



**Universidad Nacional
Autónoma de México**
Facultad de Ingeniería
División de Ingeniería Mecánica e
Industrial



**Laboratorio de Cómputo de Ingeniería Mecatrónica
(1472)**

*Profesor: Miguel Serrano Reyes
Semestre 2021-1*

Práctica No. 2

Nombre de la práctica

Implementación de una Interfaz de Programación de
Aplicaciones(API)

Grupo: 2

Nombre del Estudiante:
Celaya González David Alejandro

Actividad 1:

#Conteo de palabras

```
import requests #Importo librerias
from bs4 import BeautifulSoup
```

Obtenemos las librerías correspondientes para poder trabajar

```
#Contador de palabras
pagina = requests.get("https://www.york.ac.uk/teaching/cws/wws/webpage1.html")#Obtengo codigo html
soup = BeautifulSoup(pagina.content,"html.parser")#Objeto de la clase BeautifulSoup
```

En esta sección obtendremos el código HTML de la pagina de nuestro interés. Seguido de esto se crea un objeto BeautifulSoup, el cual recibirá como argumento la pagina y el parser. En este caso "html.parser" aunque existen otros tipos como lxml y htl5lib. Que basicamente son analizadores de sintaxis.

```
texto_pagina=soup.find_all("div")#Buscamos todos los div
#type(texto_pagina)
```

Una vez que tenemos el código HTML, buscaremos todas las secciones 'div' que es donde se encuentra la información de nuestro interés.

```
contador = 0
for texto in texto_pagina:
    #print(texto.get_text())
    palabras = texto.get_text()#Obtengo texto completo
    #print(palabras)
    palabras_totales = palabras.split() #Cuento con split()
    #print(palabras_totales)
    #print(len(palabras_totales))
    contador = contador +len(palabras_totales)#Acumulo el total
```

En esta parte comenzaremos a iterar por nuestro texto obtenido en las secciones 'div' que encontramos anteriormente. Guardaremos el texto en una variable llamada palabras que almacenara temporalmente las palabras y debido a que es de tipo 'str' podremos contarlas con el metodo split().Despues lo almacenaremos en una variable que será de tipo 'list' para poder contar su longitud con la función len().

Al final acumularemos todas la longitudes y guardaremos en la variable contador con el fin de obtener un total de palabras.

```
[7] print("Las palabras en la pagina web son:",contador)
```

```
Las palabras en la pagina web son: 643
```

Finalizando la actividad, imprimiremos el contador que tendrá el total de palabras en la pagina web.

Actividad 2:

#Obtención de nombres de una tabla

```
import requests #Importo librerías
from bs4 import BeautifulSoup
```

Obtenemos las librerías correspondientes para poder trabajar

```
pagina = requests.get("https://en.wikipedia.org/wiki/List_of_current_National_Football_League_stadiums") #Obtengo código HTML
print(pagina)
```

Esta parte nos ayuda a obtener el código HTML de la página de nuestro interés.

```
[85] soup = BeautifulSoup(pagina.content, "html.parser") #Objeto BeautifulSoup
print(soup.prettify()) #Ordeno código HTML
```

En esta sección obtendremos el código HTML de la página de nuestro interés. Seguido de esto se creará un objeto BeautifulSoup, el cual recibirá como argumento la página y el parser. En este caso "html.parser" aunque existen otros tipos como lxml y html5lib. Que básicamente son analizadores de sintaxis.

Prettify() ordenará el código HTML de manera que lo podamos observar de una mejor manera.

```
table = soup.find_all("table")[1] #Accedo a la tabla de Estadios
for fila in table.find_all('tr'): #Accedo a cada fila
    for columna in fila.find_all('th'): #Accedo a los cuadros de nombre
        for titulo in columna.find_all('a'): #Obtengo los títulos
            name = titulo.text #Obtengo el texto de títulos
            print(name)
```

```
Allegiant Stadium
Arrowhead Stadium
AT&T Stadium
Bank of America Stadium
Bills Stadium
Empower Field at Mile High
FedExField
FirstEnergy Stadium
Ford Field
Gillette Stadium
Hard Rock Stadium
Heinz Field
Lambeau Field
Levi's Stadium
Lincoln Financial Field
Lucas Oil Stadium
Lumen Field
M&T Bank Stadium
Mercedes-Benz Stadium
Mercedes-Benz Superdome
MetLife Stadium
Nissan Stadium
NRG Stadium
Paul Brown Stadium
Raymond James Stadium
SoFi Stadium
Soldier Field
State Farm Stadium
TIAA Bank Field
U.S. Bank Stadium
```

En esta última sección se busca la tabla de interés con `find_all('table')[1]` y se guarda en una variable que nos ayudará a iterar por nuestra tabla. Primero accederemos a la cada fila de la tabla buscando todas las 'tr', después todas las 'th'. Estas últimas serán la sección de código donde se guarda la información de nuestro interés, pero dentro de la misma buscaremos todas las secciones 'a' para poder llegar al nombre del estadio o en su defecto la información que buscamos. Ya que llegamos a este sitio, guardaremos el texto de la sección 'a' que será el nombre del estadio, para finalmente obtener como salida todos los nombres, de todos los estadios, contenidos en una tabla de una página web.