### S12 T02: Tasca de web scraping

```
In [2]: import pandas as pd
import numpy as np
import requests
from bs4 import BeautifulSoup
from selenium.webdriver import Firefox
from selenium.webdriver.firefox.options import Options as options
from selenium.webdriver.firefox.service import Service
from selenium import webdriver
from selenium.webdriver.common.by import By
from scrapy import Selector
import scrapy
from scrapy.crawler import CrawlerProcess
```

# Exercici 1. Realitza web scraping d'una pàgina de la borsa de Madrid (<a href="https://www.bolsamadrid.es">https://www.bolsamadrid.es</a> (<a href="https://www.bolsamadrid.es">https://www.bolsamadrid.es</a>)) utilitzant BeautifulSoup i Selenium.

Escollim la pàgina que conté l'última valoració de l'Íbex 35. La url seria la següent:

https://www.bolsamadrid.es/esp/aspx/Mercados/Precios.aspx?indice=ESI100000000 (https://www.bolsamadrid.es/esp/aspx/Mercados/Precios.aspx?indice=ESI100000000)

L'inici de la consulta (query) es fa amb el paràmetre ?; això vol dir que "indice=ESI100000000" està codificat com a "key=value" pairs, i és el resultat de la consulta.

Fem una exploració amb shortcut ctrl+shift+l (tinc linux, però també és vàlid per windows), per veure l'estructura amb la consola del navegador (Firefox). Identifiquem html i javascript i els fulls d'estil css a primer cop d'ull.

#### Web scraping amb Beautiful Soup

```
In [3]: URL = "https://www.bolsamadrid.es/esp/aspx/Mercados/Precios.aspx?indice=ESI100000000"
        page = requests.get(URL)
        print(page.text)
        <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-tra</pre>
        nsitional.dtd">
        <html xmlns="http://www.w3.org/1999/xhtml">
        <head data-idioma="esp" data-hora-act="Sun, 22 May 2022 15:16:05 GMT" data-app-path="/" data-bolsa="BMadri</pre>
        d" data-analytics-id="UA-35966870-2"><meta http-equiv="X-UA-Compatible" content="IE=11" /><meta http-equiv
        ="Content-Type" content="text/html; charset=UTF-8" /><meta id="ctl00 copyright" name="copyright" content="
        Copyright © BME 2022" /><title>
                Bolsa de Madrid - Precios de la sesión
        </title><link id="ctl00 RSSLink1" rel="alternate" type="application/rss+xml" href="/esp/aspx/RSS/RSS.ashx?</pre>
        feed=Todo" title="Bolsa de Madrid: Todos los contenidos agregados" /><link id="ctl00 RSSLink2" rel="altern
        ate" type="application/rss+xml" href="/esp/aspx/RSS/RSS.ashx?feed=NotasPrensa" title="Bolsa de Madrid: Not
        as de Prensa" /><link id="ctl00 RSSLink3" rel="alternate" type="application/rss+xml" href="/esp/aspx/RSS/R
        SS.ashx?feed=Regulacion" title="Bolsa de Madrid: Regulación: Circulares e Instrucciones Operativas" /><lin
        k id="ctl00 RSSLink4" rel="alternate" type="application/rss+xml" href="/esp/aspx/RSS/RSS.ashx?feed=Indice
        s" title="Bolsa de Madrid: Avisos Índices IBEX" /><link id="ctl00 RSSLink5" rel="alternate" type="applicat
        ion/rss+xml" href="/esp/aspx/RSS/RSS.ashx?feed=Revista" title="Bolsa de Madrid: Artículos de la Revista Bo
        lsa" /><script src="/js/Master?v=thBxh-xvKtDQpWbHp0k4A3E9sxEywhoh56HqL0sWH3I1"></script>
        <link href="/css/css?v=e0iHiLxmglVKER4v5gQI2K9pdMEvzhww7tRrYIcX2w41" rel="stylesheet"/>
        <script src="/aspx/Mercados/Precios.js" language="javascript" type="text/javascript"></script></head>
         والمصطالة المائية والمصادر
```

Observem la informació de les empreses integrants de l'Íbex-35, diferents menús, capçaleres i peus de pàgina.

Per exemple, *Telefónica* està inserida amb aquesta programació a dins de la taula (veiem també la seva capçalera).

Es pot fer transformar la presentació de la informació amb pàgines web especialitzades com <u>webformatter (https://webformatter.com/html)</u>, i visualitzar el codi de forma més intel·ligible.

```
<div class="Menu2 MenuDiv Off" menu="m335"><div class="MenuLink MenuConOp Off"><a href="#" target="</pre>
self">Formación</a></div>
<div class="Menu3 MenuSubOp" menu="m335">
<div class="Menu3 MenuDiv Off" menu="m336"><div class="MenuLink Off"><a href="/esp/Inversores/Forma</pre>
cion/QueeslaBolsa.aspx" target=" self">¿Qué es la Bolsa?</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m337"><div class="MenuLink Off"><a href="/esp/aspx/Inversores/"</pre>
Formacion/Formacion.aspx?tipo=T" target=" self">Cursos y Masters</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m338"><div class="MenuLink Off"><a href="/esp/Inversores/Forma</pre>
cion/Invertir/PerfilInversor.aspx" target=" self">Cómo invertir en Bolsa</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m346"><div class="MenuLink Off"><a href="/esp/Inversores/Forma</pre>
cion/Guias/Guias.aspx" target=" self">Guías del inversor</a></div>
<div class="Menu3 MenuDiv Off" menu="m347"><div class="MenuLink Off"><a href="/esp/Inversores/Forma">
cion/Glosario/A.aspx" target=" self">Glosario</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m375"><div class="MenuLink Off"><a href="/esp/Inversores/Forma">
cion/Invertir/FAQS.aspx" target=" self">Preguntas Frecuentes</a></div></div>
</div>
<div class="Menu2 MenuDiv Off" menu="m376"><div class="MenuLink Off"><a href="/esp/Inversores/Merca">
doEsp.aspx" target=" self">El Mercado de Valores Español</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m380"><div class="MenuLink Off"><a href="http://biblioteca.bol">
sasymercados.es/opac/" target=" blank">Biblioteca BME</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m381"><div class="MenuLink Off"><a href="https://www.bolsasyme</pre>
rcados.es/esp/Estudios-Publicaciones/Libros-y-Guias" target=" blank">Libros y quías</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m382"><div class="MenuLink Off"><a href="/esp/aspx/Inversores/"</pre>
Agenda/Agenda.aspx" target=" self">Agenda del Inversor</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m388"><div class="MenuLink Off"><a href="/esp/Inversores/ServI</pre>
nversores.aspx" target=" self">Servicios para Inversores</a></div></div>
</div>
<div class="Menu1 MenuDiv Blanco Off" menu="m389"><div class="MenuLink MenuConOp Off"><a href="#" t</pre>
arget=" self">PARA EMPRESAS</a></div></div>
<div class="Menu2 MenuSubOp" menu="m389">
<div class="Menu2 MenuDiv Off" menu="m390"><div class="MenuLink Off"><a href="/esp/ParaEmp/ComoCoti</pre>
zar/ComoCotizar1.aspx" target=" self">Cómo cotizar en Bolsa</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m391"><div class="MenuLink Off"><a href="/esp/ParaEmp/VenCotiz">
ar/VenCotizar1.aspx" target=" self">Ventajas de Cotizar en Bolsa</a></div></div>
```

```
<div class="Menu2 MenuDiv Off" menu="m398"><div class="MenuLink Off"><a href="http://www.bolsasymer</pre>
cados.es/BME 4Companies/esp" target=" blank">BME 4Companies</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m399"><div class="MenuLink Off"><a href="http://www.entornopre">
mercado.es" target=" blank">Entorno Pre Mercado</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m400"><div class="MenuLink Off"><a href="/esp/ParaEmp/Eventos-</pre>
Corporativos.aspx" target=" self">Eventos Corporativos</a></div></div>
</div>
<div class="Menu1 MenuDiv Blanco Off" menu="m401"><div class="MenuLink MenuConOp Off"><a href="#" t</pre>
arget=" self">PARA MIEMBROS Y PROFESIONALES</a></div></div>
<div class="Menu2 MenuSubOp" menu="m401">
<div class="Menu2 MenuDiv Off" menu="m402"><div class="MenuLink Off"><a href="/esp/Miembros/Ibex/IB</pre>
EXContratoLicenciaUso.aspx" target=" self">Contratos IBEX</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m403"><div class="MenuLink Off"><a href="/esp/Miembros/Ibex/Co</pre>
ntacto.aspx" target=" self">Contacto</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m404"><div class="MenuLink Off"><a href="javascript:sacaVentan">
a('/docs/SBolsas/InformesSB/listadodevalores.pdf')" target=" self">Listado de Valores</a></div></di
V>
<div class="Menu2 MenuDiv Off" menu="m405"><div class="MenuLink MenuConOp Off"><a href="#" target="</pre>
self">Modelo de Mercado</a></div></div>
<div class="Menu3 MenuSub0p" menu="m405">
<div class="Menu3 MenuDiv Off" menu="m406"><div class="MenuLink Off"><a href="javascript:sacaVentan"</pre>
a('/docs/SBolsas/docsSubidos/SIBE/modelomercado.pdf')" target=" self">Acciones y Derechos</a></div>
</div>
<div class="Menu3 MenuDiv Off" menu="m407"><div class="MenuLink Off"><a href="javascript:sacaVentan">
a('/docs/SBolsas/docsSubidos/SIBE/modelomercadoETF.pdf')" target=" self">ETFs</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m408"><div class="MenuLink Off"><a href="javascript:sacaVentan")</pre>
a('/docs/SBolsas/docsSubidos/SIBE/mdlomdowarrcertotros.pdf')" target=" self">Warrants</a></div></di
٧>
</div>
<div class="Menu2 MenuDiv Off" menu="m409"><div class="MenuLink MenuConOp Off"><a href="/esp/Miembr</pre>
os/AccesoMercado.aspx" target=" self">Acceso al Mercado</a></div></div>
<div class="Menu3 MenuSub0p" menu="m409">
<div class="Menu3 MenuDiv Off" menu="m410"><div class="MenuLink Off"><a href="/esp/Miembros/AplicEx</pre>
ternas.aspx" target=" self">Aplicaciones externas</a></div></div>
```

```
</div>
<div class="Menu2 MenuDiv Off" menu="m411"><div class="MenuLink MenuConOp Off"><a href="#" target="</pre>
self">Tarifas</a></div></div>
<div class="Menu3 MenuSubOp" menu="m411">
<div class="Menu3 MenuDiv Off" menu="m412"><div class="MenuLink Off"><a href="/docs/Miembros/tarifa</pre>
sSB.pdf" target=" blank">Sociedad de Bolsas</a></div></div>
<div class="Menu3 MenuDiv Off" menu="m413"><div class="MenuLink Off"><a href="/docs/Miembros/Tarifa</pre>
sRectoras.pdf" target=" blank">Rectoras</a></div></div>
</div>
<div class="Menu2 MenuDiv Off" menu="m414"><div class="MenuLink Off"><a href="javascript:sacaVentan">
a('/docs/SBolsas/InformesSB/miembrosmercado.pdf')" target=" self">Miembros del Mercado</a></div></d
iv>
<div class="Menu2 MenuDiv Off" menu="m415"><div class="MenuLink Off"><a href="/esp/aspx/Regulacion/</pre>
Regulacion.aspx?cod=SBolsas" target=" self">Regulación</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m416"><div class="MenuLink Off"><a href="javascript:sacaVentan">
a('/docs/Acciones/sect.pdf')" target=" self">Clasificación sectorial</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m417"><div class="MenuLink Off"><a href="/esp/Miembros/Servici</pre>
os/MiembrosServicios.aspx" target=" self">Servicios para miembros y profesionales</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m419"><div class="MenuLink Off"><a href="/esp/aspx/Miembros/MI</pre>
FID-II.aspx" target=" self">MiFID II</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m429"><div class="MenuLink Off"><a href="/esp/Miembros/BMESecu">
ritiesSolutions.aspx" target=" self">BME Securities Solutions</a></div></div>
</div>
<div class="Menu1 MenuDiv Blanco Off" menu="m430"><div class="MenuLink MenuConOp Off"><a href="#" t</pre>
arget=" self">PARA MEDIOS DE COMUNICACIÓN</a></div></div>
<div class="Menu2 MenuSubOp" menu="m430">
<div class="Menu2 MenuDiv Off" menu="m431"><div class="MenuLink Off"><a href="/esp/aspx/Prensa/Sala</pre>
Prensa.aspx?tipo=07" target=" self">Notas de Prensa</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m432"><div class="MenuLink Off"><a href="/esp/aspx/Prensa/Cont</pre>
acte.aspx" target=" self">Contacto de Prensa</a></div></div>
<div class="Menu2 MenuDiv Off" menu="m433"><div class="MenuLink Off"><a href="http://www.bolsasymer">
cados.es/esp/Sala-Comunicacion" target="Prensa">Sala de Prensa</a></div></div>
</div>
</div>
```

```
</div>
<script language="javascript">
var MenuIzgLoad MenusSel = [ "m196","m197","m198" ];
var MenuIzqLoad MenusDespl = [ ];
MenuIzqLoad(false, 'General');
</script>
<div class="col6"><form method="post" action="/esp/aspx/Mercados/Precios.aspx?indice=ESI100000000"</pre>
id="aspnetForm">
<input type="hidden" name=" VIEWSTATE" id=" VIEWSTATE" value="jGhsci5wqdIrzZqch+0DS5YQa68TFABnHjX</pre>
sAwTRqpNSHLwDJAxmKU3Rh0qwrPlzXhQlWStonlDGb7uw6reUR8L0yJomwY3t1Tuwq6q2qJbH7UdvtqQ7iUGAo4PT+qqhNM9wkk
ndV5V/nzhDCKh0tlQaf3WcHu5rE2tdvSCvt5Hs5/4T6HQzG5kTivPJiohWqhqreKhzYHAnpEosEH8mFfYit2yiRYswXY7qZNfZj
TnfxTCWsSF54qPQTWkbqDv9dmohvw==" />
<input type="hidden" name=" VIEWSTATEGENERATOR" id=" VIEWSTATEGENERATOR" value="DC05C2DC" />
<input type="hidden" name=" EVENTVALIDATION" id=" EVENTVALIDATION" value="fh96g569rPozDIuBsdEXy2s</pre>
rM4y79F1nnhPP0jRNp6rK+rnxUql7vDKjdFYa+AhVI0qhvIQcj9W7IExG7PbszXyJGxBAhH4sXFa6jsoi7A6NciKv+vVlUZ4pnq
Q387tiuLv83xIFrqlnejwkIcF2z/q16qC232d/5pnIIpHI8jNzxOtZqB8RkOm8pRV63hbwvIZLAVHstitsZkZi+laDrR8X4LBAs
0E4QpPSWcIYzyenrSYeVgWRp5mpFi0AJl9LEvEnSa8s+CnNT1w83DgVkSPm7rhWWIlHAyuG+0j0MupgegAUwkLS8l3JHPIuxQwG
N/5flR3Tz4R7Xs0zRj9v8EXssgrDv7udYhve09ehHdm02HoTYZDzRFuem8L1hs63gwVt+3InmlvtquwNWMn3iZjfW9T2dB0i6LU
qZdaIXsnkAkBbSKoNDW72GuZxU3tbklRn7Vlt89bpvliu7ur5y5ceBr8jIHVwxs3stp0zZ8NJ5m49y+TfPckMjsEx0MHrdA5e9D
RfK1iwoITwIuyLkJbZJ2xYsqHyKXWPMc6z/ny+yWxcL/Vi8Pn6bBxPKj4dX7xNbWy9PgzRSC9dMYDe2+bxbu9lrYuh+w0lbM3Yz
qv04q4P9r/4mAw24Zvct6oV2dL/bMdtnBBTxy/IUV6qP9R5LBAN8xtcFGo+y8z5nOeqmS2GTZHtqoV57r//JGNNaIzDqMGNnjk5
enmFMOjabO3FVKCaO3dN2mnVmbxbpGwrq9YoYQtPMHEGu3mIKSTtmAhK2ywo1D3IoqRud0a70GhMXhxTALHdZL3nUKUqbfqCvXd
4uUek4fLpBhHhkq1jqN49Gq==" />
<div class="TituloPag">Precios de la sesión</div>
Mercado <select name="ctl00$Contenido$SelMercado" id
="SelMercado" onchange="rstSel(0);">
```

```
<option selected="selected" value="" style="color: #999999;">Seleccione mercado</option>
   <option value="MC">Mercado Continuo</option>
   <option value="ETF">ETFs</option>
   <option value="LT">Latibex</option>
   <option value="GROWTH">BME Growth
   <option value="CMAD">Corros Madrid
   <option value="CBAR">Corros Barcelona</option>
   <option value="CBIL">Corros Bilbao</option>
   <option value="CVAL">Corros Valencia</option>
</select> 
 Índice <select name="ctl00$Contenido$SelIndice" id="S
elIndice" onchange="rstSel(1);">
   <option value="" style="color: #999999;">Seleccione &#237;ndice</option>
   <option selected="selected" value="ESI100000000">IBEX 35&#174;</option>
   <option value="ESI060000000">IBEX MEDIUM CAP&#174;</option>
   <option value="ESI240000000">IBEX SMALL CAP&#174;</option>
   <option value="ESI260000000">IBEX TOP Dividendo&#174;</option>
   <option value="ES0S00000901">IBEX 35&#174; Bancos/option>
   <option value="ES0S00000919">IBEX 35&#174; Energ&#237;a/option>
   <option value="ES0S00000927">IBEX 35&#174; Construcci&#243;n/option>
   <option value="ES0S00001586">IBEX Gender Equality/option>
   <option value="ES0S00001594">IBEX Gender Equality Total Return
   <option value="ES0S00001602">IBEX Gender Equality Net Return
   <option value="ESI230000000">FTSE4Good IBEX</option>
   <option value="ESI200000000">FTSE Latibex All Share
   <option value="ESI220000000">FTSE Latibex TOP</option>
   <option value="ESI250000000">FTSE Latibex BRASIL
</select> 
 Sector <select name="ctl00$Contenido$SelSector" id="S
elSector" onchange="rstSel(2);">
   <option value="" style="color: #999999;">Selectione sector</option>
   <option value="01">Petr&#243;leo v Energ&#237;a</option>
```

```
<option value="02">Mat.Basicos, Industria y Construcci&#243;n/option>
  <option value="03">Bienes de Consumo</option>
  <option value="04">Servicios de Consumo</option>
  <option value="05">Servicios Financieros</option>
  <option value="06">Tecnolog&#237;a y Telecomunicaciones/option>
  <option value="07">Servicios Inmobiliarios</option>
</select> 
 <input type="submit" name="ctl00$Contenido$Consultar" value="Consultar" onclick="return chkSel</pre>
(); "id="ctl00 Contenido Consultar" />
 <br /><br />
<div class="LinkAntSig">
<div class="Ctr">IBEX 35®</div>
<div class="finFila"></div>
</div>
< div >
  <table class="TblPort" cellspacing="0" cellpadding="3" align="Center" id="ctl00 Contenido tblÍn
dice" width="80%" style="margin-bottom: 20px;">
     NombreAnteriorÚltimo
>% Dif.MáximoMínimo<th sco
pe="col">FechaHora% Dif.<br/>Año 2022
     d class="DifClSb">0,938.546,908.456,1020/05/2022td><td
align="center">17:38:00-2,63
```

```
</div>
<div>
    100%">
        NombreÚlt.% Dif.<th</th>
scope="col">M\&\#225; x.scope="col">M\&\#237; n.scope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="col">Volumenscope="c
>Efectivo (miles €)FechaHora
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES012")</pre>
5220311">ACCIONA</a>183,10001000182,900
089.92116.536,7720/05/202220/05/202220/05/202220/05/2022
ter">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES013")</pre>
2105018">ACERINOX</a>10,6500class="DifClBj">-0,7010,950010,6250
877.0259.480,3220/05/2022
er">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES016")</pre>
7050915">ACS</a>25,4300class="DifClSb">1,0725,840025,3300
624.55715.948,5320/05/2022
Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES010")</pre>
5046009">AENA</a>135,80001,46136,3000133,4000
td>100.81614.174,9720/05/2022<td class="Ult" align="center"
r">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES015")</pre>
7097017">ALMIRALL</a>10.5800class="DifClSb">1.7310.600010.3700
346.4303.642,88align="center">20/05/2022class="Ult" align="center"
```

```
er">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES010")</pre>
9067019">AMADEUS</a>59,10000,9660,040058,7600
td>516.07130.610,30align="center">20/05/2022<td class="Ult" align="center"
r">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=LU159"]</pre>
8757687">ARCELORMIT.</a>27,4000class="DifClSb">1,0328,045027,33
00615.96517.103,2020/05/2022<td class="Ult" align="c
enter">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES011")</pre>
3900J37">B.SANTANDER</a>2,7440class="DifClSb">0,132,79052,7320
40.067.111110.999,2920/05/2022<td class="Ult" align
="center">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES011")</pre>
3860A34">BA.SABADELL</a>0,7314-0,570,75040,7314
22.843.01116.918,40align="center">20/05/2022class="Ult" align="
center">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES011")</pre>
3679I37">BANKINTER</a>5,5640class="DifClBi">-0,685,67805,5560
td>2.052.97811.519,40align="center">20/05/2022class="Ult" align="center">20/05/2022
ter">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES011</pre>
3211835">BBVA</a>4,7475-0,404,82854,7355+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+td>+
d>24.941.682119.088,6120/05/2022<td class="Ult" align="center"
r">Cierre
        <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES014">
0609019">CAIXABANK</a>3,1200-0,163,184040
```

```
td>18.086.44756.753,7320/05/202220/05/202256.753,73
nter">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES010")</pre>
5066007">CELLNEX</a>40,4200class="DifClSb">2,2040,890039,4100
td>934.47937.760,7420/05/2022<td class="Ult" align="center"
r">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES010")</pre>
5630315">CIE AUTOMOT.</a>21,3000class="DifClSb">0,9521,700021,1
400258.0605.525,1620/05/2022<td class="Ult" align="c
enter">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES013")</pre>
0960018">ENAGAS</a>21,47000,2821,630021,2900
d>831.19417.823,6420/05/2022<td class="Ult" align="center"
r">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES013")</pre>
0670112">ENDESA</a>20,6700class="DifClSb">0,0520,920020,5800
d>902.92418.693,3320/05/2022<td class="Ult" align="center"
r">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES011")</pre>
8900010">FERROVIAL</a>23,3100class="DifClSb">0,5623,640023,1100
791.47618.479,1320/05/202220/05/2022
ter">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES013")</pre>
7650018">FLUIDRA</a>22,7800class="DifClSb">1,2423,240022,5400
td>310.4757.128,0320/05/2022<td class="Ult" align="center"
r">Cierre
    <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES017")</pre>
```

1996087">GRIFOLS CL.A</a>19,8800-0,1520,180019, 6800779.00615.476,2320/05/2022<td class="Ult" align ="center">Cierre <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES017</pre> 7542018">IAG</a>1,4790class="DifClSb">2,211,50151,4540+td>+td 14.614.24321.631,2920/05/2022C ierre <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES014")</pre> 4580Y14">IBERDROLA</a>11,10003,1111,180010,8600 22.219.036241.268,1520/05/202220/05/2022 ="center">Cierre <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES014")</pre> 8396007">INDITEX</a>20,4200class="DifClSb">0,5420,710020,0600 td>3.627.42974.407,1720/05/202220/05/2022 ter">Cierre 

< a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0118594417">INDRA A</a><math display="block"> 10<td

< a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES012 4244E34">MAPFRE</a>1,7730111,80901,7730111

< a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0176252718">MELIA HOTELS</a>6252718">MELIA HOTELS</a>6252718">DifClSb">0,947,07006252718">Class="DifClSb">0,947,07006252718">Class="DifClSb">0,9462

 $<td\ class="DifFlSb"\ align="left"><a\ href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES016">0501022">PHARMA\ MAR</a><math display="block">1.18</t$ 

 $<td\ class="DifFlSb"\ align="left"><a\ href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0173093024">R.E.C.</a><math display="block">19,5250<math display="block">100$ 

```
<a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES014")</pre>
3416115">SIEMENS GAME</a>16,745016,4
5502.112.90435.052,2920/05/202220/05/2022
="center">Cierre
     <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES016")</pre>
5386014">SOLARIA</a>21,8600class="DifClBi">-3,7022,760021,7100
614.09013.580,9720/05/202220/05/2022
ter">Cierre
     <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES017</pre>
8430E18">TELEFONICA</a>4,69201,104,70804,6310
td>10.879.20350.798,3920/05/202220/05/202220/05/202220/05/2022
nter">Cierre
     </div>
<div class="Nota">
  Información diferida 15 minutos.<br />
<u>Precios</u> expresados en euros.<br />
<u>Efectivo</u> expresado en miles de euros.<br />
El <u>volumen y efectivo</u> para cada valor incluye todas las operaciones realizadas hasta el cier
re de la sesión de contratación. El <u>volumen y efectivo total</u> incluyendo las operaciones espe
ciales realizadas después del cierre de la sesión está disponible en la consulta histórica.<br/>>br />
</div>
</form></div>
```

```
<div class="finFila"></div>
     <div id="capaIconosRRSS" class="noimpr">
        Síquenos en:<br />
        class="mclick"><a href="https://twitter.com/GrupoBME" target=" blank">&nbsp;<im
g src="/images/Twitter.png" border="0" align="absmiddle" alt="Twitter" /> </a>class="
mclick"><a href="https://www.flickr.com/photos/archivobme/" target=" blank">&nbsp;<img src="/images</pre>
/Flickr.png" border="0" align="absmiddle" alt="Flickr" /> </a><a href
="https://www.linkedin.com/company/grupo-bme" target=" blank"> <img src="/images/Linkedin.png"
border="0" align="absmiddle" alt="Linkedin" /> </a><a href="https://ww
w.youtube.com/user/GrupoBME?ob=0&feature=results main" target=" blank"> <img src="/images/Yout
ube.png" border="0" align="absmiddle" alt="YouTube" /> </a>
     </div>
     <div>
        <div id="PieOps" class="noimpr"><a href="/esp/aspx/Market-Status"</pre>
/Market-Status.aspx" target=" self"> Estado del Mercado </a><a hr
ef="/esp/aspx/Comun/MapaWeb.aspx" target=" self"> Mapa web </a><a
href="javascript:Disclaimer('esp')" target=" self"> Disclaimer </a>class="mclic"
k"><a href="/esp/privacidad/Politica-de-Privacidad" target=" self">&nbsp;Política de Privacidad&nbs
p;</a><a href="/esp/privacidad/Politica-de-cookies" target=" self">&nbsp;Po
lítica de cookies </a><a href="/esp/aspx/Privacidad/Aviso-Legal.aspx?C
odBolsa=BMadrid" target=" self"> Aviso legal </a></div>
        <div id="PieWebs" class="noimpr"><form id="formWeb">
           <select name="ctl00$webs" id="ctl00 webs">
  <option value="">Webs del grupo BME</option>
```

```
<option value="https://www.bolsasymercados.es">BME</option>
```

Creem un Beutiful Soup object, amb "html.parser" com a analitzador (sintàctic) del contingut de la pàgina.

```
In [4]: soup = BeautifulSoup(page.content, "html.parser")
In [5]: print(soup)
        <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-tra</pre>
        nsitional.dtd">
        <html xmlns="http://www.w3.org/1999/xhtml">
        <head data-analytics-id="UA-35966870-2" data-app-path="/" data-bolsa="BMadrid" data-hora-act="Sun, 22 May</pre>
        2022 15:16:05 GMT" data-idioma="esp"><meta content="IE=11" http-equiv="X-UA-Compatible"/><meta content="te
        xt/html; charset=utf-8" http-equiv="Content-Type"/><meta content="Copyright © BME 2022" id="ctl00 copyright"
        t" name="copyright"/><title>
                Bolsa de Madrid - Precios de la sesión
        </title><link href="/esp/aspx/RSS/RSS.ashx?feed=Todo" id="ctl00 RSSLink1" rel="alternate" title="Bolsa de</pre>
        Madrid: Todos los contenidos agregados" type="application/rss+xml"/><link href="/esp/aspx/RSS/RSS.ashx?fee"
        d=NotasPrensa" id="ctl00 RSSLink2" rel="alternate" title="Bolsa de Madrid: Notas de Prensa" type="applicat
        ion/rss+xml"/><link href="/esp/aspx/RSS/RSS.ashx?feed=Regulacion" id="ctl00 RSSLink3" rel="alternate" titl
        e="Bolsa de Madrid: Regulación: Circulares e Instrucciones Operativas" type="application/rss+xml"/><link h
        ref="/esp/aspx/RSS/RSS.ashx?feed=Indices" id="ctl00 RSSLink4" rel="alternate" title="Bolsa de Madrid: Avis
        os Índices IBEX" type="application/rss+xml"/><link href="/esp/aspx/RSS/RSS.ashx?feed=Revista" id="ctl00 RS
        SLink5" rel="alternate" title="Bolsa de Madrid: Artículos de la Revista Bolsa" type="application/rss+xml"/
        ><script src="/js/Master?v=thBxh-xvKtDQpWbHp0k4A3E9sxEywhoh56HqL0sWH3I1"></script>
        <link href="/css/css?v=e0iHiLxmglVKER4v5gQI2K9pdMEvzhww7tRrYIcX2w41" rel="stylesheet"/>
```

div

Veiem un resultat molt semblant al aconseguit amb [webformatter].

Volem llistar les empreses de l'Íbex-35, les quals estan tabulades sota els següents elements:

```
table class="TblPort"
               id="ctl00 Contenido tblÍndice"
               id="ctl00 Contenido tblAcciones"
               Nombre
In [6]: results = soup.find(id='ctl00 Contenido tblAcciones')
In [7]: print(results)
           NombreÚlt.% Dif.Máx.<t</pre>
           h scope="col">Mín.VolumenEfectivo (miles €)
           FechaHora
           <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0125220311">ACCIONA</a
          >183,1000183,1000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000183,9000
           6.536,7720/05/2022Cierre
           <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0132105018">ACERINOX</
           a>10,6500625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250625062506250
           9.480,3220/05/2022Cierre
           <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0167050915">ACS</a></t
           d>25,4300624.55715.948,
           5320/05/2022Cierre
           <a href="/esp/aspx/Empresas/FichaValor.aspx?ISIN=ES0105046009">AENA</a></
In [8]: print(results.prettify())
```

```
Nombre
      Últ.
      % Dif.
      Máx.
      Mín.
      z+h ccono="col">
In [9]: #Observem les columnes de la capçalera i les llistem
     ibex elements = results.find all("th")
In [10]: print(ibex elements)
     [Nombre, Últ., % Dif., Má
     x., Mín., Volumen, Efectivo (miles €),
     Fecha, Hora]
In [11]: #Extraiem els títols de les columnes
     index col=[]
     for element in results.find all("th"):
       index col.append(element.get text())
     print(index col)
     ['Nombre', 'Últ.', '% Dif.', 'Máx.', 'Mín.', 'Volumen', 'Efectivo (miles €)', 'Fecha', 'Hora']
In [12]: files = []
     for element in results.find_all("td"):
```

files.append(element.text)

print(files) ['ACCIONA', '183,1000', '0,83', '185,3000', '182,9000', '89.921', '16.536,77', '20/05/2022', 'Cierre', 'AC ERINOX', '10,6500', '-0,70', '10,9500', '10,6250', '877.025', '9.480,32', '20/05/2022', 'Cierre', 'ACS', 25,4300', '1,07', '25,8400', '25,3300', '624.557', '15.948,53', '20/05/2022', 'Cierre', 'AENA', '135,8000 ', '1,46', '136,3000', '133,4000', '100.816', '14.174,97', '20/05/2022', 'Cierre', 'ALMIRALL', '10,5800', '1,73', '10,6000', '10,3700', '346.430', '3.642,88', '20/05/2022', 'Cierre', 'AMADEUS', '59,1000', '0,96', '60,0400', '58,7600', '516.071', '30.610,30', '20/05/2022', 'Cierre', 'ARCELORMIT.', '27,4000', '1,03', '2 8,0450', '27,3300', '615.965', '17.103,20', '20/05/2022', 'Cierre', 'B.SANTANDER', '2,7440', '0,13', '2,79 05', '2,7320', '40.067.111', '110.999,29', '20/05/2022', 'Cierre', 'BA.SABADELL', '0,7314', '-0,57', '0,75 04', '0,7314', '22.843.011', '16.918,40', '20/05/2022', 'Cierre', 'BANKINTER', '5,5640', '-0,68', '5,6780 ', '5,5560', '2.052.978', '11.519,40', '20/05/2022', 'Cierre', 'BBVA', '4,7475', '-0,40', '4,8285', '4,735 5', '24.941.682', '119.088,61', '20/05/2022', 'Cierre', 'CAIXABANK', '3,1200', '-0,16', '3,1840', '3,1090 '18.086.447', '56.753,73', '20/05/2022', 'Cierre', 'CELLNEX', '40,4200', '2,20', '40,8900', '39,4100', '934.479', '37.760,74', '20/05/2022', 'Cierre', 'CIE AUTOMOT.', '21,3000', '0,95', '21,7000', '21,1400', ' 258.060', '5.525,16', '20/05/2022', 'Cierre', 'ENAGAS', '21,4700', '0,28', '21,6300', '21,2900', '831.194 ', '17.823,64', '20/05/2022', 'Cierre', 'ENDESA', '20,6700', '0,05', '20,9200', '20,5800', '902.924', '18. 693,33', '20/05/2022', 'Cierre', 'FERROVIAL', '23,3100', '0,56', '23,6400', '23,1100', '791.476', '18.479, 13', '20/05/2022', 'Cierre', 'FLUIDRA', '22,7800', '1,24', '23,2400', '22,5400', '310.475', '7.128,03', '2 0/05/2022', 'Cierre', 'GRIFOLS CL.A', '19,8800', '-0,15', '20,1800', '19,6800', '779.006', '15.476,23', '2 0/05/2022', 'Cierre', 'IAG', '1,4790', '2,21', '1,5015', '1,4540', '14.614.243', '21.631,29', '20/05/2022 ', 'Cierre', 'IBERDROLA', '11,1000', '3,11', '11,1800', '10,8600', '22.219.036', '241.268,15', '20/05/2022 ', 'Cierre', 'INDITEX', '20,4200', '0,54', '20,7100', '20,0600', '3.627.429', '74.407,17', '20/05/2022', Cierre', 'INDRA A', '9,6150', '1,10', '9,7700', '9,4950', '629.913', '6.081,13', '20/05/2022', 'Cierre', ' INM.COLONIAL', '7,6800', '0,92', '7,7150', '7,6000', '571.711', '4.387,52', '20/05/2022', 'Cierre', 'MAPFR E', '1,7730', '-0,11', '1,8090', '1,7730', '4.653.735', '8.316,98', '20/05/2022', 'Cierre', 'MELIA HOTELS ', '6,9650', '0,94', '7,0700', '6,9250', '719.012', '5.023,26', '20/05/2022', 'Cierre', 'MERLIN', '10,3400 ', '0,98', '10,5000', '10,1500', '841.403', '8.740,20', '20/05/2022', 'Cierre', 'NATURGY', '27,6800', '0,5 8', '28,0200', '27,5600', '337.127', '9.346,64', '20/05/2022', 'Cierre', 'PHARMA MAR', '70,1000', '1,18', '70,8000', '69,2400', '48.329', '3.380,44', '20/05/2022', 'Cierre', 'R.E.C.', '19,5250', '0,08', '19,6100 ', '19,4050', '921.833', '17.988,95', '20/05/2022', 'Cierre', 'REPSOL', '14,4100', '-0,21', '14,7500', '1 4,4050', '6.414.647', '92.102,10', '20/05/2022', 'Cierre', 'ROVI', '59,9000', '4,08', '60,6000', '58,1500 ', '131.909', '7.877,93', '20/05/2022', 'Cierre', 'SIEMENS GAME', '16,7450', '1,12', '16,9300', '16,4550', '2.112.904', '35.052,29', '20/05/2022', 'Cierre', 'SOLARIA', '21,8600', '-3,70', '22,7600', '21,7100', '61 4.090', '13.580,97', '20/05/2022', 'Cierre', 'TELEFONICA', '4,6920', '1,10', '4,7080', '4,6310', '10.879.2 03', '50.798,39', '20/05/2022', 'Cierre']

Es pot comprovar que al llistat "files" tenim un canvi cada 9 dades, que equival a les 9 columnes de l'índex. Ara hem d'extreure aquesta informació per crear un dataframe de 9 columnes (atributs), i afegir l'índex a posteriori.

```
In [13]: dades={}
    for i in range(9):
        dades[i]=files[i::9]
    print(dades)
```

{0: ['ACCIONA', 'ACERINOX', 'ACS', 'AENA', 'ALMIRALL', 'AMADEUS', 'ARCELORMIT.', 'B.SANTANDER', 'BA.SABADE LL', 'BANKINTER', 'BBVA', 'CAIXABANK', 'CELLNEX', 'CIE AUTOMOT.', 'ENAGAS', 'ENDESA', 'FERROVIAL', 'FLUIDR A', 'GRIFOLS CL.A', 'IAG', 'IBERDROLA', 'INDITEX', 'INDRA A', 'INM.COLONIAL', 'MAPFRE', 'MELIA HOTELS', 'M ERLIN', 'NATURGY', 'PHARMA MAR', 'R.E.C.', 'REPSOL', 'ROVI', 'SIEMENS GAME', 'SOLARIA', 'TELEFONICA'], 1: ['183,1000', '10,6500', '25,4300', '135,8000', '10,5800', '59,1000', '27,4000', '2,7440', '0,7314', '5,564 0', '4,7475', '3,1200', '40,4200', '21,3000', '21,4700', '20,6700', '23,3100', '22,7800', '19,8800', '1,47 90', '11,1000', '20,4200', '9,6150', '7,6800', '1,7730', '6,9650', '10,3400', '27,6800', '70,1000', '19,52 50', '14,4100', '59,9000', '16,7450', '21,8600', '4,6920'], 2: ['0,83', '-0,70', '1,07', '1,46', '1,73', ' 0,96', '1,03', '0,13', '-0,57', '-0,68', '-0,40', '-0,16', '2,20', '0,95', '0,28', '0,05', '0,56', '1,24', '-0,15', '2,21', '3,11', '0,54', '1,10', '0,92', '-0,11', '0,94', '0,98', '0,58', '1,18', '0,08', '-0,21', '4,08', '1,12', '-3,70', '1,10'], 3: ['185,3000', '10,9500', '25,8400', '136,3000', '10,6000', '60,0400', '28,0450', '2,7905', '0,7504', '5,6780', '4,8285', '3,1840', '40,8900', '21,7000', '21,6300', '20,9200', ' 23,6400', '23,2400', '20,1800', '1,5015', '11,1800', '20,7100', '9,7700', '7,7150', '1,8090', '7,0700', '1 0.5000', '28,0200', '70,8000', '19,6100', '14,7500', '60,6000', '16,9300', '22,7600', '4,7080'], 4: ['182, 9000', '10.6250', '25,3300', '133,4000', '10,3700', '58,7600', '27,3300', '2,7320', '0,7314', '5,5560', ' 4,7355', '3,1090', '39,4100', '21,1400', '21,2900', '20,5800', '23,1100', '22,5400', '19,6800', '1,4540', '10,8600', '20,0600', '9,4950', '7,6000', '1,7730', '6,9250', '10,1500', '27,5600', '69,2400', '19,4050', '14,4050', '58,1500', '16,4550', '21,7100', '4,6310'], 5: ['89.921', '877.025', '624.557', '100.816', '34 6.430', '516.071', '615.965', '40.067.111', '22.843.011', '2.052.978', '24.941.682', '18.086.447', '934.47 9', '258.060', '831.194', '902.924', '791.476', '310.475', '779.006', '14.614.243', '22.219.036', '3.627.4 29', '629.913', '571.711', '4.653.735', '719.012', '841.403', '337.127', '48.329', '921.833', '6.414.647', '131.909', '2.112.904', '614.090', '10.879.203'], 6: ['16.536,77', '9.480,32', '15.948,53', '14.174,97', ' 3.642,88', '30.610,30', '17.103,20', '110.999,29', '16.918,40', '11.519,40', '119.088,61', '56.753,73', '3 7.760,74', '5.525,16', '17.823,64', '18.693,33', '18.479,13', '7.128,03', '15.476,23', '21.631,29', '241.2 68,15', '74.407,17', '6.081,13', '4.387,52', '8.316,98', '5.023,26', '8.740,20', '9.346,64', '3.380,44', ' 17.988,95', '92.102,10', '7.877,93', '35.052,29', '13.580,97', '50.798,39'], 7: ['20/05/2022', '20/05/2022 , '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/0 5/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2 022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '20/05/2022', '2 0/05/2022', '20/05/2022', '20/05/2022'], 8: ['Cierre', 'Cierre', ' Cierre', 'Cierre', 'Cierre ', 'Cierre', 'Ci erre', 'Cierre', 'Cierre', 'Cierre', 'Cierre', 'Cierre', 'Cierre']}

In [14]: #Creem el dataframe i li afegim el títol de les columnes

ibex\_df = pd.DataFrame(dades)
ibex\_df.columns = index\_col

In [15]: ibex\_df

Out[15]:

|    | Nombre       | Últ.     | % Dif. | Máx.     | Mín.     | Volumen    | Efectivo (miles €) | Fecha      | Hora   |
|----|--------------|----------|--------|----------|----------|------------|--------------------|------------|--------|
| 0  | ACCIONA      | 183,1000 | 0,83   | 185,3000 | 182,9000 | 89.921     | 16.536,77          | 20/05/2022 | Cierre |
| 1  | ACERINOX     | 10,6500  | -0,70  | 10,9500  | 10,6250  | 877.025    | 9.480,32           | 20/05/2022 | Cierre |
| 2  | ACS          | 25,4300  | 1,07   | 25,8400  | 25,3300  | 624.557    | 15.948,53          | 20/05/2022 | Cierre |
| 3  | AENA         | 135,8000 | 1,46   | 136,3000 | 133,4000 | 100.816    | 14.174,97          | 20/05/2022 | Cierre |
| 4  | ALMIRALL     | 10,5800  | 1,73   | 10,6000  | 10,3700  | 346.430    | 3.642,88           | 20/05/2022 | Cierre |
| 5  | AMADEUS      | 59,1000  | 0,96   | 60,0400  | 58,7600  | 516.071    | 30.610,30          | 20/05/2022 | Cierre |
| 6  | ARCELORMIT.  | 27,4000  | 1,03   | 28,0450  | 27,3300  | 615.965    | 17.103,20          | 20/05/2022 | Cierre |
| 7  | B.SANTANDER  | 2,7440   | 0,13   | 2,7905   | 2,7320   | 40.067.111 | 110.999,29         | 20/05/2022 | Cierre |
| 8  | BA.SABADELL  | 0,7314   | -0,57  | 0,7504   | 0,7314   | 22.843.011 | 16.918,40          | 20/05/2022 | Cierre |
| 9  | BANKINTER    | 5,5640   | -0,68  | 5,6780   | 5,5560   | 2.052.978  | 11.519,40          | 20/05/2022 | Cierre |
| 10 | BBVA         | 4,7475   | -0,40  | 4,8285   | 4,7355   | 24.941.682 | 119.088,61         | 20/05/2022 | Cierre |
| 11 | CAIXABANK    | 3,1200   | -0,16  | 3,1840   | 3,1090   | 18.086.447 | 56.753,73          | 20/05/2022 | Cierre |
| 12 | CELLNEX      | 40,4200  | 2,20   | 40,8900  | 39,4100  | 934.479    | 37.760,74          | 20/05/2022 | Cierre |
| 13 | CIE AUTOMOT. | 21,3000  | 0,95   | 21,7000  | 21,1400  | 258.060    | 5.525,16           | 20/05/2022 | Cierre |
| 14 | ENAGAS       | 21,4700  | 0,28   | 21,6300  | 21,2900  | 831.194    | 17.823,64          | 20/05/2022 | Cierre |
| 15 | ENDESA       | 20,6700  | 0,05   | 20,9200  | 20,5800  | 902.924    | 18.693,33          | 20/05/2022 | Cierre |
| 16 | FERROVIAL    | 23,3100  | 0,56   | 23,6400  | 23,1100  | 791.476    | 18.479,13          | 20/05/2022 | Cierre |
| 17 | FLUIDRA      | 22,7800  | 1,24   | 23,2400  | 22,5400  | 310.475    | 7.128,03           | 20/05/2022 | Cierre |
| 18 | GRIFOLS CL.A | 19,8800  | -0,15  | 20,1800  | 19,6800  | 779.006    | 15.476,23          | 20/05/2022 | Cierre |
| 19 | IAG          | 1,4790   | 2,21   | 1,5015   | 1,4540   | 14.614.243 | 21.631,29          | 20/05/2022 | Cierre |

|    | Nombre       | Últ.    | % Dif. | Máx.    | Mín.    | Volumen    | Efectivo (miles €) | Fecha      | Hora   |
|----|--------------|---------|--------|---------|---------|------------|--------------------|------------|--------|
| 20 | IBERDROLA    | 11,1000 | 3,11   | 11,1800 | 10,8600 | 22.219.036 | 241.268,15         | 20/05/2022 | Cierre |
| 21 | INDITEX      | 20,4200 | 0,54   | 20,7100 | 20,0600 | 3.627.429  | 74.407,17          | 20/05/2022 | Cierre |
| 22 | INDRA A      | 9,6150  | 1,10   | 9,7700  | 9,4950  | 629.913    | 6.081,13           | 20/05/2022 | Cierre |
| 23 | INM.COLONIAL | 7,6800  | 0,92   | 7,7150  | 7,6000  | 571.711    | 4.387,52           | 20/05/2022 | Cierre |
| 24 | MAPFRE       | 1,7730  | -0,11  | 1,8090  | 1,7730  | 4.653.735  | 8.316,98           | 20/05/2022 | Cierre |
| 25 | MELIA HOTELS | 6,9650  | 0,94   | 7,0700  | 6,9250  | 719.012    | 5.023,26           | 20/05/2022 | Cierre |
| 26 | MERLIN       | 10,3400 | 0,98   | 10,5000 | 10,1500 | 841.403    | 8.740,20           | 20/05/2022 | Cierre |
| 27 | NATURGY      | 27,6800 | 0,58   | 28,0200 | 27,5600 | 337.127    | 9.346,64           | 20/05/2022 | Cierre |
| 28 | PHARMA MAR   | 70,1000 | 1,18   | 70,8000 | 69,2400 | 48.329     | 3.380,44           | 20/05/2022 | Cierre |
| 29 | R.E.C.       | 19,5250 | 0,08   | 19,6100 | 19,4050 | 921.833    | 17.988,95          | 20/05/2022 | Cierre |
| 30 | REPSOL       | 14,4100 | -0,21  | 14,7500 | 14,4050 | 6.414.647  | 92.102,10          | 20/05/2022 | Cierre |
| 31 | ROVI         | 59,9000 | 4,08   | 60,6000 | 58,1500 | 131.909    | 7.877,93           | 20/05/2022 | Cierre |
| 32 | SIEMENS GAME | 16,7450 | 1,12   | 16,9300 | 16,4550 | 2.112.904  | 35.052,29          | 20/05/2022 | Cierre |
| 33 | SOLARIA      | 21,8600 | -3,70  | 22,7600 | 21,7100 | 614.090    | 13.580,97          | 20/05/2022 | Cierre |

```
In [16]: #Guardem el resultat
  ibex_df.to_csv("//home/rusi/Escritorio/rubenIT/DataSources/ibex35_20220222.csv")
```

### Web scraping amb Selenium

```
In [17]: #comprovo la versió Selenium instal·lada.
import selenium
selenium.__version__

Out[17]: '4.1.5'

In [18]: opts = options()
    opts.add_argument("--headless")
    opts.headless # Operating in headless mode
    #assert opts.add argument("--headless")
```

```
browser = Firefox(options=opts)
browser = Firefox(executable path = '/usr/local/bin/geckodriver')
url = 'https://www.bolsamadrid.es'
browser.get(url)
WebDriverException
                                          Traceback (most recent call last)
/tmp/ipykernel 25892/1324337208.py in <module>
      3 opts.headless # Operating in headless mode
      4 #assert opts.add argument("--headless")
----> 5 browser = Firefox(options=opts)
      6 browser = Firefox(executable path = '/usr/local/bin/geckodriver')
      7 url = 'https://www.bolsamadrid.es'
~/anaconda3/lib/python3.9/site-packages/selenium/webdriver/firefox/webdriver.py in init (self, firefox
profile, firefox binary, capabilities, proxy, executable path, options, service log path, service args, se
rvice, desired capabilities, log path, keep alive)
                        service args=service args,
    173
    174
                        log path=service log path)
--> 175
                self.service.start()
    176
    177
                executor = FirefoxRemoteConnection(
~/anaconda3/lib/python3.9/site-packages/selenium/webdriver/common/service.py in start(self)
     96
                count = 0
               while True:
     97
                    self.assert process still running()
---> 98
                    if self.is connectable():
     99
    100
                        break
~/anaconda3/lib/python3.9/site-packages/selenium/webdriver/common/service.py in assert process still runni
na(self)
                return code = self.process.poll()
    108
    109
                if return code:
--> 110
                    raise WebDriverException(
                        'Service %s unexpectedly exited. Status code was: %s'
    111
    112
                        % (self.path, return code)
WebDriverException: Message: Service geckodriver unexpectedly exited. Status code was: 1
```

Després de molts intents, no puc fer servir Selenium, em dona un error. Tanmateix, trobo finalment la solució a:

https://www.youtube.com/watch?v=l\_4Tmb2arVg (https://www.youtube.com/watch?v=l\_4Tmb2arVg)

https://pythonshowcase.com/question/webdriverexception-selenium-common-exceptions-webdriverexception-message-service-webdrivers-geckodriver-unexpectedly-exited-status-code-was-1 (https://pythonshowcase.com/question/webdriverexception-selenium-common-exceptions-webdriverexception-message-service-webdrivers-geckodriver-unexpectedly-exited-status-code-was-1)

```
In [19]: driver_path = '/usr/local/bin/geckodriver'
driver = Firefox(service=Service(driver_path))
driver.get("https://www.bolsamadrid.es")
```

El menú "Acciones" es troba a:

```
<div class="seg"><a href="/?id=esp" target="_self">Inicio</a> / <a href="#" target="_self">MERCADOS
Y COTIZACIONES</a> / <a href="/esp/aspx/Mercados/Precios.aspx?indice=ESI1000000000" target="_self">A
cciones</a> / <span class="segUlt">Precios de la sesión &gt;</span></div>
```

Molt important: cal clickar al link del navegador-selenium abans de seguir els passos!!

 $25 ext{ de } 39$  22/5/22, 16:42

<bound method WebElement.click of <selenium.webdriver.remote.webelement.WebElement (session="b9e2f365-e4bf</pre> In [21]: # un cop a la nova finestra busquem els valors de la taula results = driver.find element by id('ctl00 Contenido tblAcciones') /tmp/ipykernel 25892/180685495.py:2: DeprecationWarning: find element by \* commands are deprecated. Please use find element() instead results = driver.find element by id('ctl00 Contenido tblAcciones') In [22]: #Extraiem els títols de les columnes index col=[] for element in results.find elements by tag name("th"): index col.append(element.text) print(index col) ['Nombre', 'Últ.', '% Dif.', 'Máx.', 'Mín.', 'Volumen', 'Efectivo (miles €)', 'Fecha', 'Hora'] /tmp/ipykernel 25892/690601767.py:4: DeprecationWarning: find elements by tag name is deprecated. Please u se find elements(by=By.TAG NAME, value=name) instead for element in results.find elements by tag name("th"): In [24]: files = [] for element in results.find elements by tag name("td"): files.append(element.text) print(files)

/tmp/ipykernel 25892/3486389068.py:2: DeprecationWarning: find elements by tag name is deprecated. Please

 $26 ext{ de } 39$  22/5/22, 16:42

use find elements(by=By.TAG NAME, value=name) instead

for element in results.find elements by tag name("td"):

```
['ACCIONA', '183,1000', '0,83', '185,3000', '182,9000', '89.921', '16.536,77', '20/05/2022', 'Cierre', 'AC
ERINOX', '10,6500', '-0,70', '10,9500', '10,6250', '877.025', '9.480,32', '20/05/2022', 'Cierre', 'ACS', '
25,4300', '1,07', '25,8400', '25,3300', '624.557', '15.948,53', '20/05/2022', 'Cierre', 'AENA', '135,8000
 , '1,46', '136,3000', '133,4000', '100.816', '14.174,97', '20/05/2022', 'Cierre', 'ALMIRALL', '10,5800',
'1,73', '10,6000', '10,3700', '346.430', '3.642,88', '20/05/2022', 'Cierre', 'AMADEUS', '59,1000', '0,96'
'60,0400', '58,7600', '516.071', '30.610,30', '20/05/2022', 'Cierre', 'ARCELORMIT.', '27,4000', '1,03', '2
8,0450', '27,3300', '615.965', '17.103,20', '20/05/2022', 'Cierre', 'B.SANTANDER', '2,7440', '0,13', '2,79
05', '2,7320', '40.067.111', '110.999,29', '20/05/2022', 'Cierre', 'BA.SABADELL', '0,7314', '-0,57', '0,75
04', '0,7314', '22.843.011', '16.918,40', '20/05/2022', 'Cierre', 'BANKINTER', '5,5640', '-0,68', '5,6780
   '5,5560', '2.052.978', '11.519,40', '20/05/2022', 'Cierre', 'BBVA', '4,7475', '-0,40', '4,8285', '4,735
5', '24.941.682', '119.088,61', '20/05/2022', 'Cierre', 'CAIXABANK', '3,1200', '-0,16', '3,1840', '3,1090
  '18.086.447', '56.753,73', '20/05/2022', 'Cierre', 'CELLNEX', '40,4200', '2,20', '40,8900', '39,4100',
'934.479', '37.760,74', '20/05/2022', 'Cierre', 'CIE AUTOMOT.', '21,3000', '0,95', '21,7000', '21,1400', '
258.060', '5.525,16', '20/05/2022', 'Cierre', 'ENAGAS', '21,4700', '0,28', '21,6300', '21,2900', '831.194
  . '17.823,64', '20/05/2022', 'Cierre', 'ENDESA', '20,6700', '0,05', '20,9200', '20,5800', '902.924', '18.
693,33', '20/05/2022', 'Cierre', 'FERROVIAL', '23,3100', '0,56', '23,6400', '23,1100', '791.476', '18.479,
13', '20/05/2022', 'Cierre', 'FLUIDRA', '22,7800', '1,24', '23,2400', '22,5400', '310.475', '7.128,03', '2
0/05/2022', 'Cierre', 'GRIFOLS CL.A', '19,8800', '-0,15', '20,1800', '19,6800', '779.006', '15.476,23', '2
0/05/2022', 'Cierre', 'IAG', '1,4790', '2,21', '1,5015', '1,4540', '14.614.243', '21.631,29', '20/05/2022
  'Cierre', 'IBERDROLA', '11,1000', '3,11', '11,1800', '10,8600', '22.219.036', '241.268,15', '20/05/2022
', 'Cierre', 'INDITEX', '20,4200', '0,54', '20,7100', '20,0600', '3.627.429', '74.407,17', '20/05/2022', '
Cierre', 'INDRA A', '9,6150', '1,10', '9,7700', '9,4950', '629.913', '6.081,13', '20/05/2022', 'Cierre', '
INM.COLONIAL', '7,6800', '0,92', '7,7150', '7,6000', '571.711', '4.387,52', '20/05/2022', 'Cierre', 'MAPFR
E', '1,7730', '-0,11', '1,8090', '1,7730', '4.653.735', '8.316,98', '20/05/2022', 'Cierre', 'MELIA HOTELS
   '6,9650', '0,94', '7,0700', '6,9250', '719.012', '5.023,26', '20/05/2022', 'Cierre', 'MERLIN', '10,3400
', '0,98', '10,5000', '10,1500', '841.403', '8.740,20', '20/05/2022', 'Cierre', 'NATURGY', '27,6800', '0,5
```

```
In [25]: dades={}
    for i in range(9):
        dades[i]=files[i::9]
    print(dades)
```

{0: ['ACCIONA', 'ACERINOX', 'ACS', 'AENA', 'ALMIRALL', 'AMADEUS', 'ARCELORMIT.', 'B.SANTANDER', 'BA.SABADE LL', 'BANKINTER', 'BBVA', 'CAIXABANK', 'CELLNEX', 'CIE AUTOMOT.', 'ENAGAS', 'ENDESA', 'FERROVIAL', 'FLUIDR A', 'GRIFOLS CL.A', 'IAG', 'IBERDROLA', 'INDITEX', 'INDRA A', 'INM.COLONIAL', 'MAPFRE', 'MELIA HOTELS', 'M ERLIN', 'NATURGY', 'PHARMA MAR', 'R.E.C.', 'REPSOL', 'ROVI', 'SIEMENS GAME', 'SOLARIA', 'TELEFONICA'], 1: ['183,1000', '10,6500', '25,4300', '135,8000', '10,5800', '59,1000', '27,4000', '2,7440', '0,7314', '5,564 0', '4,7475', '3,1200', '40,4200', '21,3000', '21,4700', '20,6700', '23,3100', '22,7800', '19,8800', '1,47 90', '11,1000', '20,4200', '9,6150', '7,6800', '1,7730', '6,9650', '10,3400', '27,6800', '70,1000', '19,52 50', '14,4100', '59,9000', '16,7450', '21,8600', '4,6920'], 2: ['0,83', '-0,70', '1,07', '1,46', '1,73', 0,96', '1,03', '0,13', '-0,57', '-0,68', '-0,40', '-0,16', '2,20', '0,95', '0,28', '0,05', '0,56', '1,24' '-0,15', '2,21', '3,11', '0,54', '1,10', '0,92', '-0,11', '0,94', '0,98', '0,58', '1,18', '0,08', '-0,21', '4,08', '1,12', '-3,70', '1,10'], 3: ['185,3000', '10,9500', '25,8400', '136,3000', '10,6000', '60,0400', '28,0450', '2,7905', '0,7504', '5,6780', '4,8285', '3,1840', '40,8900', '21,7000', '21,6300', '20,9200', ' 23,6400', '23,2400', '20,1800', '1,5015', '11,1800', '20,7100', '9,7700', '7,7150', '1,8090', '7,0700', '1 0,5000', '28,0200', '70,8000', '19,6100', '14,7500', '60,6000', '16,9300', '22,7600', '4,7080'], 4: ['182, 9000', '10,6250', '25,3300', '133,4000', '10,3700', '58,7600', '27,3300', '2,7320', '0,7314', '5,5560', 4,7355', '3,1090', '39,4100', '21,1400', '21,2900', '20,5800', '23,1100', '22,5400', '19,6800', '1,4540', '10,8600', '20,0600', '9,4950', '7,6000', '1,7730', '6,9250', '10,1500', '27,5600', '69,2400', '19,4050', '14,4050', '58,1500', '16,4550', '21,7100', '4,6310'], 5: ['89.921', '877.025', '624.557', '100.816', '34 6.430', '516.071', '615.965', '40.067.111', '22.843.011', '2.052.978', '24.941.682', '18.086.447', '934.47 9', '258.060', '831.194', '902.924', '791.476', '310.475', '779.006', '14.614.243', '22.219.036', '3.627.4 29', '629.913', '571.711', '4.653.735', '719.012', '841.403', '337.127', '48.329', '921.833', '6.414.647', '131.909', '2.112.904', '614.090', '10.879.203'], 6: ['16.536,77', '9.480,32', '15.948,53', '14.174,97', 3.642,88', '30.610,30', '17.103,20', '110.999,29', '16.918,40', '11.519,40', '119.088,61', '56.753,73', '3 7.760,74', '5.525,16', '17.823,64', '18.693,33', '18.479,13', '7.128,03', '15.476,23', '21.631,29', '241.2 68,15', '74.407,17', '6.081,13', '4.387,52', '8.316,98', '5.023,26', '8.740,20', '9.346,64', '3.380,44', 17.988,95', '92.102,10', '7.877,93', '35.052,29', '13.580,97', '50.798,39'], 7: ['20/05/2022', '20/05/2022

## In [26]: #Creem el dataframe i li afegim el títol de les columnes ibex\_df = pd.DataFrame(dades) ibex\_df.columns = index\_col

### In [27]: ibex\_df

| _   |    |       |  |
|-----|----|-------|--|
| (1) | 11 | 17/   |  |
| υu  | ı  | L Z / |  |

|   | Nombre   | Últ.     | % Dif. | Máx.     | Mín.     | Volumen | Efectivo (miles €) | Fecha      | Hora   |
|---|----------|----------|--------|----------|----------|---------|--------------------|------------|--------|
| 0 | ACCIONA  | 183,1000 | 0,83   | 185,3000 | 182,9000 | 89.921  | 16.536,77          | 20/05/2022 | Cierre |
| 1 | ACERINOX | 10,6500  | -0,70  | 10,9500  | 10,6250  | 877.025 | 9.480,32           | 20/05/2022 | Cierre |
| 2 | ACS      | 25,4300  | 1,07   | 25,8400  | 25,3300  | 624.557 | 15.948,53          | 20/05/2022 | Cierre |
| 3 | AENA     | 135,8000 | 1,46   | 136,3000 | 133,4000 | 100.816 | 14.174,97          | 20/05/2022 | Cierre |

|    | Nombre       | Últ.    | % Dif. | Máx.    | Mín.    | Volumen    | Efectivo (miles €) | Fecha      | Hora   |
|----|--------------|---------|--------|---------|---------|------------|--------------------|------------|--------|
| 4  | ALMIRALL     | 10,5800 | 1,73   | 10,6000 | 10,3700 | 346.430    | 3.642,88           | 20/05/2022 | Cierre |
| 5  | AMADEUS      | 59,1000 | 0,96   | 60,0400 | 58,7600 | 516.071    | 30.610,30          | 20/05/2022 | Cierre |
| 6  | ARCELORMIT.  | 27,4000 | 1,03   | 28,0450 | 27,3300 | 615.965    | 17.103,20          | 20/05/2022 | Cierre |
| 7  | B.SANTANDER  | 2,7440  | 0,13   | 2,7905  | 2,7320  | 40.067.111 | 110.999,29         | 20/05/2022 | Cierre |
| 8  | BA.SABADELL  | 0,7314  | -0,57  | 0,7504  | 0,7314  | 22.843.011 | 16.918,40          | 20/05/2022 | Cierre |
| 9  | BANKINTER    | 5,5640  | -0,68  | 5,6780  | 5,5560  | 2.052.978  | 11.519,40          | 20/05/2022 | Cierre |
| 10 | BBVA         | 4,7475  | -0,40  | 4,8285  | 4,7355  | 24.941.682 | 119.088,61         | 20/05/2022 | Cierre |
| 11 | CAIXABANK    | 3,1200  | -0,16  | 3,1840  | 3,1090  | 18.086.447 | 56.753,73          | 20/05/2022 | Cierre |
| 12 | CELLNEX      | 40,4200 | 2,20   | 40,8900 | 39,4100 | 934.479    | 37.760,74          | 20/05/2022 | Cierre |
| 13 | CIE AUTOMOT. | 21,3000 | 0,95   | 21,7000 | 21,1400 | 258.060    | 5.525,16           | 20/05/2022 | Cierre |
| 14 | ENAGAS       | 21,4700 | 0,28   | 21,6300 | 21,2900 | 831.194    | 17.823,64          | 20/05/2022 | Cierre |
| 15 | ENDESA       | 20,6700 | 0,05   | 20,9200 | 20,5800 | 902.924    | 18.693,33          | 20/05/2022 | Cierre |
| 16 | FERROVIAL    | 23,3100 | 0,56   | 23,6400 | 23,1100 | 791.476    | 18.479,13          | 20/05/2022 | Cierre |
| 17 | FLUIDRA      | 22,7800 | 1,24   | 23,2400 | 22,5400 | 310.475    | 7.128,03           | 20/05/2022 | Cierre |
| 18 | GRIFOLS CL.A | 19,8800 | -0,15  | 20,1800 | 19,6800 | 779.006    | 15.476,23          | 20/05/2022 | Cierre |
| 19 | IAG          | 1,4790  | 2,21   | 1,5015  | 1,4540  | 14.614.243 | 21.631,29          | 20/05/2022 | Cierre |
| 20 | IBERDROLA    | 11,1000 | 3,11   | 11,1800 | 10,8600 | 22.219.036 | 241.268,15         | 20/05/2022 | Cierre |
| 21 | INDITEX      | 20,4200 | 0,54   | 20,7100 | 20,0600 | 3.627.429  | 74.407,17          | 20/05/2022 | Cierre |
| 22 | INDRA A      | 9,6150  | 1,10   | 9,7700  | 9,4950  | 629.913    | 6.081,13           | 20/05/2022 | Cierre |
| 23 | INM.COLONIAL | 7,6800  | 0,92   | 7,7150  | 7,6000  | 571.711    | 4.387,52           | 20/05/2022 | Cierre |
| 24 | MAPFRE       | 1,7730  | -0,11  | 1,8090  | 1,7730  | 4.653.735  | 8.316,98           | 20/05/2022 | Cierre |
| 25 | MELIA HOTELS | 6,9650  | 0,94   | 7,0700  | 6,9250  | 719.012    | 5.023,26           | 20/05/2022 | Cierre |
| 26 | MERLIN       | 10,3400 | 0,98   | 10,5000 | 10,1500 | 841.403    | 8.740,20           | 20/05/2022 | Cierre |
| 27 | NATURGY      | 27,6800 | 0,58   | 28,0200 | 27,5600 | 337.127    | 9.346,64           | 20/05/2022 | Cierre |
| 28 | PHARMA MAR   | 70,1000 | 1,18   | 70,8000 | 69,2400 | 48.329     | 3.380,44           | 20/05/2022 | Cierre |

Fecha Hora

Nombre

IÍIt % Dif

Máy

Mín

|          |    | Nonible                         | Oit.    | /6 DII. | IVIAA.   | IVIIII. | Volumen   | Liectivo (illies e) | i eciia    | liola            |  |
|----------|----|---------------------------------|---------|---------|----------|---------|-----------|---------------------|------------|------------------|--|
|          | 29 | R.E.C.                          | 19,5250 | 0,08    | 19,6100  | 19,4050 | 921.833   | 17.988,95           | 20/05/2022 | Cierre           |  |
|          | 30 | REPSOL                          | 14,4100 | -0,21   | 14,7500  | 14,4050 | 6.414.647 | 92.102,10           | 20/05/2022 | Cierre           |  |
|          | 31 | ROVI                            | 59,9000 | 4,08    | 60,6000  | 58,1500 | 131.909   | 7.877,93            | 20/05/2022 | Cierre           |  |
|          | 32 | SIEMENS GAME                    | 16,7450 | 1,12    | 16,9300  | 16,4550 | 2.112.904 | 35.052,29           | 20/05/2022 | Cierre           |  |
| In [28]: |    | wardem el resu<br>x_df.to_csv(" |         | rusi/E  | Escritor | io/rube | nIT/Datas | Sources/ibex35      | 5_2022022  | 2_selenium.csv") |  |

Volumen Ffective (miles £)

### Exercici 2. Documenta en un word el teu conjunt de dades generat amb la informació que tenen els diferents arxius de Kaggle.

```
In [29]: ibex df02=pd.read csv("//home/rusi/Escritorio/rubenIT/DataSources/ibex35 20220222.csv")
In [30]: ibex df02.info()
         <class 'pandas.core.frame.DataFrame'>
         RangeIndex: 35 entries, 0 to 34
         Data columns (total 10 columns):
                                   Non-Null Count
              Column
                                                   Dtype
                                   35 non-null
              Unnamed: 0
                                                    int64
              Nombre
                                   35 non-null
                                                   object
              Últ.
                                   35 non-null
                                                   object
              % Dif.
                                   35 non-null
                                                   object
                                   35 non-null
                                                   object
              Máx.
              Mín.
                                   35 non-null
                                                   object
              Volumen
                                   35 non-null
                                                   object
              Efectivo (miles €) 35 non-null
                                                   object
                                                   object
              Fecha
                                   35 non-null
                                                   object
              Hora
                                   35 non-null
         dtypes: int64(1), object(9)
         memory usage: 2.9+ KB
```

| 31]:                                      | ibex_c   | df02.hea   | ad()  |   |   |                                     |   |                              |   |   |   |                            |       |                      |                  |        |
|---|--|--|---|---|---|-------------------------------------|---|------------------------------|---|---|---|----------------------------|-------|----------------------|------------------|--------|
| 31]:                                      | Unn  | named: 0   | Nombre  | Últ.  | % Dif.  | Máx.                                | Mín.  | Volumen                      | Efectivo (mile  | s€) F   | echa  | Hora                       |       |                      |                  |        |
|   | 0  | 0  | ACCIONA   | 183,1000  | 0,83  | 185,3000                            | 182,9000  | 89.921                       | 16.536  | 6,77 20/05/   | 2022  | Cierre                     |       |                      |                  |        |
|   | 1  | 1 .  | ACERINOX  | 10,6500   | -0,70   | 10,9500                             | 10,6250   | 877.025                      | 9.480   | ),32 20/05/   | 2022  | Cierre                     |       |                      |                  |        |
|   | 2  | 2  | ACS   | 25,4300   | 1,07  | 25,8400                             | 25,3300   | 624.557                      | 15.948  | 3,53 20/05/   | 2022  | Cierre                     |       |                      |                  |        |
|   | 3  | 3  | AENA  | 135,8000  | 1,46  | 136,3000                            | 133,4000  | 100.816                      | 14.174  | 1,97 20/05/   | 2022  | Cierre                     |       |                      |                  |        |
|   | 4  | 4  | ALMIRALL  | 10,5800   | 1,73  | 10,6000                             | 10,3700   | 346.430                      | 3.642   | 2,88 20/05/   | 2022  | Cierre                     |       |                      |                  |        |
|   |  |  | 101111111111111111111111111111111111111   | 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1                                    |   |                                     |   |                              |   |   |   |                            | ١ ٠+٠ | . ronloc             | 0/1 1 1          | ` I\ , |
|   | /tmp/i<br>alse i<br>al sti                             | ipykerne<br>in a fut<br>rings wh<br>x_df02[ˈ                                 | el_25892/<br>ture vers<br>hen regex   | 3383750<br>ion. Ir<br>=True.<br>(miles                                    | )530.py<br>addit  | -<br>/:5: Fut<br>:ion, s:           | tureWarn<br>ingle ch  | ing: Th                      | e default v<br>regular ex   | value of<br>xpressio  | rege<br>ns wi   | ex wi<br>ill *r            | l cha | e treat              | om True<br>ed as | to F   |
|   | /tmp/i<br>alse i<br>al str<br>ibex<br>(',',            | ipykerne<br>in a fut<br>rings wh<br>x_df02['<br>'.').ast                     | el_25892/<br>ture vers<br>hen regex<br>'Efectivo  | 3383750<br>ion. Ir<br>=True.<br>(miles                                    | 0530.py<br>n addit<br>s €)']                                      | <br>/:5: Fut<br>:ion, s:<br>=ibex_c | tureWarn<br>ingle ch  | ing: Th                      | e default v<br>regular ex   | value of<br>xpressio  | rege<br>ns wi   | ex wi<br>ill *r            | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| : [[                                      | /tmp/i<br>alse i<br>al str<br>ibex<br>(',',            | ipykerne<br>in a fut<br>rings wh<br>x_df02['<br>'.').ast                     | el_25892/<br>ture vers<br>hen regex<br>'Efectivo<br>type(floa   | 3383750<br>ion. Ir<br>=True.<br>(miles                                    | 0530.py<br>n addit<br>s €)']                                      | <br>/:5: Fut<br>:ion, s:<br>=ibex_c | tureWarn<br>ingle ch  | ing: Th                      | e default v<br>regular ex   | value of<br>xpressio  | rege<br>ns wi   | ex wi<br>ill *r            | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| <br> ::[::[:::::::::::::::::::::::::::::: | /tmp/i<br>alse i<br>al str<br>ibex<br>(',','           | ipykerne<br>in a fut<br>rings wh<br>x_df02['<br>'.').ast                     | el_25892/<br>ture vers<br>hen regex<br>'Efectivo<br>type(floa<br>ex_df02.d                                  | 3383750<br>ion. Ir<br>=True.<br>(miles                                    | 0530.py<br>n addit<br>s €)']                                      |                                     | tureWarn<br>ingle ch<br>df02['Ef                                      | ing: Th<br>aracter           | e default v<br>regular ex   | value of<br>xpressio  | rege<br>ns wi<br>place                                  | ex wil<br>ill *r<br>e('.', | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| ]: [                                      | /tmp/i<br>alse i<br>al str<br>ibex<br>(',','           | ipykerne<br>in a fut<br>rings wh<br>x_df02['<br>'.').ast<br>df03=ibe         | el_25892/<br>ture vers<br>hen regex<br>'Efectivo<br>type(floa<br>ex_df02.d                                  | 3383750 ion. Ir =True. (miles t) rop("Un                                  | 0530.py<br>n addit<br>s €)']<br>named:                            |                                     | tureWarn ingle ch df02['Ef  | ing: Th<br>aracter           | e default v<br>regular ex<br>(miles €)'   | value of<br>xpressio<br>].str.re                                      | regens wi   | ex wilill *re('.',         | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| ]: [<br>]: [                              | /tmp/i<br>alse i<br>al str<br>ibex<br>(',','           | ipykerne<br>in a fut<br>rings wh<br>x_df02['<br>'.').ast<br>df03=ibe         | el_25892/ ture vers nen regex 'Efectivo type(floa ex_df02.d   | 3383750 ion. Ir =True. (miles t) rop("Un                                  | 0530.py<br>n addit<br>s €)']<br>named:                            |                                     | tureWarn ingle ch df02['Ef Ls=1)  in. Volu 00 89                      | ing: Th<br>aracter<br>ectivo | e default y<br>regular ex<br>(miles €) ';<br>tivo (miles €)   | value of<br>xpressio<br>].str.re                                      | regens wi   | ex wilill *re('.',         | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| ]: [<br>]: [                              | /tmp/i<br>alse i<br>al str<br>ibex_c<br>ibex_c         | ipykerne in a fut rings wh x_df02[' '.').ast df03=ibe df03  Nom  ACCIC       | el_25892/ ture vers hen regex 'Efectivo type(floa ex_df02.d   | 3383756 ion. Ir =True. (miles t) rop("Un  t. % Dif. 0 0.83                | 0530.py<br>n addit<br>s €)']<br>named:<br>Má                      |                                     | tureWarn ingle ch df02['Ef ls=1)  in. Volu 00 89 50 877               | eing: The laracter sectivo   | e default y<br>regular ex<br>(miles €) ';<br>tivo (miles €)   | value of xpressio ].str.re  Fecha 20/05/2022                          | regens wi   | ex wilill *re('.',         | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| ]: [                                      | /tmp/i<br>alse i<br>al str<br>ibex<br>(',','           | ipykerne in a fut rings wh x_df02[' '.').ast df03=ibe df03  Nom ACCIC ACERIN | el_25892/ ture vers nen regex 'Efectivo type(floa ex_df02.d   | 3383756 ion. Ir =True. (miles t) rop("Un  t. % Dif. 0 0.83                | 0530.py<br>n addit<br>s €)']<br>named:<br>Má<br>185.300<br>10.950 |                                     | tureWarn ingle ch df02['Ef ls=1)  in. Volu 00 89 50 877 00 624        | men Efec                     | e default vergular extends from the regular e | Value of Expression of Str.re  Fecha 20/05/2022 20/05/2022 20/05/2022 | regens wind place  Hora  Cierre  Cierre  Cierre         | ex wilill *re('.',         | l cha | ange fro<br>be treat | om True<br>ed as | to F   |
| ]: [                                      | /tmp/i<br>alse i<br>al str<br>ibex<br>(',','<br>ibex_c | ipykerne in a fut rings wh x_df02[' '.').ast df03=ibe df03  Nom ACCIC ACERIN | el_25892/ ture vers hen regex 'Efectivo type(floa ex_df02.dex DNA 183.100 NOX 10.650 ACS 25.430 ENA 135.800 | 3383756 ion. Ir =True. (miles t) rop("Un  t. % Dif. 0 0.83 0 -0.70 0 1.07 | 0530.py<br>n addit<br>s €)']<br>named:<br>Má<br>185.300<br>10.950 |                                     | tureWarn ingle ch df02['Ef ls=1)  in. Volu 00 89 50 877 00 624 00 100 | men Efec<br>.921<br>.025     | e default (regular extended from the second f | Value of Expression of Str.re  Fecha 20/05/2022 20/05/2022 20/05/2022 | regens wind place  Hora  Cierre  Cierre  Cierre  Cierre | ex wilill *re('.',         | l cha | ange fro<br>be treat | om True<br>ed as | to F   |

| Nombre       | Últ.   | % Dif.   | Máx.   | Mín.  | Volumen  | Efectivo (miles €)  | Fecha  | Hora   |
|--------------|--|--|--|---|--|---|--|--|
| ARCELORMIT.  | 27.4000  | 1.03   | 28.0450  | 27.3300   | 615.965  | 17103.20  | 20/05/2022   | Cierre   |
| B.SANTANDER  | 2.7440   | 0.13   | 2.7905   | 2.7320  | 40.067.111   | 110999.29   | 20/05/2022   | Cierre   |
| BA.SABADELL  | 0.7314   | -0.57  | 0.7504   | 0.7314  | 22.843.011   | 16918.40  | 20/05/2022   | Cierre   |
| BANKINTER    | 5.5640   | -0.68  | 5.6780   | 5.5560  | 2.052.978  | 11519.40  | 20/05/2022   | Cierre   |
| BBVA         | 4.7475   | -0.40  | 4.8285   | 4.7355  | 24.941.682   | 119088.61   | 20/05/2022   | Cierre   |
| CAIXABANK    | 3.1200   | -0.16  | 3.1840   | 3.1090  | 18.086.447   | 56753.73  | 20/05/2022   | Cierre   |
| CELLNEX      | 40.4200  | 2.20   | 40.8900  | 39.4100   | 934.479  | 37760.74  | 20/05/2022   | Cierre   |
| CIE AUTOMOT. | 21.3000  | 0.95   | 21.7000  | 21.1400   | 258.060  | 5525.16   | 20/05/2022   | Cierre   |
| ENAGAS       | 21.4700  | 0.28   | 21.6300  | 21.2900   | 831.194  | 17823.64  | 20/05/2022   | Cierre   |
| ENDESA       | 20.6700  | 0.05   | 20.9200  | 20.5800   | 902.924  | 18693.33  | 20/05/2022   | Cierre   |
| FERROVIAL    | 23.3100  | 0.56   | 23.6400  | 23.1100   | 791.476  | 18479.13  | 20/05/2022   | Cierre   |
| FLUIDRA      | 22.7800  | 1.24   | 23.2400  | 22.5400   | 310.475  | 7128.03   | 20/05/2022   | Cierre   |
| GRIFOLS CL.A | 19.8800  | -0.15  | 20.1800  | 19.6800   | 779.006  | 15476.23  | 20/05/2022   | Cierre   |
| IAG          | 1.4790   | 2.21   | 1.5015   | 1.4540  | 14.614.243   | 21631.29  | 20/05/2022   | Cierre   |
| IBERDROLA    | 11.1000  | 3.11   | 11.1800  | 10.8600   | 22.219.036   | 241268.15   | 20/05/2022   | Cierre   |
| INDITEX      | 20.4200  | 0.54   | 20.7100  | 20.0600   | 3.627.429  | 74407.17  | 20/05/2022   | Cierre   |
| INDRA A      | 9.6150   | 1.10   | 9.7700   | 9.4950  | 629.913  | 6081.13   | 20/05/2022   | Cierre   |
| INM.COLONIAL | 7.6800   | 0.92   | 7.7150   | 7.6000  | 571.711  | 4387.52   | 20/05/2022   | Cierre   |
| MAPFRE       | 1.7730   | -0.11  | 1.8090   | 1.7730  | 4.653.735  | 8316.98   | 20/05/2022   | Cierre   |
| MELIA HOTELS | 6.9650   | 0.94   | 7.0700   | 6.9250  | 719.012  | 5023.26   | 20/05/2022   | Cierre   |
| MERLIN       | 10.3400  | 0.98   | 10.5000  | 10.1500   | 841.403  | 8740.20   | 20/05/2022   | Cierre   |
| NATURGY      | 27.6800  | 0.58   | 28.0200  | 27.5600   | 337.127  | 9346.64   | 20/05/2022   | Cierre   |
| PHARMA MAR   | 70.1000  | 1.18   | 70.8000  | 69.2400   | 48.329   | 3380.44   | 20/05/2022   | Cierre   |
| R.E.C.       | 19.5250  | 0.08   | 19.6100  | 19.4050   | 921.833  | 17988.95  | 20/05/2022   | Cierre   |
| REPSOL       | 14.4100  | -0.21  | 14.7500  | 14.4050   | 6.414.647  | 92102.10  | 20/05/2022   | Cierre   |
|              | ARCELORMIT. B.SANTANDER BA.SABADELL BANKINTER BBVA CAIXABANK CELLNEX CIE AUTOMOT. ENAGAS ENDESA FERROVIAL FLUIDRA GRIFOLS CL.A IAG IBERDROLA INDITEX INDRA A INM.COLONIAL MAPFRE MELIA HOTELS MERLIN NATURGY PHARMA MAR R.E.C. | ARCELORMIT. 27.4000 B.SANTANDER 2.7440 BA.SABADELL 0.7314 BANKINTER 5.5640 BBVA 4.7475 CAIXABANK 3.1200 CELLNEX 40.4200 CIE AUTOMOT. 21.3000 ENAGAS 21.4700 ENDESA 20.6700 FERROVIAL 23.3100 FLUIDRA 22.7800 GRIFOLS CL.A 19.8800 IAG 1.4790 IBERDROLA 11.1000 INDITEX 20.4200 INDRA A 9.6150 INM.COLONIAL 7.6800 MAPFRE 1.7730 MELIA HOTELS 6.9650 MERLIN 10.3400 NATURGY 27.6800 PHARMA MAR 70.1000 R.E.C. 19.5250 | ARCELORMIT.       27.4000       1.03         B.SANTANDER       2.7440       0.13         BA.SABADELL       0.7314       -0.57         BANKINTER       5.5640       -0.68         BBVA       4.7475       -0.40         CAIXABANK       3.1200       -0.16         CELLNEX       40.4200       2.20         CIE AUTOMOT.       21.3000       0.95         ENAGAS       21.4700       0.28         ENDESA       20.6700       0.05         FERROVIAL       23.3100       0.56         FLUIDRA       22.7800       1.24         GRIFOLS CL.A       19.8800       -0.15         IAG       1.4790       2.21         IBERDROLA       11.1000       3.11         INDITEX       20.4200       0.54         INDRA A       9.6150       1.10         INM.COLONIAL       7.6800       0.92         MAPFRE       1.7730       -0.11         MELIA HOTELS       6.9650       0.94         MERLIN       10.3400       0.98         NATURGY       27.6800       0.58         PHARMA MAR       70.1000       1.18         R.E.C.       19.5250 <th>ARCELORMIT.       27.4000       1.03       28.0450         B.SANTANDER       2.7440       0.13       2.7905         BA.SABADELL       0.7314       -0.57       0.7504         BANKINTER       5.5640       -0.68       5.6780         BBVA       4.7475       -0.40       4.8285         CAIXABANK       3.1200       -0.16       3.1840         CELLNEX       40.4200       2.20       40.8900         CIE AUTOMOT.       21.3000       0.95       21.7000         ENAGAS       21.4700       0.28       21.6300         ENDESA       20.6700       0.05       20.9200         FERROVIAL       23.3100       0.56       23.6400         FLUIDRA       22.7800       1.24       23.2400         GRIFOLS CL.A       19.8800       -0.15       20.1800         IAG       1.4790       2.21       1.5015         IBERDROLA       11.1000       3.11       11.1800         INDITEX       20.4200       0.54       20.7100         INMACOLONIAL       7.6800       0.92       7.7150         MAPFRE       1.7730       -0.11       1.8090         MELIA HOTELS       6.9650       0.94<!--</th--><th>ARCELORMIT.         27.4000         1.03         28.0450         27.3300           B.SANTANDER         2.7440         0.13         2.7905         2.7320           BA.SABADELL         0.7314         -0.57         0.7504         0.7314           BANKINTER         5.5640         -0.68         5.6780         5.5560           BBVA         4.7475         -0.40         4.8285         4.7355           CAIXABANK         3.1200         -0.16         3.1840         3.1090           CELLNEX         40.4200         2.20         40.8900         39.4100           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400           ENAGAS         21.4700         0.28         21.6300         21.2900           ENDESA         20.6700         0.05         20.9200         20.5800           FERROVIAL         23.3100         0.56         23.6400         23.1100           FERROVIAL         19.8800         -0.15         20.1800         19.6800           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800           IBERDROLA         11.1000         3.11         11.1800         10.8600           INDITEX         20.420</th><th>ARCELORMIT.         27.4000         1.03         28.0450         27.3300         615.965           B.SANTANDER         2.7440         0.13         2.7905         2.7320         40.067.111           BA.SABADELL         0.7314         -0.57         0.7504         0.7314         22.843.011           BANKINTER         5.5640         -0.68         5.6780         5.5560         2.052.978           BBVA         4.7475         -0.40         4.8285         4.7355         24.941.682           CAIXABANK         3.1200         -0.16         3.1840         3.1090         18.086.447           CELLNEX         40.4200         2.20         40.8900         39.4100         934.479           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400         258.060           ENAGAS         21.4700         0.28         21.6300         21.2900         831.194           ENDESA         20.6700         0.05         20.9200         20.5800         902.924           FERROVIAL         23.3100         0.56         23.6400         23.1100         791.476           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800         779.006           IBER</th><th>ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 CIE AUTOMOT. 21.3000 0.95 21.7000 21.1400 258.060 5525.16 ENAGAS 21.4700 0.28 21.6300 21.2900 831.194 17823.64 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 FLUIDRA 22.7800 1.24 23.2400 22.5400 310.475 7128.03 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 IAG 1.4790 2.21 1.5015 1.4540 14.614.243 21631.29 IBERDROLA 11.1000 3.11 11.1800 10.8600 22.219.036 241268.15 INDITEX 20.4200 0.54 20.7100 20.0600 3.627.429 74407.17 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 INM.COLONIAL 7.6800 0.92 7.7150 7.6000 571.711 4387.52 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 MELIA HOTELS 6.9650 0.94 7.0700 6.9250 719.012 5023.26 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 NATURGY 27.6800 0.58 28.0200 27.5600 337.127 9346.64 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44</th><th>ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 20/05/2022 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 20/05/2022 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 20/05/2022 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 20/05/2022 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 20/05/2022 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 20/05/2022 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 20/05/2022 ENAGAS 21.4700 0.95 21.7000 21.1400 258.060 5525.16 20/05/2022 ENAGAS 21.4700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 INDRA 4 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHAR</th></th> | ARCELORMIT.       27.4000       1.03       28.0450         B.SANTANDER       2.7440       0.13       2.7905         BA.SABADELL       0.7314       -0.57       0.7504         BANKINTER       5.5640       -0.68       5.6780         BBVA       4.7475       -0.40       4.8285         CAIXABANK       3.1200       -0.16       3.1840         CELLNEX       40.4200       2.20       40.8900         CIE AUTOMOT.       21.3000       0.95       21.7000         ENAGAS       21.4700       0.28       21.6300         ENDESA       20.6700       0.05       20.9200         FERROVIAL       23.3100       0.56       23.6400         FLUIDRA       22.7800       1.24       23.2400         GRIFOLS CL.A       19.8800       -0.15       20.1800         IAG       1.4790       2.21       1.5015         IBERDROLA       11.1000       3.11       11.1800         INDITEX       20.4200       0.54       20.7100         INMACOLONIAL       7.6800       0.92       7.7150         MAPFRE       1.7730       -0.11       1.8090         MELIA HOTELS       6.9650       0.94 </th <th>ARCELORMIT.         27.4000         1.03         28.0450         27.3300           B.SANTANDER         2.7440         0.13         2.7905         2.7320           BA.SABADELL         0.7314         -0.57         0.7504         0.7314           BANKINTER         5.5640         -0.68         5.6780         5.5560           BBVA         4.7475         -0.40         4.8285         4.7355           CAIXABANK         3.1200         -0.16         3.1840         3.1090           CELLNEX         40.4200         2.20         40.8900         39.4100           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400           ENAGAS         21.4700         0.28         21.6300         21.2900           ENDESA         20.6700         0.05         20.9200         20.5800           FERROVIAL         23.3100         0.56         23.6400         23.1100           FERROVIAL         19.8800         -0.15         20.1800         19.6800           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800           IBERDROLA         11.1000         3.11         11.1800         10.8600           INDITEX         20.420</th> <th>ARCELORMIT.         27.4000         1.03         28.0450         27.3300         615.965           B.SANTANDER         2.7440         0.13         2.7905         2.7320         40.067.111           BA.SABADELL         0.7314         -0.57         0.7504         0.7314         22.843.011           BANKINTER         5.5640         -0.68         5.6780         5.5560         2.052.978           BBVA         4.7475         -0.40         4.8285         4.7355         24.941.682           CAIXABANK         3.1200         -0.16         3.1840         3.1090         18.086.447           CELLNEX         40.4200         2.20         40.8900         39.4100         934.479           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400         258.060           ENAGAS         21.4700         0.28         21.6300         21.2900         831.194           ENDESA         20.6700         0.05         20.9200         20.5800         902.924           FERROVIAL         23.3100         0.56         23.6400         23.1100         791.476           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800         779.006           IBER</th> <th>ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 CIE AUTOMOT. 21.3000 0.95 21.7000 21.1400 258.060 5525.16 ENAGAS 21.4700 0.28 21.6300 21.2900 831.194 17823.64 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 FLUIDRA 22.7800 1.24 23.2400 22.5400 310.475 7128.03 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 IAG 1.4790 2.21 1.5015 1.4540 14.614.243 21631.29 IBERDROLA 11.1000 3.11 11.1800 10.8600 22.219.036 241268.15 INDITEX 20.4200 0.54 20.7100 20.0600 3.627.429 74407.17 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 INM.COLONIAL 7.6800 0.92 7.7150 7.6000 571.711 4387.52 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 MELIA HOTELS 6.9650 0.94 7.0700 6.9250 719.012 5023.26 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 NATURGY 27.6800 0.58 28.0200 27.5600 337.127 9346.64 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44</th> <th>ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 20/05/2022 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 20/05/2022 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 20/05/2022 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 20/05/2022 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 20/05/2022 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 20/05/2022 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 20/05/2022 ENAGAS 21.4700 0.95 21.7000 21.1400 258.060 5525.16 20/05/2022 ENAGAS 21.4700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 INDRA 4 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHAR</th> | ARCELORMIT.         27.4000         1.03         28.0450         27.3300           B.SANTANDER         2.7440         0.13         2.7905         2.7320           BA.SABADELL         0.7314         -0.57         0.7504         0.7314           BANKINTER         5.5640         -0.68         5.6780         5.5560           BBVA         4.7475         -0.40         4.8285         4.7355           CAIXABANK         3.1200         -0.16         3.1840         3.1090           CELLNEX         40.4200         2.20         40.8900         39.4100           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400           ENAGAS         21.4700         0.28         21.6300         21.2900           ENDESA         20.6700         0.05         20.9200         20.5800           FERROVIAL         23.3100         0.56         23.6400         23.1100           FERROVIAL         19.8800         -0.15         20.1800         19.6800           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800           IBERDROLA         11.1000         3.11         11.1800         10.8600           INDITEX         20.420 | ARCELORMIT.         27.4000         1.03         28.0450         27.3300         615.965           B.SANTANDER         2.7440         0.13         2.7905         2.7320         40.067.111           BA.SABADELL         0.7314         -0.57         0.7504         0.7314         22.843.011           BANKINTER         5.5640         -0.68         5.6780         5.5560         2.052.978           BBVA         4.7475         -0.40         4.8285         4.7355         24.941.682           CAIXABANK         3.1200         -0.16         3.1840         3.1090         18.086.447           CELLNEX         40.4200         2.20         40.8900         39.4100         934.479           CIE AUTOMOT.         21.3000         0.95         21.7000         21.1400         258.060           ENAGAS         21.4700         0.28         21.6300         21.2900         831.194           ENDESA         20.6700         0.05         20.9200         20.5800         902.924           FERROVIAL         23.3100         0.56         23.6400         23.1100         791.476           GRIFOLS CL.A         19.8800         -0.15         20.1800         19.6800         779.006           IBER | ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 CIE AUTOMOT. 21.3000 0.95 21.7000 21.1400 258.060 5525.16 ENAGAS 21.4700 0.28 21.6300 21.2900 831.194 17823.64 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 FLUIDRA 22.7800 1.24 23.2400 22.5400 310.475 7128.03 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 IAG 1.4790 2.21 1.5015 1.4540 14.614.243 21631.29 IBERDROLA 11.1000 3.11 11.1800 10.8600 22.219.036 241268.15 INDITEX 20.4200 0.54 20.7100 20.0600 3.627.429 74407.17 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 INM.COLONIAL 7.6800 0.92 7.7150 7.6000 571.711 4387.52 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 MELIA HOTELS 6.9650 0.94 7.0700 6.9250 719.012 5023.26 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 NATURGY 27.6800 0.58 28.0200 27.5600 337.127 9346.64 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 | ARCELORMIT. 27.4000 1.03 28.0450 27.3300 615.965 17103.20 20/05/2022 B.SANTANDER 2.7440 0.13 2.7905 2.7320 40.067.111 110999.29 20/05/2022 BA.SABADELL 0.7314 -0.57 0.7504 0.7314 22.843.011 16918.40 20/05/2022 BANKINTER 5.5640 -0.68 5.6780 5.5560 2.052.978 11519.40 20/05/2022 BBVA 4.7475 -0.40 4.8285 4.7355 24.941.682 119088.61 20/05/2022 CAIXABANK 3.1200 -0.16 3.1840 3.1090 18.086.447 56753.73 20/05/2022 CELLNEX 40.4200 2.20 40.8900 39.4100 934.479 37760.74 20/05/2022 ENAGAS 21.4700 0.95 21.7000 21.1400 258.060 5525.16 20/05/2022 ENAGAS 21.4700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 ENDESA 20.6700 0.05 20.9200 20.5800 902.924 18693.33 20/05/2022 FERROVIAL 23.3100 0.56 23.6400 23.1100 791.476 18479.13 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 GRIFOLS CL.A 19.8800 -0.15 20.1800 19.6800 779.006 15476.23 20/05/2022 INDRA 4 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 20.0600 3.627.429 74407.17 20/05/2022 INDRA A 9.6150 1.10 9.7700 9.4950 629.913 6081.13 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MAPFRE 1.7730 -0.11 1.8090 1.7730 4.653.735 8316.98 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 MERLIN 10.3400 0.98 10.5000 10.1500 841.403 8740.20 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHARMA MAR 70.1000 1.18 70.8000 69.2400 48.329 3380.44 20/05/2022 PHAR |

```
Últ. % Dif.
                                                            Volumen Efectivo (miles €)
                    Nombre
                                             Máx.
                                                      Mín.
                                                                                             Hora
                                                                                      Fecha
           31
                       ROVI
                                                             131.909
                             59.9000
                                     4.08
                                           60.6000
                                                   58.1500
                                                                            7877.93 20/05/2022 Cierre
           32 SIEMENS GAME
                                           16.9300
                             16.7450
                                     1.12
                                                   16.4550
                                                           2.112.904
                                                                           35052.29 20/05/2022 Cierre
In [35]: ibex df03.info()
          <class 'pandas.core.frame.DataFrame'>
          RangeIndex: 35 entries, 0 to 34
          Data columns (total 9 columns):
                Column
                                      Non-Null Count
                                                       Dtype
                Nombre
                                      35 non-null
                                                        object
               Últ.
                                                        float64
                                      35 non-null
                                      35 non-null
                                                        float64
               % Dif.
           3
               Máx.
                                      35 non-null
                                                        float64
                                                        float64
           4
                Mín.
                                      35 non-null
                                                        object
                Volumen
                                      35 non-null
                                                        float64
               Efectivo (miles €) 35 non-null
                Fecha
                                      35 non-null
                                                        object
                                                        object
                Hora
                                      35 non-null
          dtypes: float64(5), object(4)
          memory usage: 2.6+ KB
```

Pugem el csv a Kaggle (és de google, aplicatiu web que no cal instal·lar, i permet visualitzar dades o fer-ho servir de notebook). Ho documentem en un doc de libroffice, tal i com es demana.

## Exercici 3. Tria una página web que tu vulguis i realitza web scraping mitjançant la llibreria Scrapy.

```
In [36]: url="https://cadenaser.com/"
  page02 = requests.get(url).content

In [37]: sel = Selector(text = page02)
```

Busquem la class que conté enllaços de notícies, i en concret les de "Deportes". A continuació, un extracte:

<div class="mnu" id="mnu"><div class="c-mnu"><a href="/actualidad/" title="Actualidad">Actualidad</a>
a><a href="/actualidad/" title="Actualidad">Actualidad</a>
<a href="/actualidad/politica/" title="Pol</a>

id="/deportes/futbol"><button>Volver</button><a href="/deportes/futbol/" title="F\xc3\xbatbol">
F\xc3\xbatbol</a><a href="/deportes/futbol/liga-santander/" title="Liga Santander">Liga
Santander</a><a href="/deportes/futbol/liga-smartbank/" title="Liga Smartbank">Liga Smartbank</a><a href="/deportes/futbol/champions-league/" title="Champions League">Champions League">Champions League</a><a href="/deportes/futbol/europa-league/" title="Europa League">Europa League</a><a href="/deportes/futbol/europa-league/" title="Motociclismo">Motociclismo</a><a href="/deportes/formula-1/" title="F\xc3\xb3rmula 1">F\xc3\xb3rmula 1</a><a href="/deportes/formula-1/" title="F\xc3\xb3rmula 1">F\xc3\xb3rmula 1</a><a href="/deportes/tenis/" title="T enis">Tenis</a><a href="/deportes/mas-deportes/" title="M\xc3\xa1s deportes">M\xc3\xa1s deportes">M\xc3\xa1s deportes</a>

```
In [40]: # seleccionem la divisió dels continguts principals
links = sel.xpath ('//div[@class="mnu"]//a/@href').extract()
links
```

Out[40]:

```
['/actualidad/',
           '/actualidad/',
           '/actualidad/politica/',
           '/actualidad/tribunales/',
           '/actualidad/internacional/',
           '/actualidad/sociedad/',
           '/actualidad/economia-y-negocios/',
           '/actualidad/ciencia-y-tecnologia/',
           '/deportes/',
           '/deportes/',
           '/deportes/futbol/',
           '/deportes/futbol/liga-santander/',
           '/deportes/futbol/liga-smartbank/',
           '/deportes/futbol/champions-league/',
           '/deportes/futbol/europa-league/',
           '/deportes/motociclismo/',
           '/deportes/formula-1/',
           '/deportes/baloncesto/',
           '/deportes/tenis/',
           '/deportes/mas-deportes/',
           '/humor/',
           '/ocio-y-cultura/',
           '/ocio-y-cultura/',
           '/ocio-y-cultura/ocio/',
           '/ocio-y-cultura/cultura-y-sociedad/',
           '/ocio-y-cultura/gastro/',
           '/ocio-y-cultura/gente/',
           '/ocio-y-cultura/cine-y-tv/',
           '/ocio-y-cultura/historia/',
           '/ocio-y-cultura/musica/',
           '/ocio-y-cultura/salud-y-bienestar/',
           '/opinion/',
           '/ccaa/andalucia/',
In [42]: outcome = []
         for link in links:
             #busquem aquells links on apareixi el string "deportes"
             if 'deportes/futbol/' in link:
                  outcome.append(link)
                  break
```

```
In [43]: outcome
 Out[43]: ['/deportes/futbol/']
 In [44]: # anem a la secció "deportes"
          page03 = requests.get(url+outcome[0]).content
 In [77]: # parsegem el contingut amb la llibreria Selector de scrapy
          sel02 = Selector(text = page03)
 In [78]: sel02
 Out[78]: <Selector xpath=None data='<html lang="es"><head><meta charset="...'>
          Fem un extracte de la informació continguda:
              <div class="ele-4"
              <div class="s-s">
                                              <article class="sc t-s">
                                                  <figure>
                                                          href="/nacional/2022/05/21/la-segunda-division-se-resue
              lve-los-resultados-en-directo-cadena-ser/"
                                                          title="El Almer\xc3\xada no logra el ascenso directo tr
              as empatar contra el Alcorc\xc3\xb3n en la jornada de la salvaci\xc3\xb3n para M\xc3\xa1laga y Spor
              ting"
In [117]: # volem trobar tots els tags "articles" i el seu text associat
          news headlines = sel02.xpath('//article/div/h2/a/text()').extract()
          # volem trobar tots els tags "articles" i el seu text associat
          #news headlines = sel.xpath ('//article/div/h2/a/text()').extract()
```

```
In [118]: news headlines[:20]
Out[118]: []
 In [98]: sel02.xpath("//div").extract()
 Out[98]: ['<div id="fusion-app"><script type="application/ld+json">[{"@context":"https://schema.org","@type":"ItemL
          ist", "url": "https://cadenaser.com", "itemListElement": [{ "@type": "ListItem", "position": 1, "url": "https://cade
          naser.com/nacional/2022/05/22/la-lucha-por-el-descenso-en-directo-laliga-se-resuelve-en-carrusel-deportivo
          -cadena-ser/"},{"@type":"ListItem","position":2,"url":"https://cadenaser.com/nacional/2022/05/21/ultima-ho
          ra-de-mbappe-y-el-real-madrid-en-directo-cadena-ser/"},{"@type":"ListItem","position":3,"url":"https://cad
          enaser.com/nacional/2022/05/22/me-parece-el-mayor-gatillazo-de-la-historia-del-futbol-cadena-ser/"},{"@typ
          e":"ListItem", "position":4, "url": "https://cadenaser.com/nacional/2022/05/22/ander-herrera-explica-en-carru
          sel-como-se-entero-el-vestuario-del-psq-de-la-renovacion-de-mbappe-cadena-ser/"},{"@type":"ListItem","posi
          tion":5, "url": "https://cadenaser.com/nacional/2022/05/22/manu-carreno-ni-el-real-madrid-ni-ningun-club-de-
          futbol-puede-competir-contra-todo-un-estado-como-catar-cadena-ser/"},{"@type":"ListItem","position":6,"ur
          l": "https://cadenaser.com/nacional/2022/05/21/la-segunda-division-se-resuelve-los-resultados-en-directo-ca
          dena-ser/"},{"@type":"ListItem","position":7,"url":"https://cadenaser.com/nacional/2022/05/21/kylian-mbapp
          e-anuncia-su-renovacion-con-el-psg-hasta-2025-cadena-ser/"},{"@type":"ListItem","position":8,"url":"http
          s://cadenaser.com/nacional/2022/05/21/asi-fue-el-show-de-la-renovacion-de-mbappe-en-el-parque-de-los-princ
          ipes-cadena-ser/"},{"@type":"ListItem","position":9,"url":"https://cadenaser.com/nacional/2022/05/21/lalig
          a-denunciara-al-psg-ante-la-uefa-tras-la-renovacion-de-mbappe-cadena-ser/"},{"@type":"ListItem","positio
          n":10, "url": "https://cadenaser.com/nacional/2022/05/21/barcelona-olympique-de-lyon-la-final-de-la-champion
          s-league-en-directo-cadena-ser/"},{"@type":"ListItem","position":11,"url":"https://cadenaser.com/nacional/
          2022/05/21/tebas-lo-que-va-a-hacer-el-psg-renovando-a-mbappe-es-un-insulto-al-futbol-cadena-ser/"},{"@typ
In [104]: #extraiem informació de div class="ele-4", que pertany a Deportes
          sel03=sel02.xpath("//div[@class='ele-4']").extract()
          print(sel03)
```

```
In [113]: #If we want to get html node
sel02.xpath("/html").extract()
```

Out[113]: ['<html lang="es"><head><meta charset="utf-8"><meta http-equiv="X-UA-Compatible" content="ie=edge"><meta h</pre> ttp-equiv="Content-Type" content="text/html; charset=utf-8"><meta name="viewport" content="width=device-wi dth, initial-scale=1"><meta name="referrer" content="no-referrer-when-downgrade"><meta name="lang" content ="es-ES"><meta name="robots" content="index, follow, noarchive"><meta name="robots" content="max-video-pre view:3"><meta name="robots" content="max-image-preview:large"><meta name="format-detection" content="addre</pre> ss=no;email=no,telephone:no"><meta name="robots" content="max-snippet:-1"><meta name="organization" content t="Cadena SER."><meta property="og:image" content="https://cadenaser.com/pf/resources/dist/img/ser-cl1.sv q?d=217"><meta property="og:image:width" content="300"><meta property="og:image:height" content="300"><met a property="og:type" content="website"><meta property="og:site name" content="cadena SER"><meta property=" fb:app id" content="229574677178613"><meta property="fb:pages" content="15658775846,32974073100,1411837632 66,297798423468,260097275665,154038834626728,126239340759913,152716074759665,286241832727,108754747127102 1,397748630349260,260807093963798,218355098183367,192630137538936,209200415772655,9060610628506,1594098194 2404,130512786996530,1930066643884232,767055920005326,965792190174136,1414168792217838,124134164273538,107 895409321804, 168883333672674, 503744749685810, 263858326757, 258738666530, 319811467873, 222909084390267, 943306 019036011,534829370024405,175022845914285,597264423810487,105594696642793,426479974076402,417737195051925, 1586904454928885, 161976080524709, 202089616485001, 391992824157969, 275214532511758, 119523764775530, 141266639 267590, 198050246911, 278472948668, 394007950541, 1571792933045547, 157613800916527, 586964561374374, 8724060227 2,1766650093650479,1884462045171442,192810794093275,154946021197794,154185008018557,345200755667666,838686 046216716,178334868928834,230508573734771 ,188156367978437,337959822922824,839188389472025 ,18685351468776

```
In [114]: #If you want to get all div descendant of this html
sel02.xpath("/html//div").extract()
```

Out[114]:

['<div id="fusion-app"><script type="application/ld+json">[{"@context": "https://schema.org", "@type": "ItemList", "url": "https://cadenaser.com", "itemListElement": [{"@type": "ListItem", "position": 1, "url": "https://cadenaser.com/nacional/2022/05/22/la-lucha-por-el-descenso-en-directo-laliga-se-resuelve-en-carrusel-deportivo-cadena-ser/"}, {"@type": "ListItem", "position": 2, "url": "https://cadenaser.com/nacional/2022/05/21/ultima-hora-de-mbappe-y-el-real-madrid-en-directo-cadena-ser/"}, {"@type": "ListItem", "position": 3, "url": "https://cadenaser.com/nacional/2022/05/22/me-parece-el-mayor-gatillazo-de-la-historia-del-futbol-cadena-ser/"}, {"@type": "ListItem", "position": 4, "url": "https://cadenaser.com/nacional/2022/05/22/ander-herrera-explica-en-carrusel-como-se-entero-el-vestuario-del-psg-de-la-renovacion-de-mbappe-cadena-ser/"}, {"@type": "ListItem", "position": 5, "url": "https://cadenaser.com/nacional/2022/05/22/manu-carreno-ni-el-real-madrid-ni-ningun-club-de-futbol-puede-competir-contra-todo-un-estado-como-catar-cadena-ser/"}, {"@type": "ListItem", "position": 6, "url": "https://cadenaser.com/nacional/2022/05/21/la-segunda-division-se-resuelve-los-resultados-en-directo-cadena-ser/"}, {"@type": "ListItem", "position": 6, "url": "https://cadenaser.com/nacional/2022/05/21/kylian-mbappe-anuncia-su-renovacion-con-el-psg-hasta-2025-cadena-ser/"}, {"@type": "ListItem", "position": 8, "url": "https://cadenaser.com/nacional/2022/05/21/kylian-mbappe-en-el-psg-hasta-2025-cadena-ser/"}, {"@type": "ListItem", "position": 8, "url": "https://cadenaser.com/nacional/2022/05/21/asi-fue-el-show-de-la-renovacion-de-mbappe-en-el-parque-de-los-princ

In [ ]: