

```
In [54]: import requests
          from bs4 import BeautifulSoup

          r = requests.get("http://www.oabrij.org.br/")
          c = r.content

          c
```

[illegible]

```
In [55]: soup=BeautifulSoup(c,"html.parser")
```

```
In [66]: all=soup.find_all("div",{"class":""})
```

In [67]: all

```
<a href="/visitas-as-sessoes-2017/visitas-as-sessoes-2017/47">
</li>
</ul>
</li>
<li class="menu-horizontal-item-tem-submenus">
<a href="/sn/paginas/58/Calend%C3%A1rio%20de%20sess%C3%B5es.html">Tribunal de Ética</a>
<ul class="menu-horizontal-submenu">
<li>
<a href="/apresentacao-ted">Apresentação (TED)</a>
</li>
<li>
<a href="/calendario-de-sessoes">Calendário de sessões</a>
</li>
<li>
<a href="/composicao">Composição</a>
</li>
<li>
<a href="/regimento-interno">Regimento interno</a>
</li>
<li>
```

In [69]: `import pandas as pd`

```
df= pd.DataFrame(all)
df.to_dict()
```

df

Out[69]:

0

```
0    <div id="wrapper"> <div id="header"> <div id="...
1    <div id="header"> <div id="header-barra-extra"...
2    <div id="header-barra-extra"> <div id="header-...
3    <div id="header-pesquisa"> <form action="/busc...
4    <div> <input name="expressao" type="text" valu...
5    <div id="header-links"> <a class="header-barra...
6        <div id="logo"> <a href="/" title="Ir para a p...
7    <div id="menu-secoes-especiais"> <div id="menu...
8    <div id="menu-secoes-especiais-inner"> <ul> <l...
9    <div id="menu-horizontal"> <ul class="menu-hor...
10   <div id="container"> <div id="banner-topo"> <a...
11   <div id="banner-topo"> <a href="http://posgrad...
12   <div id="lateral-esquerda"> <div style="margin...
13   <div style="margin: 0 0 12px 0"> <a href="http...
14   <div id="webmail" style="width:141px;height:35...
15   <div id="destaques"> <div class="carousel"> <d...
16   <div id="tv-oab"> <div class="chapeu"> <div cl...
17   <div id="busca-de-inscritos"> <div class="chap...
18   <div id="lista-principal"> <div class="chapeu"...
19   <div id="destaques-laterais"> <ul class="notic...
```

0

```

20 <div id="redes-sociais"> <div class="rede-soci...
21 <div id="rodape-banners"> </div>
22 <div id="rodape"> <div>OAB/RJ - Av. Marechal C...
23 <div>OAB/RJ - Av. Marechal Câmara, 150 - Rio d...
24 <div><a href="mailto:atendimento@oabrj.org.br"...
25 <div id="fb-root"></div>
26 <div id="modal"></div>
27 <div id="popupContact"> <a href="javascript:vo...
28 <div id="retornoAjax"></div>
29 <div id="backgroundPopup"></div>

```

```
In [70]: df.to_csv('all.csv', index=False)
print('Done')
```

Done

```
In [73]: df.head(5)
```

Out[73]:

0

```

0 <div id="wrapper"> <div id="header"> <div id="...
1 <div id="header"> <div id="header-barra-extra"...
2 <div id="header-barra-extra"> <div id="header-...
3 <div id="header-pesquisa"> <form action="/busc...
4 <div> <input name="expressao" type="text" valu...

```

```
In [74]: import psycopg2

conn = psycopg2.connect("host=localhost dbname=diario_oficial user=postgres")
cur = conn.cursor()
cur.execute('''CREATE TABLE dados (
                id INT,
                name TEXT
            )''')

conn.commit()
conn.close()
```

```
In [76]: import psycopg2

conn = psycopg2.connect("host=localhost dbname=diario_oficial user=postgres")
cur = conn.cursor()
with open('all.csv', 'r') as f:
    next(f)
    cur.copy_from(f, 'dados', sep=',')

conn.commit()
```

```
-----
DataError                                Traceback (most recent call last)
<ipython-input-76-5c06ab10b3b8> in <module>()
      5 with open('all.csv', 'r') as f:
      6     next(f)
----> 7     cur.copy_from(f, 'dados', sep=',')[1]
      8
      9 conn.commit()
```

```
DataError: invalid input syntax for integer: ""<div id=""wrapper"">"
CONTEXT:  COPY dados, line 1, column id: ""<div id=""wrapper"">"
```

```
In [ ]:
```

```
In [ ]:
```

