

# Pràctica 1

## 1. Context. Explicar en quin context s'ha recol·lectat la informació.

Explicar per què el lloc web triat proporciona aquesta informació.

Els dos membres de l'equip ens trobem en una situació similar, on estem cercant pisos per independitzar-nos. Per fer-ho, cerquem en diferents portals web dedicats a la venda i lloguer d'habitatges. Quan hem fet aquesta recerca per trobar un nou habitatge, hem vist que s'ha de dedicar molt de temps per tal de trobar aquells pisos que encaixen amb les necessitats i expectatives de l'usuari, ja que trobem massa oferta de pisos que no s'ajusten als nostres criteris.

Hi ha diferents llocs web on es poden trobar pisos. Després d'analitzar els diferents portals més populars, hem observat que un dels que ofereix pisos a un preu més competitiu i utilitza una interfície amigable pels usuaris és *pisos.com*. A més, una funcionalitat que ens va cridar l'atenció és la possibilitat de poder guardar en favorits aquells pisos que ens resultin més interessants, de manera que els tinguem més controlats i poder rebre notificacions sobre aquests. Per aquest motiu hem decidit triar aquesta pàgina web per fer aquesta pràctica.

Finalment, com que els dos vivim a la comarca del Baix Llobregat i ens sembla que Esparreguera satisfà les nostres necessitats així com concentra una bona part del serveis essencials, hem escollit aquest poble per fer la cerca d'habitatges.

## 2. Definir un títol pel dataset. Triar un títol que sigui descriptiu.

Dades dels pisos d'Esparreguera.

## 3. Descripció del dataset. Desenvolupar una descripció breu del conjunt de dades que s'ha extret (és necessari que aquesta descripció tingui sentit amb el títol triat).

El dataset conté les dades relacionades amb els immobles del poble d'Esparreguera situat a la comarca del Baix Llobregat. Cadascun dels registres recollirà un subconjunt de les dades que es mostren al portal web i ens permetrà visualitzar ràpidament aquells immobles que són d'interés i agilitzar la cerca d'un pis que s'ajusti a les nostres necessitats.

#### 4. Representació gràfica. Presentar una imatge o esquema que identifiqui el dataset visualment.

El dataset que hem presentat té un total de 340 tuples (habitatges), i para cada d'aquestes desm 12 atributs. Aquest atributs són títol, zona, preu, número d'habitacions, número de banys, metres quadrats, número de planta, preu per metre quadrat, antiguitat, estat de conservació, url i descripció. Adjuntem una captura de pantalla de les 10 primeres tuples del fitxer .CSV resultant.

341 l nes (341 sloc) | 322 KB

RawBlameHistory




 Search this file...

1	Titulo	Zona	Precio	Habitaciones	Aseos	Metros cuadrados
2	Casa adosada en venta en La Plana	Esparreguera	178000	4	2	217
3	Casa adosada en venta en Esparreguera	Esparreguera	275000	3	2	165
4	Casa en venta en Carrer Gran, n� 82	Esparreguera	217000	-	2	236
5	Chalet en venta en Can Rial	Esparreguera	195000	3	1	121
6	Piso en venta en Avinguda de Francesc Maci�, cerca de Carrer de Ponent	Esparreguera	234500	3	2	100
7	Piso en venta en Carrer del Cam� Ral, cerca de Carrer del Doctor Fleming	Esparreguera	165000	3	1	88
8	Casa en venta en Esparreguera, Zona de - Esparreguera	Esparreguera	200000	4	1	120
9	Casa en venta en Esparreguera	Esparreguera	200000	3	1	90
10	Casa en venta en Esparreguera	Esparreguera	199000	4	1	150

341 l nes (341 sloc) | 322 KB

RawBlameHistory

 Search this file...

Planta	Precio (�/m <sup>2</sup> )	Antig�edad	Estado conservacion	URL	
-	820	Entre 30 y 50 a�os	En buen estado	<a href="https://www.pisos.com/comprar/casa_adosada-esparreguera_centro_urbano-26927192_500732/">https://www.pisos.com/comprar/casa_adosada-esparreguera_centro_urbano-26927192_500732/</a>	Ca
-	1.666	Entre 20 y 30 a�os	En buen estado	<a href="https://www.pisos.com/comprar/casa_adosada-esparreguera_centro_urbano-100036323922_100500/">https://www.pisos.com/comprar/casa_adosada-esparreguera_centro_urbano-100036323922_100500/</a>	Se
-	919	-	-	<a href="https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-100006217030_109200/">https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-100006217030_109200/</a>	Ca
-	1.611	Entre 30 y 50 a�os	En buen estado	<a href="https://www.pisos.com/comprar/chalet-esparreguera_centro_urbano-100067745906_101800/">https://www.pisos.com/comprar/chalet-esparreguera_centro_urbano-100067745906_101800/</a>	**
4�	2.345	Entre 10 y 20 a�os	En buen estado	<a href="https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-2547570943_999081/">https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-2547570943_999081/</a>	Pis
-	1.875	Entre 10 y 20 a�os	En buen estado	<a href="https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-1698625027_501243/">https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-1698625027_501243/</a>	Et
-	1.666	-	En buen estado	<a href="https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-95013243918_100500/">https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-95013243918_100500/</a>	Ur
Bajo	2.222	Entre 30 y 50 a�os	En buen estado	<a href="https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-99247113175_519327/">https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-99247113175_519327/</a>	90
S�tano	1.326	Entre 30 y 50 a�os	Reformado	<a href="https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-90036078325_513853/">https://www.pisos.com/comprar/casa-esparreguera_centro_urbano-90036078325_513853/</a>	Pr
1�	2.215	Entre 10 y 20 a�os	En buen estado	<a href="https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-95025960710_519351/">https://www.pisos.com/comprar/piso-esparreguera_centro_urbano-95025960710_519351/</a>	Pis

341 lines (341 sloc) | 322 KB

Raw	Blame	History
-----	-------	---------



🔍 Search this file...

Description

<p>Casa a reformar adosada 116 m<sup>2</sup> de 2 plantas en parcela de 120 m<sup>2</sup>, en esparraguera, la zona de la plana. Planta baja: salón comedor. Cocina independiente con salida a patio. Una habitación.</p> <p>Se vende casa adosada en la zona del castell de esparraguera, zona muy tranquila a la vez que tienes todos los servicios necesarios a poca distancia, como escuelas e instituto, transporte público.</p> <p>Casa de 236m<sup>2</sup> Muy céntrica, con local comercial en la planta baja totalmente reformado, con baño Local p. Baja de 96m<sup>2</sup> + patio de 70m<sup>2</sup> P. Primera de 72m<sup>2</sup> (3 habitaciones) P. Segundo de 72m<sup>2</sup> (3 habitaciones).</p> <p>** casa de 12 m<sup>2</sup> construidos. 88 m<sup>2</sup> útiles. + Garaje de 24 m<sup>2</sup>** Construida sobre parcela de 520 m<sup>2</sup>. 3 habitaciones. 1 baño con bañera. Calefacción gasoil. Ventanas de aluminio. Mosquiteras.</p> <p>Piso en venta en Francesc Macià. Zona Vinyals Esparraguera. ¿Cuánto tiempo llevas buscando un elegante piso para ti y tu familia con unas vistas panorámicas a la escultural montaña de Collserola?</p> <p>Et presentem una llar encantador que reuneix comoditat, proximitat i estades equilibrades, situat a escassos metres de el centre de la població d'Esparraguera. La finca disposa d'ascensor.</p> <p>Urb. Mas d'en gall - parcela 600 m<sup>2</sup> - vivienda 80 m<sup>2</sup> + 30 m<sup>2</sup>. Garaje. . 4 locales comerciales. Salón comedor 22 m<sup>2</sup> con salida a porche, 4 habitaciones (3 dobles), cocina office, baño nuevo.</p> <p>90 m<sup>2</sup>. Vivienda - parcela plana 562 m<sup>2</sup> - antigüedad 40 años Salón comedor 24 m<sup>2</sup>. Con salida balcón y con chimenea. Cocina office de origen, baño completo con bañera, 3 habitaciones.</p> <p>Propietaris.cat ®, líder en catalunya para la venta entre particulares sin comisiones, ofrece: - fantástica casa en una sola planta con vistas panorámicas espectaculares y privacidad total.</p> <p>Piso muy centrico , cerca de transporte publico y todos los servicios, piso para entrar a vivir , tiene osmosis y descalcificador , cocina americana piso muy amplio</p>
---

Per veure'l amb més detall, es pot consultar al nostre repository [github](#).

A continuació mostrem d'on estem obtenint la informació del nostre CSV. Per cada pis, obtenim les dades dels requadres que es mostra en la següent captura:

< Ver más pisos y apartamentos en Esparreguera

> 46 de 122 Pisos en Esparreguera >

**TITULO**

Piso en venta en Esparreguera

**ZONA**

Esparreguera

**PRECIO**

98.000 €

Solicita tu hipoteca ✓ Aviso activado

pisos.com

Fotos (1/10) Mapa

**METROS CUADRADOS**

74 m <sup>2</sup>	4 habs.	1 baño	3ª planta	1.324 €/m <sup>2</sup>
-------------------	---------	--------	-----------	------------------------

**PRECIO (€/m<sup>2</sup>)**

**HABITACIONES**

ASEOS PLANTA

DESCRIPCION CARACTERISTICAS HIPOTECAS GALERIA UBICACION

**DESCRIPCION**

Descripción del piso en venta en Esparreguera

Traducciones disponibles: Español

- 74 m<sup>2</sup>. Vivienda - Salón comedor 19 m<sup>2</sup>, 4 Habitaciones (2 Dobles), Cocina office, Baño con plato de ducha, Carpintería int. haya, exterior aluminio climatit con mosquiteros, Suelo gres, Antena parabólica, Calefacción gas. Reformado hace 8 años.

Pide más información

**Datos básicos**

Superficie construida: 74 m <sup>2</sup>	Superficie útil: 74 m <sup>2</sup>
Habitaciones: 4	Baños: 1
Planta: 3ª	Interior

**ANTIGÜEDAD**

Antigüedad: Entre 10 y 20 años

**ESTADO CONSERVACION**

Conservación: Reformado

**Inmuebles de FINQUES EL POBLE**

Actualizado el 25/03/2020

publicar

Adapta Reformas: Depósitos en todo tipo Reformas  
Empresa de Reformas Integrales y Construcción en Barcelona | ADAPTA REFORMAS

Especialistas en Reformas Integrales y Construcción (PISO de Privacidad Control)

PROVINCIAL DEL PORTA NOU DIAGONAL DEL MARITIM EL POBLE

DISTRITO DE

SITIO WEB LLAMAR

Cal dir que al .CSV hi posem un paràmetre més que és la URL de l'habitatge, per si es desitja consultar aquell pis.

5. Contingut. Explicar els camps que inclou el dataset, el període de temps de les dades i com s'ha recollit.

El dataframe té un total de 16 atributs. Per tant, cada fila o registre del dataframe que correspon a un pis contindrà la següent informació:

- **Títol:** És una cadena de caràcters (String) que conté una petita descripció de l'immoble.
- **Zona:** És una cadena de caràcters (String) que referencia al barri on es troba l'habitatge.
- **Preu:** És una dada de tipus numèric enter (Integer) que fa referencia al cost total de l'immoble.
- **Habitacions:** És una dada de tipus numèric enter (Integer) que fa referència al nombre d'habitacions amb les quals compte el pis.
- **Lavabos:** És una dada de tipus numèric enter (Integer) que fa referència al nombre de banys amb els quals compte el pis.
- **Metres quadrats:** És una dada de tipus numèric enter (Integer) que fa referencia al nombre de metres quadrats habitables amb els quals compte el pis.
- **Planta:** És una cadena de caràcters (String) que fa referència a la planta de l'edifici on es troba situat el pis
- **Preu/m2:** És una dada de tipus numèric decimal (Double) que fa referència al cost del metre, és a dir, la proporció entre el preu total i els metres quadrats.
- **Descripció:** És una cadena de caràcters (String) que descriu en un parell de frase l'habitatge, les seves principals característiques i situació.
- **Antiguitat:** És una dada de caràcter numèric enter (Integer) que descriu els anys des de que va ser construït el pis.
- **Estat de conservació:** És una cadena de caràcters (String) que especifica les condicions en què es troba el pis.
- **URL:** És una cadena de caràcters (String) que referencia al pis scrapejat.

6. Agraïments. Presentar el propietari del conjunt de dades. És necessari incloure cites de recerca o anàlisis anteriors (si n'hi ha).

El propietari d'aquest conjunt de dades és *Pisos.com*. Els habitatges que es publiquen a la pàgina web són de particulars o d'agències immobiliàries, i són ells els encarregats de que aquestes dades, tal i com s'explica en l'apartat d'avisos legals:

*“En todo caso el Usuario será el único responsable de las manifestaciones falsas o inexactas que realice y de los perjuicios que cause a pisos.com o a terceros por la información que facilite.”<sup>[1]</sup>*

També ens diu al seu apartat d'avisos legals:

*“Los contenidos, elementos e información a los que el usuario pueda acceder a través de pisos.com están sujetos a derechos de propiedad industrial e intelectual, patentes marcas, copyright del titular de pisos.com o de terceros titulares de los mismos. En consecuencia, el acceso a estos contenidos o elementos no otorga al Usuario el derecho de alteración, modificación, explotación, reproducción, distribución o comunicación pública o cualquier otro derecho que corresponda al titular del derecho afectado.”<sup>[1]</sup>*

Pel que ens diu que com a usuaris podem accedir al seu contingut, però no podem fer cap alteració, modificació, explotació, reproducció o distribució de les seves dades.

7. Inspiració. Explicar per què és interessant aquest conjunt de dades i quines preguntes es pretenen respondre.

Aquest dataframe primer de tot, recull tota la informació referent als pisos de la zona d'Esparreguera en un únic document. De manera que, amb una primera ullada ens podem adonar d'aquells que són del nostre interès sense haver de navegar entre les diferents pàgines que té el lloc web.

*Pisos.com*, de la mateixa manera que ho fan la resta de portals web dedicats a la recerca de pisos, ofereixen diferents filtres per trobar habitatges. Tot i això, a vegades els filtres són massa generals, alenteixen la cerca i no incorporen la combinació de filtres desitjats. Degut això, amb la implementació del nostre dataset podem fer consultes més precises.

Per posar un exemple, si des de *Pisos.com* volem cercar pisos que valguin menys de 135.000€, veiem que no podem ja que funciona per rangs de valors. Tal i com podem veure



a la següent imatge, o bé hem de cercar pisos fins a 150.000€ o 120.000€, cosa que no és exactament el que desitgem.



En canvi, al nostre dataset si que podem fer les consultes tal i com volem, ja que es pot fer servir el potencial d'un CSV i els fulls de càlcul.

8. Llicència. Seleccionar una d'aquestes llicències pel dataset resultant i explicar el motiu de la seva selecció:

- Released Under CC0: Public Domain License
- Released Under CC BY-NC-SA 4.0 License
- Released Under CC BY-SA 4.0 License
- Database released under Open Database License, individual contents under Database Contents License
- Other (specified above)
- Unknown License

El dataset resultant estarà sota la llicència CC BY-NC-SA 4.0 License. Tal i com es diu a Creative Commons (organització sense ànim de lucre dedicada a construir un conjunt públic de coneixement i cultura accessible globalment), aquesta llicència és la més restrictiva de les llicències estàndard, ja que aquesta només permet la descàrrega de les obres i que

només es puguin compartir amb el públic sempre que en reconeguin l'autoria, però no poden ser modificades de cap manera ni ser utilitzades amb finalitats comercials.<sup>2,3</sup>

Hem escollit aquesta llicència ja que la web *Pisos.com*, tal i com hem explicat en l'apart 6, ens permet com a usuaris accedir al seu contingut, però no ens deixa fer cap alteració, modificació, explotació, reproducció o distribució de les seves dades. Nosaltres ho hem fet amb l'objectiu de realitzar una pràctica acadèmica, però no volem que es pugui utilitzar aquest dataset amb finalitats comercials.

9. Codi. Adjuntar el codi amb el qual s'ha generat el dataset, preferiblement en Python o, alternativament, en R.

El codi d'aquesta pràctica està disponible al github públic següent:  
<https://github.com/paucampana/pisosScrapper/>

Hem desenvolupat el projecte en Python 3.

10. Dataset. Publicar el dataset en format CSV a Zenodo amb una xicoteta descripció.

Hem publicat el dataset en format .CSV a Zenodo. Es pot trobar a:  
<https://doi.org/10.5281/zenodo.3747646>



## Webgrafia

1. Pisos.com. 2020. *Aviso Legal*. [online] Available at: <<https://www.pisos.com/avisolegal>> [Accessed 5 April 2020].
2. Creativecommons.org. 2020. *About The Licenses - Creative Commons*. [online] Available at: <<https://creativecommons.org/licenses/>> [Accessed 8 April 2020].
3. Creativecommons.org. 2020. *Creative Commons - Attribution-NonCommercial-NoDerivatives 4.0 International*. [online] Available at: <<https://creativecommons.org/licenses/by-nc-nd/4.0/>> [Accessed 8 April 2020].