



Universidad Nacional Autónoma de
México

Facultad de Ingeniería

División de Ingeniería Mecánica e Industrial

Laboratorio de Cómputo de Ingeniería Mecatrónica (1964)



Profesor: Miguel Serrano Reyes

Semestre 2022-1

Práctica No. 2

Implementación de una Interfaz de Programación de Aplicaciones (API)

Nombre del Estudiante:

Jasso Garduño Juan José

1. Obtener el código html del sitio web "<https://www.arthurleej.com/e-history.html>"; y acceder al título del documento

```
import requests
from bs4 import BeautifulSoup

print('Actividad 1')
print('\nObtener el código html del sitio web "https://www.arthurleej.com/e-history.html" y acceder al título del documento\n')

#Solicitamos entrar a la pagina y descargar la información
pagina=requests.get("https://www.arthurleej.com/e-history.html")

#Me da el código html de este sitio web
# pagina.content

#Para organizar el código html, creamos un objeto
soup=BeautifulSoup(pagina.content,'html.parser')

#**** Accedemos al título del documento y lo imprimimos*****
#Accedemos a algunas regiones usando soup
#Encuentrame la primera coincidencia que tenga la letra b
#Me regresa toda la etiqueta
titulo=soup.find('b')

#Si solo quiero el texto dentro de la etiqueta
print(titulo.get_text())
```

```
Python 3.8.5 (default, Sep 3 2020, 21:29:08) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.

IPython 7.22.0 -- An enhanced Interactive Python.

In [1]: runfile('C:/Users/User/Google Drive/UNAM/9no Semestre/TSP II - Ciencia de datos/Tareas
Semestre/TSP II - Ciencia de datos/Tareas sketch/Prácticas/P2')
Actividad 1

Obtener el código html del sitio web "https://www.arthurleej.com/e-history.html" y accede
History
```

Actividad 2. Obtener el código html del sitio web "<https://github.com/MiguelSerranoReyes>"; y acceder al nombre del usuario

```
print("-----")
print('Actividad 2')
print('\nObtener el código html del sitio web "https://github.com/MiguelSerranoReyes" y acceder al nombre del usuario')
#Solicitamos entrar a la pagina y descargar la información
pagina=requests.get("https://www.github.com/MiguelSerranoReyes")
#Para organizar el código html, creamos un objeto y lo imprimimos
soup=BeautifulSoup(pagina.content,'html.parser')

#Sabemos que el nombre está en una etiqueta llamada span
#Quiero un span en particular, con ciertos atributos
nombre_usuario=soup.find('span',attrs={'class':'vcard-username'})
print(nombre_usuario.get_text())
```

```
-----
Actividad 2

Obtener el código html del sitio web "https://github.com/MiguelSerranoReyes" y acceder al nombre del usuario

MiguelSerranoReyes
```

Actividad 3. Navegar por el DOM del sitio web "<https://www.arthurleej.com/e-history.html>" e imprimir nuevamente el título.

```
print("-----")
print('Actividad 3')
print('\nNavegar por el DOM del sitio web "https://www.arthurleej.com/e-history.html" e imprimir nuevamente el título.\n')

pagina=requests.get("https://www.arthurleej.com/e-history.html")
soup=BeautifulSoup(pagina.content,'html.parser')

#Me da la idea que es una lista y puedo acceder a sus elementos
#soup.children

#Ahora quiero ver el tipo de archivos que contiene esa lista
#Me dice que en el tercer elemento son clase bs4.element.Tag
'''
for child in soup.children:
    print(type(child))
'''

html=list(soup.children)[2] #También es una lista
#Podemos ver como estamos navegando para ver cómo está estructurado el código html
'''
for section in html:
    print('\n\nNueva Sección')
    print(section)
'''

#De la salida del ciclo podemos ver que la cabecera <head> está en la sección 1
head=list(html.children)[1]
#Veamos las etiquetas que hay en la cabecera
'''
for tag in head:
    print('\n\nNueva Etiqueta')
    print(tag)
'''

#Si quisiera sacar el título, sé que está en el elemento 1
titulo=list(head.children)[1]
print(titulo.get_text())
```

```
-----
Actividad 3

Navegar por el DOM del sitio web "https://www.arthurleej.com/e-history.html" e imprimir nuevamente el título.

Essay on History by Arthur Lee Jacobson
```