For the example I will inspect the website of the **Madrid Stock** Exchange

Usually I can inspect by right clicking on the area that interests me.

In our case I am interested in obtaining the value of the row named **IBEX35®** and the value of the **"last" or “ultimo”** column.

Python code - Let's start!

Let's see in code how I will do to access that piece of text.

First I import the Python libraries that I will use:

|  |  |
| --- | --- |
| ***1***  ***2***  ***3***  ***4*** | ***import requests***  ***from bs4 import BeautifulSoup***  ***import csv***  ***from datetime import datetime*** |

I indicate the route of the web that I want to access:

|  |  |
| --- | --- |
| ***1***  ***2*** | ***# indicate the route***  ***url\_page = 'http://www.bolsamadrid.es/esp/aspx/Indices/Resumen.aspx'*** |

And now I will make the request to that route and process the HTML using an object of type BeautifulSoap:

|  |  |
| --- | --- |
| ***1***  ***2***  ***3*** | ***#***  ***page = requests.get(url\_page).text***  ***soup = BeautifulSoup(page, "lxml")*** |

Well, now it's time to think about the strategy to access value. In my case I am first interested in accessing the table, and from there to its cells. Luckily the table has a unique id

|  |  |
| --- | --- |
| ***1***  ***2***  ***3*** | ***# Obtenemos la tabla por un ID específico***  ***tabla = soup.find('table', attrs={'id': 'ctl00\_Contenido\_tblÍndices'})***  ***tabla*** |

Well, now within the table and being that in this case I do not have a direct access to the cells by unique ids or by classes, I only have to iterate ... Then, I will access the first row and get the name of the index cells and its value:

|  |  |
| --- | --- |
| ***1***  ***2***  ***3***  ***4***  ***5***  ***6***  ***7***  ***8***  ***9***  ***10***  ***11***  ***12***  ***13***  ***14***  ***15*** | ***name=""***  ***price=""***  ***nroFila=0***  ***for fila in tabla.find\_all("tr"):***  ***if nroFila==1:***  ***nroCelda=0***  ***for celda in fila.find\_all('td'):***  ***if nroCelda==0:***  ***name=celda.text***  ***print("Indice:", name)***  ***if nroCelda==2:***  ***price=celda.text***  ***print("Valor:", price)***  ***nroCelda=nroCelda+1***  ***nroFila=nroFila+1*** |

I will see how to exit:

|  |  |
| --- | --- |
| 1 | Index: **IBEX 35® <br>Valor: 9.194.50** |

I only have to save the data to use in the future.

Save CSV and view in Excel

I will assume that I will execute this script once a day, so what I will do is write a new line at the end of the file each time.

|  |  |
| --- | --- |
| ***1***  ***2***  ***3***  ***4*** | ***# Abrimos el csv con append para que pueda agregar contenidos al final del archivo***  ***with open('bolsa\_ibex35.csv', 'a') as csv\_file:***  ***writer = csv.writer(csv\_file)***  ***writer.writerow([name, price, datetime.now()])*** |

Finally I get the file called ***«ibex35.csv*** »ready to be used in our project