada por mi propia cuenta, pero me esmeré en comentar todo el Código y moverle a ver cómo es que funciona, espero que si alguien que es nuevo ve mi trabajo pueda comprender exactamente qué hace cada cosa con todas las notas que le integré.



The screenshot shows a Jupyter Notebook interface. The top bar includes a menu with '+ Código' and '+ Texto', a 'Conectar' dropdown, and an 'Editar' button. The notebook content is organized into sections: 'Act 3: pandas y el clima' (author: Por Marcel Herrera Rendón) and 'Paso 1'. The code cell is titled 'Importar librerías y cargar base de datos' and contains the following Python code:

```
[ ] import pandas as pd
import numpy as np

# Tomar la dirección URL del archivo en formato 'raw'
#Hacer una variable url para luego no tener que escribir la url abajo
url = 'https://raw.githubusercontent.com/mariron42/Los-trabajos-de-marcel/main/act3/2033.txt'

#variable que da nombre a las columnas
my_cols = ['Fecha', 'Precip', 'Evap', 'Tmax', 'Tmin'] #Nombres de las columnas

df_dist = pd.read_csv(url, names=my_cols, encoding='cp1251', sep='\s+', header=None, skiprows=19, skipfooter=1
# El separador de datos es de uno o más espacios, sep='\s+'
# El archivo tiene códigos ASCII extendido (Español), encoding='cp1251' Windows (Europa Occidental)
# Se brincarán los primeros 19 renglones y el último, afirmamos que no hay encabezado
# Se procesan los datos utilizando Python, engine='python'
```

2. Bibliotecas

En el ensayo anterior hablamos larga y tendidamente sobre varias librerías nuevas a usar, por otro lado esta vez solo integramos una novedoso añadido. La librería Pandas es una potente Herramienta para el manejo de las bases de datos que te permite trabajar con archivos de varios formatos desde CSV, Excel y SQL y otros tipos de bases de datos. Tiene varios tipos de arreglos y un sin fin de funciones que sirven para obtener información pertinente rápida y eficazmente. Puedes manejar series, data frame y panel. La primera es solo de una dimensión mientras que la segunda es como las tablas de Excel y la tercera no la he manejado, pero en internet dice que se asemeja a manejar cubos de tres dimensiones. Se pueden buscar y remplazar cosas automáticamente con una función determinada, además permite majar las variables de tiempo con un formato que Python entiende perfectamente lo que nos vuelve más fácil todo a la hora de buscar cosas en función a sus datos cronológicos. Hay muchas más funciones, pero algo interesante es que las puedes combinar con el resto de opciones en Python por ejemplo pueden interactuar con un ciclo.



3. Experiencia y retroalimentación

Cómo ya había dicho antes en este ensayo esta semana estuve atemorizado, todo se vía demasiado complicado y seguirle el ritmo a las explicaciones me resultaba imposible, sin embargo realmente era bastante sencillo, creo que lo que me ayudó fue ver varios ejemplos de cómo funcionaban las cosas y darme cuenta que lo realmente importante era concentrarme en aprender mientras hacía el trabajo y no antes porque si no nunca te sentirás listo. Yo le asignaría un grado de complejidad intermedio.

4. Links

Mi trabajo

<https://colab.research.google.com/drive/1JTgIDmye2RWKppHZdqPvPd1IoTfFzuWn?usp=sharing>

Fuentes

<https://github.com/carloslizarragac/FisicaComputacional1/blob/master/Actividad3.ipynb>

<https://raw.githubusercontent.com/mariron42/Los-trabajos-de-marcel/main/act3/2033.txt>

<https://pythonbasics.org/read-csv-with-pandas/>

<https://www.listendata.com/2019/06/pandas-read-csv.html>