

Termómetro: el buscador de Grados universitarios

Adrián Rico, Juan Tomás y Marc Vicedo © 2023

Idea inicial

Demasiadas carreras que estudiar en la Comunitat Valenciana, ¿cómo decidirse solo por una al hacer la preinscripción?



Idea inicial

Demasiadas carreras que estudiar en la Comunitat Valenciana, ¿cómo decidirse solo por una al hacer la preinscripción?

... con un **comparador de Grados** de las cinco universidades públicas valencianas



VNIVERSITAT
DE VALÈNCIA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT
JAUME I



Universitat d'Alacant
Universidad de Alicante



UNIVERSITAS
Miguel Hernández

¿Y cómo llevaremos a cabo
nuestro **proyecto**?

Utilizando **web scraping**



Extracción de información de
Internet, accediendo al
protocolo HTTP manualmente
(y no a través de un buscador)

Desarrollo (I)

Las cinco universidades tienen la información acerca de sus estudios disponible en sus páginas web, aunque de distinta forma. Para la presentación tomaremos de ejemplo la **Universitat Politècnica de València**



VNIVERSITAT
DE VALÈNCIA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

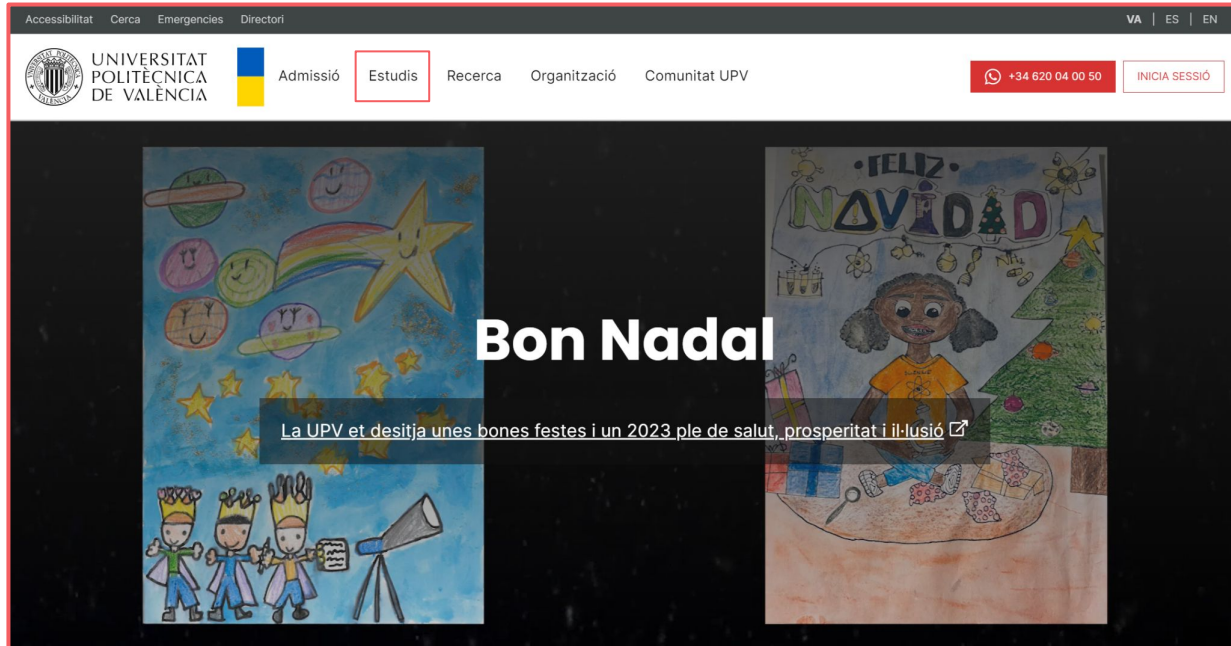


Universitat d'Alacant
Universidad de Alicante

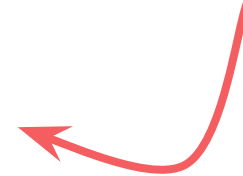


UNIVERSITAS
Miguel Hernández

Desarrollo (I)



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Desarrollo (I)

En primer lugar, obtenemos los enlaces de la página web que necesitamos, utilizando la librería Python **BeautifulSoup**

```
7
8 import requests
9 from bs4 import BeautifulSoup
10 import csv
11 dic={}
12 l_links=[]
13 import re
14 result = requests.get("https://www.upv.es/noticias-upv/noticia-13015-45-grados-y-do-es.html")
15 src = result.content
16 soup = BeautifulSoup(src, 'lxml')
17 links=soup.find_all("p",class_='texto_baseGN3')
18 for link in links:
19     try:
20         link=link.find('a')
21         l_links.append(link.attrs['href'])
22     except:
23         next
24 l_links=set(l_links)
25 l_links=list(l_links)
```

Desarrollo (I)

A continuación, seleccionamos los enlaces que nos interesa conservar: los que se refieran a **estudios** de la UPV

```
27  for link in l_links:
28
29      if 'estudios' in str(link):
30          l_link=[]
31          result = requests.get(link)
32          src = result.content
33          soup = BeautifulSoup(src, 'lxml')
34          links=soup.find_all('p',class_='texto_baseGN3')
35          for link in links:
36              try:
37                  link=link.find('a')
38                  l_link.append(link.attrs['href'])
39
```


Desarrollo (I)

Por último, exportamos a un **fichero CSV** la información extraída de los enlaces que hemos seleccionado

```
83     with open('upv3.csv', 'w', newline='') as csvfile:
84         fichero = csv.writer(csvfile, delimiter=';')
85         for k, v in dic.items():
86             nombre, lugar = k
87             lista = []
88             lista.append(nombre)
89             lista.append(lugar)
90             lista += v
91             fichero.writerow(lista)
92
```

Desarrollo (I)

Este proceso de adquisición de datos se repetirá de manera análoga para las otras cuatro universidades

```
8 import requests
9 from bs4 import BeautifulSoup
10 import csv
11 dic={}
12 l_links=[]
13 result = requests.get("https://web.ua.es/es/grados-oficiales.html#dobles-grados")
14 src = result.content
15 soup = BeautifulSoup(src, 'lxml')
16
17 tablas=soup.find_all("div",class_='row no-gutters')
```

```
28 for link in l_links:
29     if 'grados' in str(link):
30         result = requests.get(link)
31         src = result.content
32         soup = BeautifulSoup(src, 'lxml')
33         nombre=soup.find('h1',class_='textoTituloPagina').text
34         lugar='Universidad de Alicante'
35         tupla=(nombre,lugar)
36         dic[tupla]=[]
37         try:
38             carac=soup.find('p',class_='aviso')
39             dic[tupla].append(carac.text)
40         except:
41             carac=soup.find('address',class_='aviso')
42             dic[tupla].append(carac.text)
43
```

Desarrollo (II)

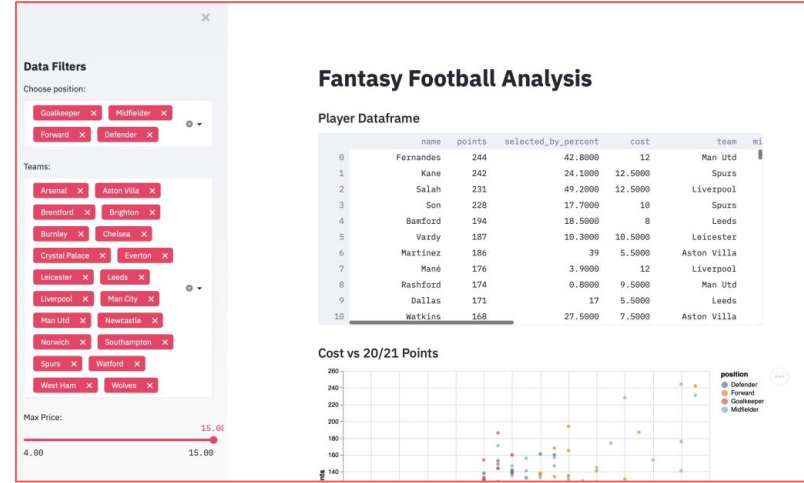
Después de tener la información en cinco ficheros distintos y sin ningún formato, necesitamos unirlos todos en un único fichero:

	A	B	C	
1	Grado en Dis	Campus de V	3 cursos 180	Créd
2	Grado en Bel	Campus de V	4 cursos 240	Créd
3	Grado en Tec	Campus de C	4 cursos 240	Créd
4	Doble grado	Campus de V	5 cursos 330	Créd
5	Doble grado	Campus de C	5 cursos 330	Créd
6	Grado en Ing	Campus de V	4 cursos 240	Créd



	A	B	C	D	
1	ID	GRADO	UNIVERSIDAD	ABREVIATUR	NO
2		1 Grado en Dis	Universitat P	UPV	
3		2 Grado en Bel	Universitat P	UPV	
4		3 Grado en Tec	Universitat P	UPV	
5		4 Doble grado	Universitat P	UPV	
6		5 Doble grado	Universitat P	UPV	

Desarrollo (II)



Para poder mostrar la información al usuario, utilizaremos una librería Python con herramientas gráficas de visualización de datos: **Streamlit**

Desarrollo (II)



```
11 class Grado:
12     def __init__(self, id, nombre, uni, uni_a, nota, plazas):
13         self.id = int(id)
14         self.nombre = nombre
15         self.uni = uni
16         self.uni_a = uni_a
17         self.nota = int(nota)
18         self.plazas = int(plazas)
19
20 class Clase:
21     def __init__(self):
22         self.carreras = {}
23
24     def nueva_carrera(self, nombre, uni_a, nota, plazas):
25         self.carreras[id] = Grado(id, nombre, uni_a, nota, plazas)
26         return True, 'Añadido correctamente'
27
28     def cargar_datos(self, fichero):
29         try:
30             with open(fichero, 'r', encoding='utf-8') as csvfile:
```

El programa leerá el fichero CSV línea a línea, para mostrar por pantalla los grados que solicite el usuario

Este es nuestro **producto final**:

Termómetro: el buscador de Grados universitarios

Adrián Rico, Juan Tomás y Marc Vicedo © 2023

Aquí podrás encontrar tu próxima carrera universitaria.

¿Qué quieres estudiar?

ciencia de datos

1 - Grado en Ciencia de Datos - Universitat Politècnica de València (Campus de Vera)

2 - Grado en Ciencia de Datos - Universitat de València

¿Quieres más información sobre algún grado?

1 - Grado en Ciencia de Datos - Universitat Politècnica de València (Campus de Vera) |

Grado en Ciencia de Datos en la UPV

Nota de corte: 12.12

Este grado tiene 120 plazas

Pulsa **aquí** para entrar a la aplicación

Posibles ampliaciones

- Incluir los estudios de **Máster**
- Incluir los estudios de las **universidades privadas**



Universidad
**Católica de
Valencia**
San Vicente Mártir



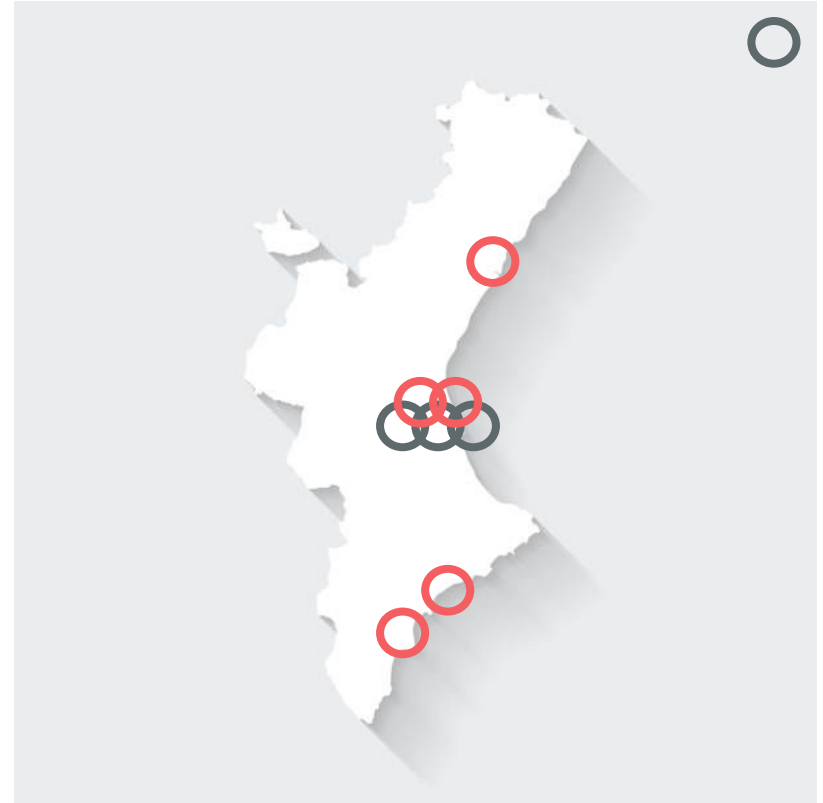
CEU
*Universidad
Cardenal Herrera*



**Universidad
Europea** VALENCIA



**Universidad
Internacional
de Valencia**



Termómetro: el buscador de Grados universitarios

Adrián Rico, Juan Tomás y Marc Vicedo © 2023