Tarea 3: Web scraping

Enunciado

Un web scraper es un método de extracción de datos de un sitio web que comúnmente funciona de forma automática. Como los scrapers no pueden ver las páginas webs como un humano, hacen búsquedas de patrones en el código HTML de la página. Una de las formas en que busca los patrones es con Expresiones regulares[2]

Por ejemplo, si tenemos un scraper que obtiene el valor de dólar diarament de algún sitio. El scraper obtendrá el valor del dólar de un día en particular creando una expresión regular que encaje con el código HTML que muestra el valor del dolar.

```
<span class="indicator-title">
  <h1>663,23</h1>
  Dólar
  <small>$</small>
</span>
```

En el caso del código anterior una expresión regular que cumpla con el objetivo sería:

```
<span class="indicator-title">\n\s+<h1>([0-9,]+)</h1>\n\s+Dólar
```

Su tarea corresponde a hacer implementar scraper que se conecte al sitio Box Office Mojo[1] que es una página web que hace un seguimiento de los ingresos de las películas.
En el deberán obtener las siguientes características asociadas a un artista:

- Ingreso bruto total en su carrera. (Lifetime Gross Total).
- El ingreso total ajustado a la inflación (Adjusted Total).
- Las últimas 5 películas del artista.
- Nombre de las 10 películas de mayor ingreso bruto.
- Nombre de las 10 películas de menor ingreso bruto.

Recomendaciones

- Se recomienda utilizar el sitio regex101 para probar sus expresiones regulares.
- Para encontrar la página asociada del artista puede usar el buscador interno de la página `https://www.boxofficemojo.com/search/?q=robert%20downey,%20jr` o ingresarlo acceder directamente a su página `https://www.boxofficemojo.com/people/chart/?view=Actor&id=robertdowneyjr.htm`

Condiciones de entrega

- La tarea es individual y debe ser entregada antes del Viernes 28 de Junio a las 23:59 hrs.
- El informe y el código fuente debe ser enviado en un archivo zip con el nombre <NombreApellido>.zip

Referencias

[1] Box office mojo. URL: `https://www.boxofficemojo.com`.

[2] Wikipedia. Expresión regular — wikipedia, la enciclopedia libre, 2019. URL: `https://es.wikipedia.org/w/index.php?title=Expresi%C3%B3n_regular`.