

Web Scraping Steam

Objetivos	2
Manual de uso / descripción de la interfaz	3
Descripción del proyecto	6
• Módulos creados	6
Clases:	6
Librerías:	7
WebScraping en Metacritic	8

Objetivos

El objetivo es desarrollar un software que realice un web scraping por la página de Steam.

OBJ-01	Mostrar las categorías
Descripción	El sistema <i>nos deberá mostrar las categorías que dicha web dispone para nosotros.</i>

OBJ-02	Mostrar los 5 juegos más valorados
Descripción	El sistema <i>nos mostrará los 5 juegos mejor valorados por los usuarios de dicha categoría</i>

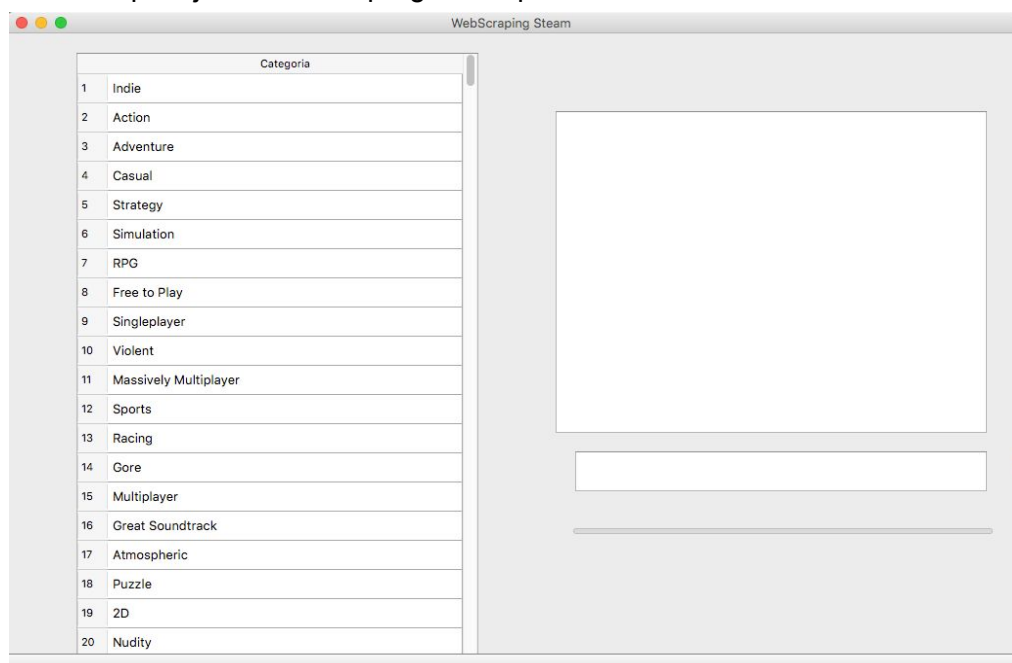
OBJ-03	Generar Excel
Descripción	El sistema <i>nos generará un excel con los 50 juegos mejor valorados por los usuarios.</i>
Comentario	Los juegos dispondrán de la siguiente información: <ul style="list-style-type: none"> A. Título B. Precio (original y con descuento, si lo hubiese) C. Breve Descripción D. URL del sitio web del juego E. Plataformas F. Idiomas: español (sí/no), inglés (sí/no)

	<p>G. Valoración de la comunidad:</p> <p>a. Análisis recientes: valoración (0-100) sobre el número de análisis de los últimos 30 días.</p> <p>b. Análisis totales: valoración (0-100) sobre el número total de análisis</p> <p>H. URL</p>
--	---

OBJ-04	Generar Excel2
Descripción	El sistema <i>nos generará un excel2 que contendrá la variación de precios de tus juegos favoritos.</i>
Comentario	Prerrequisito: <i>En el Excel1 debemos crear manualmente una segunda hoja, en ella insertar una tabla con dos columnas, una con los nombre de nuestros juegos favoritos, y otra con precio de los mismos.</i>

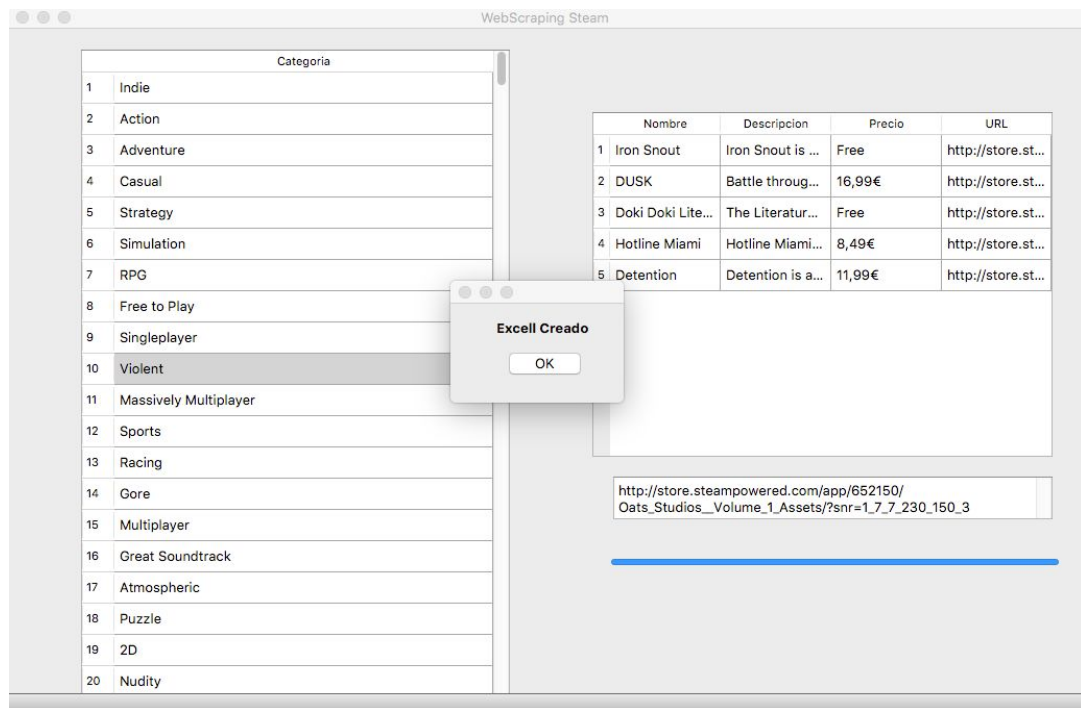
Manual de uso / descripción de la interfaz

Una vez que ejecutemos el programa aparecerá una ventana con dos tablas:

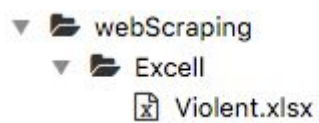


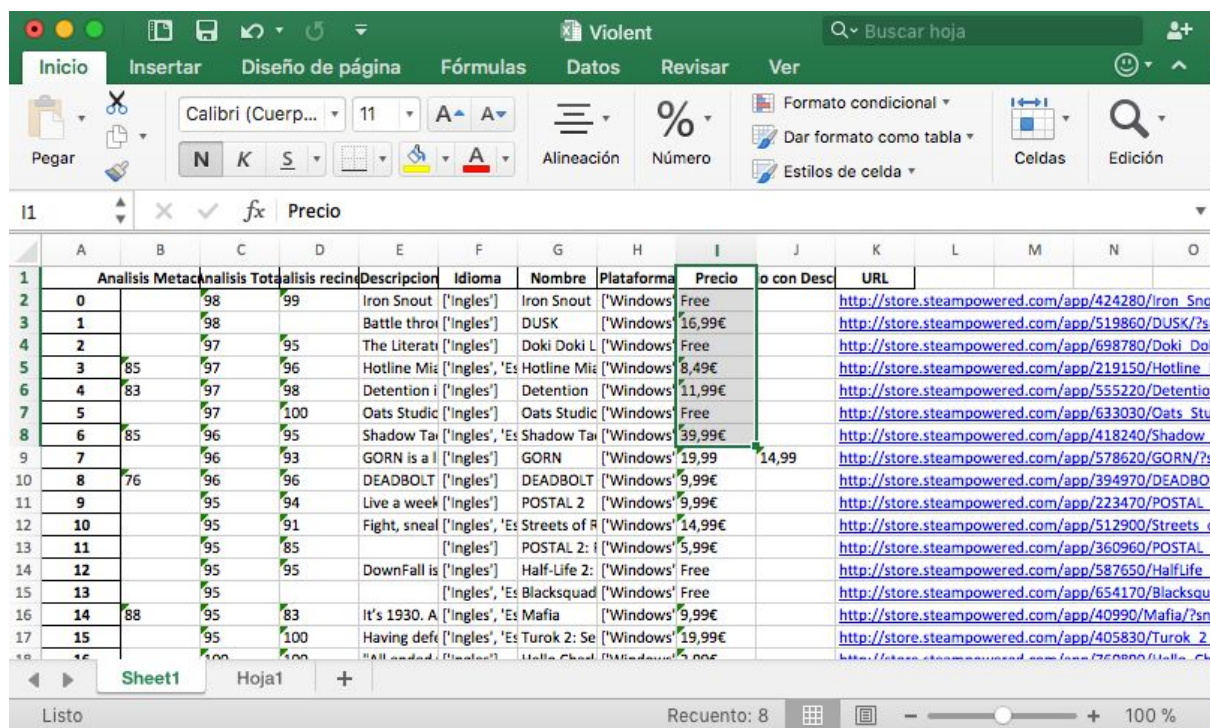
La tabla situada en el lado izquierdo corresponde a las categorías, en dicha tabla podemos observar las 379 categorías que steam pone a nuestra disposición. Haciendo click en una categoría haremos poner en marcha todo el motor de web scraping.

Tras un breve periodo de tiempo se nos mostrará el mensaje “Excel generado Correctamente”, indicando que el programa se ha ejecutado correctamente. Entonces podremos visualizar los cinco juegos mejor valorados en la tabla del lado derecho.



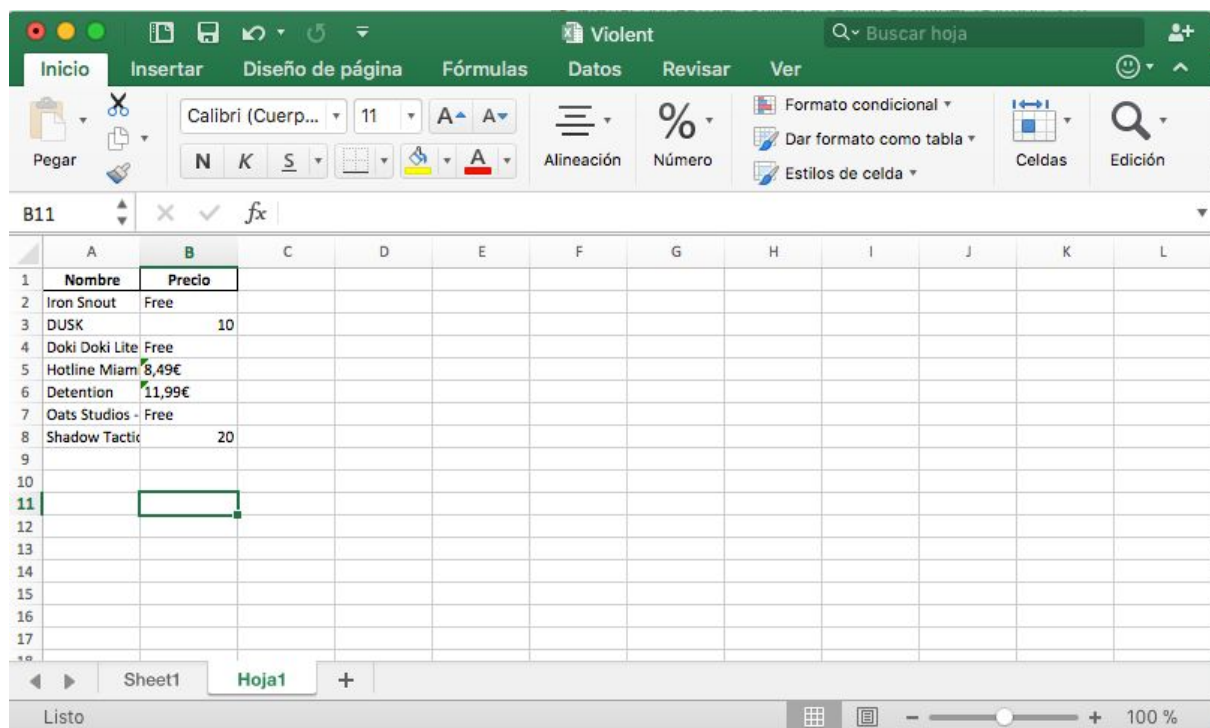
Además, dispondremos de un excel con el mismo nombre que la categoría elegida:





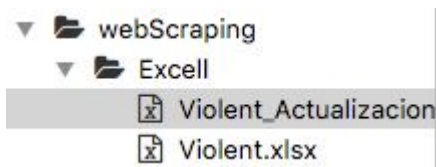
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		Analisis Metac	Analisis Total	Analisis recin	Descripcion	Idioma	Nombre	Plataforma	Precio	o con Desc	URL				
2	0		98	99	Iron Snout	['Ingles']	Iron Snout	['Windows']	Free		http://store.steampowered.com/app/424280/iron_Snout				
3	1		98		Battle throi	['Ingles']	DUSK	['Windows']	16,99€		http://store.steampowered.com/app/519860/DUSK/?s				
4	2		97	95	The Literati	['Ingles']	Doki Doki L	['Windows']	Free		http://store.steampowered.com/app/698780/Doki_Doki_Lite				
5	3	85	97	96	Hotline Mis	['Ingles'], 'Es	Hotline Mis	['Windows']	8,49€		http://store.steampowered.com/app/219150/Hotline_Miami				
6	4	83	97	98	Detention i	['Ingles']	Detention	['Windows']	11,99€		http://store.steampowered.com/app/555220/Detention_The_Theatrical_Edition				
7	5		97	100	Oats Studic	['Ingles']	Oats Studic	['Windows']	Free		http://store.steampowered.com/app/633030/Oats_Studios_-_Free				
8	6	85	96	95	Shadow Tai	['Ingles'], 'Es	Shadow Tai	['Windows']	39,99€		http://store.steampowered.com/app/418240/Shadow_Tactics				
9	7		96	93	GORN is a l	['Ingles']	GORN	['Windows']	19,99	14,99	http://store.steampowered.com/app/578620/GORN/?s				
10	8	76	96	96	DEADBOLT	['Ingles']	DEADBOLT	['Windows']	9,99€		http://store.steampowered.com/app/394970/DEADBOLT				
11	9		95	94	Live a week	['Ingles']	POSTAL 2	['Windows']	9,99€		http://store.steampowered.com/app/223470/POSTAL_2				
12	10		95	91	Fight, sneal	['Ingles'], 'Es	Streets of R	['Windows']	14,99€		http://store.steampowered.com/app/512900/Streets_of_Road				
13	11		95	85		['Ingles']	POSTAL 2: i	['Windows']	5,99€		http://store.steampowered.com/app/360960/POSTAL_2_The_Theatrical_Edition				
14	12		95	95	DownFall is	['Ingles']	Half-Life 2:	['Windows']	Free		http://store.steampowered.com/app/587650/Half-Life_2_The_Theatrical_Edition				
15	13		95			['Ingles'], 'Es	Blacksquad	['Windows']	Free		http://store.steampowered.com/app/654170/Blacksquad				
16	14	88	95	83	It's 1930. A	['Ingles'], 'Es	Mafia	['Windows']	9,99€		http://store.steampowered.com/app/40990/Mafia_2				
17	15		95	100	Having defi	['Ingles'], 'Es	Turok 2: Se	['Windows']	19,99€		http://store.steampowered.com/app/405830/Turok_2_The_Final_Chapter				
18	16		100	100	"All ended i	['Ingles']	Hells Chae	['Windows']	9,99€		http://store.steampowered.com/app/350800/Hells_Chance				

En caso de haber ejecutado el programa antes, ya dispondremos de este excel1 creado, con una segunda página que recoge nuestros juegos favoritos con sus precios,



	A	B	C	D	E	F	G	H	I	J	K	L
1	Nombre	Precio										
2	Iron Snout	Free										
3	DUSK	10										
4	Doki Doki Lite	Free										
5	Hotline Miami	8,49€										
6	Detention	11,99€										
7	Oats Studios -	Free										
8	Shadow Tactics	20										
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												

con lo que al ejecutar el programa de nuevo nos generará otro excel con la variación de precio -si la hubiera- respecto al primer precio registrado.



	A	B	C	D	E	F	G	H	I	J
1		Nombre	Precio + Reciente	Precio viejo						
2	0	Iron Snout	~	Free						
3	1	DUSK	16,99€	10						
4	2	Doki Doki Literature Club!	~	Free						
5	3	Hotline Miami	~	8,49€						
6	4	Detention	~	11,99€						
7	5	Oats Studios - Volume 1	~	Free						
8	6	Shadow Tactics: Blades of the Shogun	39,99€	15						
9										
10										
11										
12										
13										
14										
15										
16										
17										

Descripción del proyecto

• Módulos creados

Clases:

Juego:

Clase creada para almacenar toda la información del juego.

- I. Título
- J. Precio (original y con descuento, si lo hubiese)
- K. Breve Descripción
- L. URL del sitio web del juego
- M. Plataformas
- N. Idiomas: español (sí/no), inglés (sí/no)

O. Valoración de la comunidad:

- a. Análisis recientes: valoración (0-100) sobre el número de análisis de los últimos 30 días.
- b. Análisis totales: valoración (0-100) sobre el número total de análisis

P. URL

Ventana Principal:

Clase creada para soportar toda la interfaz gráfica del proyecto

Funciones:

Categorias_tags():

Función que nos devuelve un array con los nombres de las categorías y otro con la referencia que steam tiene para cada categoría. Así logramos mostrar los nombres de las categorías.

urls(tag,page):

Función que nos devuelve las diferentes url de cada juego para la categoría seleccionada.

info_Juego(url_juego):

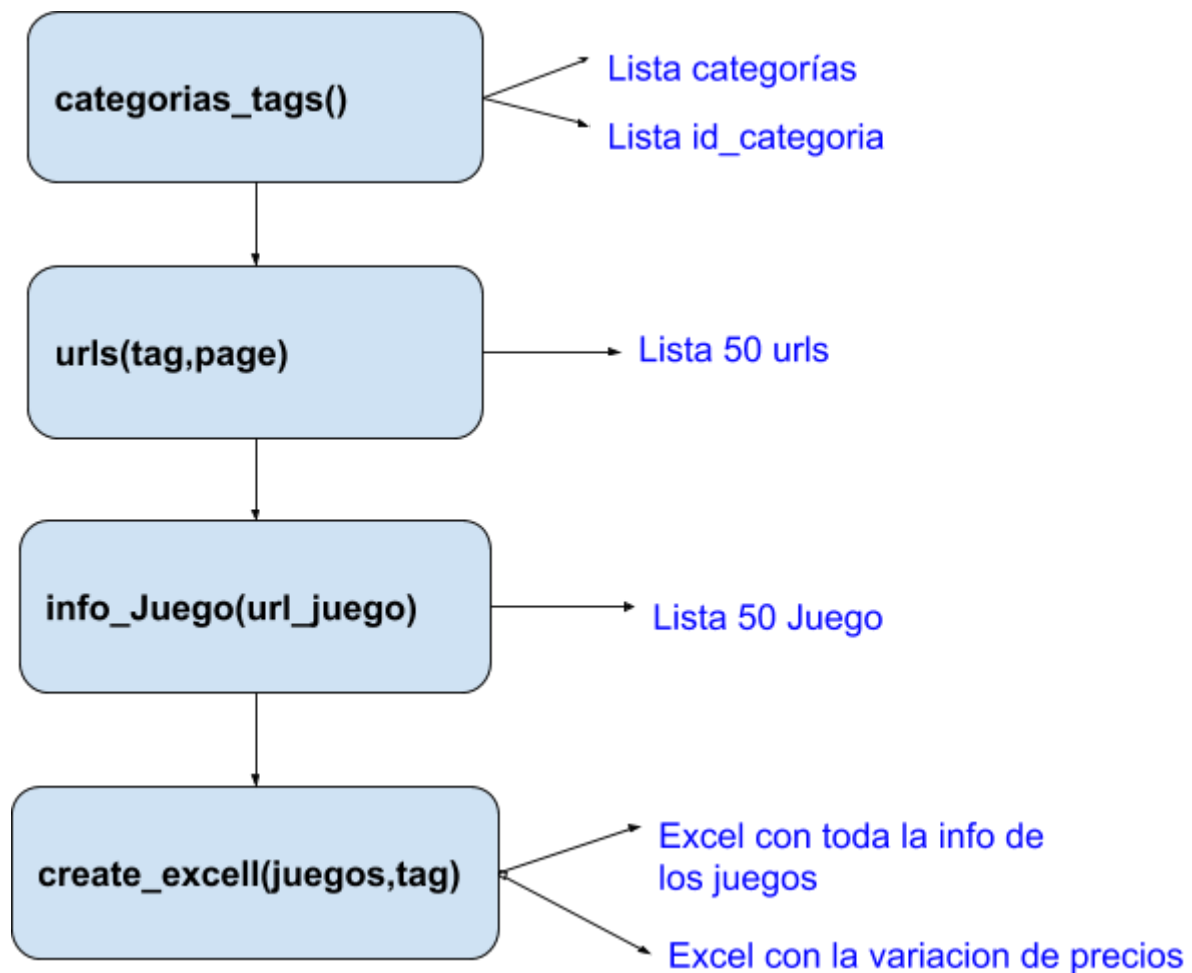
Función que a partir de una url de un juego nos devuelve toda la información del juego.

Esta función a su vez llama a otra **analisis_juego(url_juego,nombre)** para poder recuperar los análisis de los usuarios

create_excell(juegos,tag):

Función que a partir de una lista de juegos genera un excel.

Si el excel que va a crear ya existe, dicha función llama a **precion_change(xl,juego,tag):** si en la segunda página del excel disponemos de nuestros juegos favoritos con su precio, generará otro excel con la variación de los precios de esos juegos.



Librerias:

import requests:

Para poder cargar el html de la url.

import re:

Para poder realizar expresiones regulares.

from bs4 import BeautifulSoup:

Para poder realizar el web Scraping con las etiquetas.

import pandas:

Para poder realizar la escritura y lectura de los excels

WebScraping en Metacritic

Parece que el portal Metacritic mediante su nueva política de privacidad a cerrado el acceso a los web scraping. Información obtenida en foros, ejemplo

<https://www.themoviedb.org/talk/5923f2bec3a3680465009cfb>