

# Avances en Web Scraping para el Portal de Estudios de la Superintendencia de Bancos

Pedro Galindo Vera

2023-05-01

## Caso: Boletín Financiero Mensual

### Identificación del contenedor de descargas

Empleando el Inspector de código que incorpora Google Chrome identificamos rápidamente el código correspondiente al contenedor de archivos que se desea “raspar”:

```
<div class="elementor-element elementor-element-3c1c97b elementor-widget elementor-widget-wpcp-shareonee
    <div class="elementor-widget-container">
        <div id="ShareoneDrive" class=" files wpcp-has-scroll-to-top" style="display:none"><noscrip
        <i class="eva eva-home-outline"></i>
    </a>
    <div class="nav-title"><ol class="wpcp-breadcrumb wpcp wpcp-left"><li class="dropdown wpcp-drop
    <a href="javascript:void(0);" class="wpcp-toggle" data-toggle="dropdown">
        <i class="eva eva-arrow-down"></i>
    </a>
    <ul class="dropdown-menu tippy-content"><li class="is-hidden"><a href="#01ONK3WL5YZDTIKNK32ND3L2L
    </li><li class="wpcp-ellipsis is-hidden" style="">...</li><li class="first-breadcrumb"><a href="#01
        <a class="nav-search entry-info-button" tabindex="0" aria-expanded="false">
            <i class="eva eva-search"></i>
        </a>

    <a class="nav-sort entry-info-button" title="Opciones de orden" aria-expanded="false">
        <i class="eva eva-options"></i>
    </a>

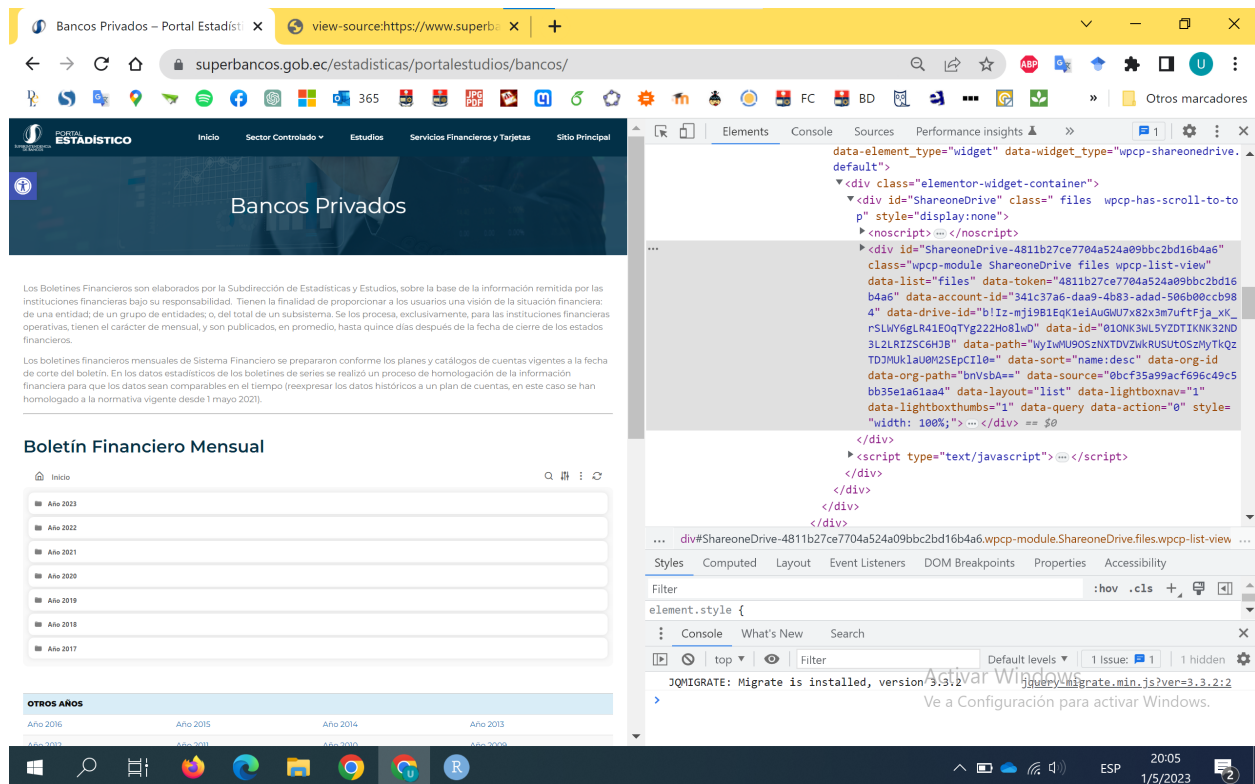
    <a class="nav-gear entry-info-button" title="Más acciones" aria-expanded="false">
        <i class="eva eva-more-vertical-outline"></i>
    </a>

        <a class="nav-refresh entry-info-button" title="Actualizar">
            <i class="eva eva-sync"></i>
        </a>
    </div>    <div class="file-container">
    <div class="loading" style="display: none;">                                <svg class="loader-spinner" viewBox="25
```

```

        <circle class="path" cx="50" cy="50" r="20" fill="none" stroke-width="3" stroke-miterlin
    </svg>
    </div>
    <div class="ajax-filelist" style=""><div class="files has-breadcrumb "><div class="folders-conta
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder" data-id="01ONK3WLY7WPUKDMR4AJCKJFNUDUIB3UXG" data-drive-id="b!Iz-mji9B1EqK1ei
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder" data-id="01ONK3WL7YZ6JTTZVXNFE2YNQ6NNUXUKQQ" data-drive-id="b!Iz-mji9B1EqK1ei
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder" data-id="01ONK3WL3Z7K26CYSM6JC2YKVXY2EREYBG" data-drive-id="b!Iz-mji9B1EqK1ei
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder " data-id="01ONK3WLZVOMAOMXU535AYNT5GPIRCYRWM" data-drive-id="b!Iz-mji9B1EqK
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder " data-id="01ONK3WL46YXOHX6G6LNALFZ3YMUDZBJB2" data-drive-id="b!Iz-mji9B1EqK
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
<div class="entry folder " data-id="01ONK3WL47BDVYGX47KZAKKGAEMN2UUE65" data-drive-id="b!Iz-mji9B1EqK
<div class="entry_block">
<div class="entry-info"><div class="entry-info-icon"><div class="description-textbox"><div class="description-file-name">A
</div>
</div>
</div><div class="files-container"></div></div></div>
    <div class="scroll-to-top scroll-to-top-bottom" style="display: none; top: 450.237px; right: 0px;
        <button class="scroll-to-top-action button button-round-icon secondary button-round-icon-lg
            </div>
    </div>
</div></div></div><script type="text/javascript">if (typeof(jQuery) !== 'undefined' && typeof(jQuery.cp
    </div>

```



El identificador del contenedor `div` refleja el módulo `ShareoneDrive`.

Expandiendo los contenedores se llega al nodo `a href` donde esperaríamos encontrar los enlaces de descarga, pero se encuentra un método que no permite obtener directamente los enlaces deseados, y que depende de la interacción con los botones del contenedor de archivos; por ejemplo al botón `Año 2023` corresponde el código:

```
<a href="javascript:void(0);" class="entry_link" title="Año 2023"><span>Año 2023 </span></a>
```

estadísticos de los boletines de series se realizó un proceso de homologación de la información financiera para que los datos sean comparables en el tiempo (reexpresar los datos históricos a un plan de cuentas, en este caso se han homologado a la normativa vigente desde 1 mayo 2021).

## Boletín Financiero Mensual

**a.entry\_link** 53.95 × 15.77

Año 2023

Año 2022

Año 2021

Año 2020

Año 2019

Año 2018

Año 2017

### OTROS AÑOS

Año 2016	Año 2015	Año 2014	Año 2013
Año 2012	Año 2011	Año 2010	Año 2009
Año 2008	Año 2007	Año 2006	Año 2005

El código con los enlaces solo se revela una vez pulsado el botón Año 2023:

```

<div class="files-container"><div class="entry file" data-id="010NK3WLZUPWHEY4AOFBJXCRTFRH2SKAV" data-
<div class="entry_block">
<div class="entry_thumbnail"><div class="entry_thumbnail-view-bottom"><div class="entry_thumbnail-view-
<div class="preloading"></div><div class="entry-info-icon "><div class="entry-info-button entry-description-button " tabindex="0" aria-expanded="false"><i cl
<div class="tippy-content-holder"><div class="description-textbox"><div class="description-file-name">B
</div>
</div>
</div>
<div class="entry file" data-id="010NK3WL7MBSJ5UW24HZAKHTEGPJTY2FND" data-name="BOLETIN BANCOS FEBRERO
<div class="entry_block">
<div class="entry_thumbnail"><div class="entry_thumbnail-view-bottom"><div class="entry_thumbnail-view-
<div class="preloading"></div><div class="entry-info-icon "><div class="entry-info-button entry-description-button " tabindex="0" aria-expanded="false"><i cl
<div class="tippy-content-holder"><div class="description-textbox"><div class="description-file-name">B
</div>
</div>
</div>
<div class="entry file" data-id="010NK3WL5KUGT56EHY6BAJVTMZU6SJVUS4" data-name="BOLETIN BANCOS ENERO 20
<div class="entry_block">
<div class="entry_thumbnail"><div class="entry_thumbnail-view-bottom"><div class="entry_thumbnail-view-
<div class="preloading"></div><div class="entry-info-icon "><div class="entry-info-button entry-description-button" tabindex="0" aria-expanded="false"><i cl
<div class="tippy-content-holder"><div class="description-textbox"><div class="description-file-name">B
</div>
</div>
</div>
</div>

```

Bancos Privados – Portal Estadístico | view-source:https://www.superbancos.gob.ec/estadisticas/portales/

superbancos.gob.ec/estadisticas/portales/estudios/bancos/

estadísticos de los boletines de series se realizó un proceso de homologación de la información financiera para que los datos sean comparables en el tiempo (reexpresar los datos históricos a un plan de cuentas, en este caso se han homologado a la normativa vigente desde 1 mayo 2021).

## Boletín Financiero Mensual

Inicio > Año 2023

div.files-container 708.61 × 129.99

BOLETIN BANCOS	FECHA	TAMANO
BOLETIN BANCOS MARZO 2023.zip	6 abril	636 KB
BOLETIN BANCOS FEBRERO 2023.zip	6 marzo	737 KB
BOLETIN BANCOS ENERO 2023.zip	13 febrero	762 KB

OTROS AÑOS

Año 2016	Año 2015	Año 2014	Año 2013
Año 2012	Año 2011	Año 2010	Año 2009
Año 2008	Año 2007	Año 2006	Año 2005
Año 2004	Año 2003	Año 2002	Año 2001
Año 2000	Año 1999	Año 1998	Año 1997

Elements Console Sources Performance insights

```

<div class="nav-header nav-header-has-breadcrumb">
</div>
<div class="file-container">
<div class="loading" style="display: none;"></div>
<div class="ajax-filelist" style="display: none;">
<div class="files has-breadcrumb">
<div class="folders-container"></div>
<div class="files-container">
<div class="entry file" data-id="010NK3MLZUPWHT
EY4AOFB3KCRTRFH2SKAV" data-name="BOLETIN BANCOS
MARZO 2023"></div>
<div class="entry file" data-id="010NK3ML7MBSJ5
UN24HZAKHTEGP3TY2FND" data-name="BOLETIN BANCOS
FEBRERO 2023"></div>
<div class="entry file" data-id="010NK3ML5KUGT5
6EHY6BAJVTMZUG65JVUS4" data-name="BOLETIN BANCOS
ENERO 2023"></div>
</div>
</div>
<div class="scroll-to-top scroll-to-top-bottom"
style="display: none; top: 450.237px; right: 0px;">
</div>
</div>

```

view div.list-container div.file-container div.ajax-filelist div.files.has-breadcrumb div.files-container

Styles Computed Layout Event Listeners DOM Breakpoints Properties Accessibility

Filter :hov .cls

Console What's New Search

top Filter

Default levels 1 Issue: 1 1 hidden

JQMIGRATE: Migrate is installed, version 3.3.2 jquery/migrate.min.js?ver=3.3.2:2

Activar Windows. Ve a Configuración para activar Windows.

20:22 1/5/2023