

UNIVERSIDAD OBERTA DE CATALUNYA
ESTUDIOS DE INFORMÁTICA, MULTIMEDIA Y TELECOMUNICACIONES.
TIPOLOGIA Y CICLO DE VIDA DE LOS DATOS.
PRACTICA1- . WEB SCRAPING.

APELLIDOS: ZAMBRANO ZAMBRANO
NOMBRES: LÍDER EUCLIDES

RESPUESTAS A LAS PREGUNTAS DE LA PRACTICA 1 – WEB SCRAPING

1. CONTEXTO. EXPLICAR EN QUÉ CONTEXTO SE HA RECOLECTADO LA INFORMACIÓN. EXPLIQUE POR QUÉ EL SITIO WEB ELEGIDO PROPORCIONA DICHA INFORMACIÓN.

Considerando que se desea realizar un estudio sobre el clima en las ciudades más importantes de Ecuador, y conociendo que Ecuador es un país donde existen 9 zonas climáticas diferentes, y relacionando el clima con el suelo y los cultivos se ha escogido el sitio **meteored.com.ec** que nos permitirá en un primer nivel conocer los valores de temperatura y poder predecir el clima en los próximos meses de las ciudades más importantes.

2. DEFINIR UN TÍTULO PARA EL DATASET. ELEGIR UN TÍTULO QUE SEA DESCRIPTIVO.

TÍTULO DEL DATASET: conjunto de datos climáticos

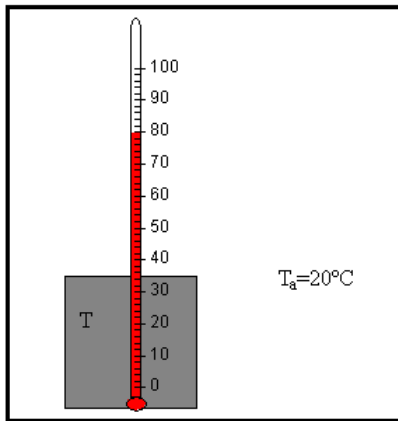
3. DESCRIPCIÓN DEL DATASET. DESARROLLAR UNA DESCRIPCIÓN BREVE DEL CONJUNTO DE DATOS QUE SE HA EXTRAÍDO (ES NECESARIO QUE ESTA DESCRIPCIÓN TENGA SENTIDO CON EL TÍTULO ELEGIDO).

El dataset tipo dataframe que hemos utilizado nos permite organizar los datos en una estructura rectangular en forma de tabla o matriz, en las matrices de un dataframe se aceptan valores alfanuméricos, por lo que esta estructura puede contener diferentes tipos de datos.

Nuestro dataset nos permite obtener tres campos, el nombre de la ciudad, la temperatura máxima y la temperatura mínima, el primer campo es de tipo cadena y los otros dos campos son de tipo numérico, este conjunto de datos nos ofrece información relacionada al clima en las ciudades más importantes del ecuador.

Para esta práctica se ha utilizado un dataframe porque nos permite además incluir el nombre de las columnas.

4. REPRESENTACIÓN GRÁFICA. PRESENTAR UNA IMAGEN O ESQUEMA QUE IDENTIFIQUE EL DATASET VISUALMENTE.



5. CONTENIDO. EXPLICAR LOS CAMPOS QUE INCLUYE EL DATASET, EL PERIODO DE TIEMPO DE LOS DATOS Y CÓMO SE HA RECOGIDO.

El campo Ciudad: Nos muestra el nombre de las ciudades más importantes del ecuador de las cuales tendremos los valores de temperatura máxima y mínima.

El campo Temperatura mínima: Este campo nos muestra los valores mínimos de temperatura y nos sirve para estudios del clima en zonas estratégicas.

El campo Temperatura máxima: Nos muestra el valor máximo de la temperatura y nos sirve para estudiar el comportamiento de los valores máximos de temperaturas máximas, así hacer evidente las zonas con poca vegetación y por lo tanto con menos humedad disponible.

El periodo de tiempo en recoger los datos demora aproximadamente unos 10 segundos.

Los datos se han recogido aplicando un dataset tipo dataframe, para ello hemos utilizado las bibliotecas como BeautifulSoup, pandas y requests.

La función findall que filtra todos los elementos de la clase a recoger los datos, y la función find que filtra cada elemento de una clase.

Luego hemos aplicado un ciclo for para recoger no solo un elemento sino todos los elementos y los presentamos los datos recogidos en un dataframe, para luego exportarlos en un archivo csv.

6. AGRADECIMIENTOS. PRESENTAR AL PROPIETARIO DEL CONJUNTO DE DATOS. ES NECESARIO INCLUIR CITAS DE INVESTIGACIÓN O ANÁLISIS ANTERIORES (SI LOS HAY).

Se agradece a **Francisco Javier López M.**, Cofundador de **Meteored.com** el portal del aficionado a la meteorología más importante de España y referente en el resto de países de habla española.

CITAS BIBLIOGRAFICAS

[Consulta: 10 de noviembre de 2019]

<<https://www.meteored.mx/noticias/divulgacion/modelos-de-prediccion-numerica-y-su-uso-en-la-meteorologia.html>>

[Consulta: 10 de noviembre de 2019] <http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers11-10/21848.pdf>

7. INSPIRACIÓN. EXPLIQUE POR QUÉ ES INTERESANTE ESTE CONJUNTO DE DATOS Y QUÉ PREGUNTAS SE PRETENDEN RESPONDER.

Sin duda alguna el clima tiene una acción decisiva sobre el mundo mineral, vegetal y animal: meteorización de las rocas, génesis y erosión de los suelos, distribución geográfica de la flora, fauna y asentamientos humanos, crecimiento de plantas y cultivos, etc.

El ECUADOR por ser un país agrícola y poseer una gama de climas ampliamente diversificados, con grandes variaciones de una a otro, muestra su interés y preocupación por las condiciones climáticas.

¿En qué ciudades las temperaturas favorecen a los agricultores en sus cultivos?

¿Cuáles son las zonas potenciales para los cultivos según el estudio del clima?

8. LICENCIA. SELECCIONE UNA DE ESTAS LICENCIAS PARA SU DATASET Y EXPLIQUE EL MOTIVO DE SU SELECCIÓN:

- RELEASED UNDER CC0: PUBLIC DOMAIN LICENSE
- RELEASED UNDER CC BY-NC-SA 4.0 LICENSE
- RELEASED UNDER CC BY-SA 4.0 LICENSE
- DATABASE RELEASED UNDER OPEN DATABASE LICENSE, INDIVIDUAL CONTENTS UNDER DATABASE CONTENTS LICENSE
- OTHER (SPECIFIED ABOVE)
- UNKNOWN LICENSE

-LICENCIA PARA NUESTRO DATASET
RELEASED UNDER CC BY-SA 4.0 LICENSE

SE HA SELECCIONADO POR LOS SIGUIENTES CARACTERISTICAS:

Atribución — Se debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios. Se puede hacer en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo del licenciante.

No Comercial — No puede hacer uso del material con propósitos comerciales.

Compartir Igual — Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original.

9. CÓDIGO. ADJUNTAR EL CÓDIGO CON EL QUE SE HA GENERADO EL DATASET, PREFERIBLEMENTE EN PYTHON O, ALTERNATIVAMENTE, EN R.

```
import pandas as pd
import requests
from bs4 import BeautifulSoup

page = requests.get("https://www.meteored.com.ec/")
soup = BeautifulSoup(page.content, "html.parser")
```

```
elementos = soup.find_all(class_="li-top-prediccion")#esta clase siempre será la que se repita
```

```
ciudad=[elemento.find(class_="anchors").get_text() for elemento in elementos]  
t_maxima =[elemento.find(class_="cMax changeUnitT").get_text() for elemento in elementos]  
t_minima =[elemento.find(class_="cMin changeUnitT").get_text() for elemento in elementos]
```

```
lider = pd.DataFrame(  
    {  
        "Ciudad":ciudad,  
        "Temperatura maxima":t_maxima,  
        "Temperatura minima":t_minima  
    })  
print(lider)  
lider.to_csv("datos_metereologicos.csv")
```

10. DATASET. PRESENTAR EL DATASET EN FORMATO CSV

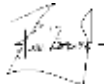
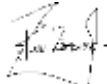
```
lider.to_csv("datos_metereologicos.csv")
```

INTEGRANTES DEL GRUPO:

No se trabajó en grupo, el trabajo realizado es personal.

NOMBRE DEL INTEGRANTE:

ZAMBRANO ZAMBRANO LIDER EUCLIDES

Contribuciones	Firma
Investigación previa	LEZZ 
Redacción de las respuestas	LEZZ 
Desarrollo código	LEZZ 