

Viviendas en Alquiler

Descripción

Se trata de un conjunto de datos que incorpora los pisos vigentes, según la fecha de extracción, en régimen de alquiler para la zona indicada.

Los atributos corresponden a las características de la vivienda como la situación, precio, metros, habitaciones y baños.

Para la construcción del conjunto de datos y con motivo de testear varias de las características del código, se ha realizado un filtro para el conjunto de todas las viviendas disponibles en la zona de Barcelona y Madrid. No se ha impuesto ningún límite de precios.

Imagen



Contexto

Tanto en la prensa económica como en foros especializados, se advierte un claro incremento en los precios del alquiler para las viviendas situadas en las grandes urbes como puede ser Barcelona, Madrid y sus alrededores.

Es por ello, que se ha realizado una extracción para estas dos zonas y así comparar tanto los precios medios como las características de estas viviendas.

Las viviendas, pueden ser pisos, apartamentos, loft, dúplex, áticos. Sin límites de precio ni de superficie.

Contenido

- **Tipo:** pisos, apartamentos, áticos, estudios, lofts, dúplex
- **Prov_Ciudad:** provincia o ciudad, según el contexto de la búsqueda
- **Zona:** barrio o zona delimitada
- **Calle:** la calle en donde se sitúa el inmueble
- **Precio:** coste del alquiler en €/mes
- **Superficie:** tamaños de la vivienda en metros cuadrados
- **Habitaciones:** número de habitaciones que dispone
- **Baños:** número de baños

No existe un periodo de tiempo como tal, pues se extrae toda la oferta de viviendas en alquiler para el momento actual y según el criterio de la búsqueda.

En este caso, corresponde a la oferta Barcelona y Madrid a fecha de 28 de Marzo de 2018. Observaciones resultantes: 5933 filas.

Agradecimientos

Todos los datos han sido obtenidos de la página web <https://www.yaencontre.com>

Se revisaron varios portales inmobiliarios pero algunos tenían restricciones explícitas para las técnicas de scraping o de automatización. Por otro lado, la estructura del portal seleccionado era la más adecuada para su tratamiento.

Para llevar a cabo la práctica se ha hecho uso de los siguientes recursos:

- *Web Scraping with Python* - Richard Lawson
- *El lenguaje Python* - David Masip Rodó
- <https://github.com>
- <https://stackoverflow.com>

Inspiración

El conjunto de datos puede ser útil para realizar un estudio de minería de datos sobre el precio de la vivienda, según zonas o barrios, como también por características de los mismo. También se puede hacer uso de modelos predictivos, por ejemplo, asignar precios para nuevas viviendas.

Por otro lado, desde una visión más particular, en conjunto de datos puede proporcionar comparaciones entre distintas viviendas, como fluctuaciones del precio entre una extracción y otra.

Licencia

La licencia es **CC BY-NC-SA** (*Reconocimiento - No Comercial – Compartir Igual*)

Permite a otros entremezclar, ajustar y construir a partir de su obra con fines no comerciales, siempre y cuando le reconozcan la autoría y sus nuevas creaciones estén bajo una licencia con los mismos términos.

Escogí esta licencia por su enfoque no comercial, pues los datos se han extraído de un negocio, además de tratarse de una práctica del ámbito educativo.

Esta licencia permite reutilizar el código y modificarlo para adaptarlo pero siempre manteniendo el mismo tipo de licencia por lo que se respeta el espíritu colaborativo y de intercambio cultural.

Código

El código se encuentra en el archivo: **scraping.py** y debe ser ejecutado bajo la versión de Python 2.7

Para la correcta ejecución son necesarios los siguientes paquetes:

```
import re
import csv
import urllib2
import requests
import argparse
import BeautifulSoup
```

Argumentos para la ejecución del fichero:

```
--lugar      dónde se realizará la búsqueda, puede ser ciudad o provincia
--paginas    número de páginas a realizar scraping, por defecto será 50
--precio     precio máximo del alquiler en euros/mes
--acumular   se mantienen los registros antiguos, por defecto está deshabilitado
```

Ejemplos

```
[Python scraping.py --lugar valencia --paginas 10 --precio 800]
```

En la ciudad de Valencia hasta 10 páginas de consultas, con un precio máximo 800 €/mes.

```
[Python scraping.py --lugar Madrid --paginas 100]
```

En la ciudad de Madrid hasta 100 páginas de consultas, sin límite de precio.

```
[Python scraping.py --lugar Barcelona-provincia --precio 1000 --acumular 1]
```

En la provincia de Barcelona, con un precio máximo 1000 €/mes. No se borrará los registros anteriores del archivo por lo que se añadirán estos nuevos.

```
[Python scraping.py]
```

Opción por defecto, ciudad de Barcelona, hasta 50 páginas y sin límite de precios.

Dataset

El conjunto de datos se ha formado con los siguientes argumentos:

```
>python scraping.py --lugar barcelona --paginas 300
```

```
>python scraping.py --lugar madrid-provincia --paginas 300 --acumular 1
```

El archivo creado es el siguiente: **dataset.csv**

Data Preview:							
Tipo (character)	Prov_Ciudad (character)	Zona (character)	Calle (character)	Precio (integer)	Superficie (integer)	Habitaciones (integer)	Baños (integer)
Piso	Barcelona	Horta	calle Horta	950	75	2	1
Piso	Barcelona	Eixample dreta	ronda Sant Pere	2550	126	4	2
Piso	Barcelona	Horta	paseo Maragall	825	80	3	1
Apartamento	Barcelona	NA	El Coll	600	45	2	1
Piso	Barcelona	NA	El Coll	650	50	2	1
Piso	Barcelona	Barcelona	calle De Cotonat	910	40	1	1
Piso	Barcelona	La Font d'en Fargues	calle Frederic Rahola	750	65	3	1
Piso	Barcelona	El Raval	calle De Balmes	3500	85	2	2
Estudio	Barcelona	La Salut	calle De Mercedes	550	31	1	0
Estudio	Barcelona	Sant Ramon-La Maternitat	calle Avinguda de Xile	850	25	1	0
Estudio	Barcelona	Barcelona	calle Avinguda del Torrent Gornal	780	20	1	0
Piso	Barcelona	Vila de Gràcia	calle Vila de Gràcia	700	40	1	1
Piso	Barcelona	El Guinardó	calle Sant Pere Dabanto	975	71	2	1
Piso	Barcelona	El Raval	calle Xucia	1000	66	2	1
Ático	Barcelona	Eixample esquerra	calle	2800	100	2	2
Dúplex	Barcelona	NA	Sant Gervasi – Galvany	7500	325	5	3
Piso	Barcelona	Sant Gervasi – La Bonanova	calle Doctor Roux	4600	215	5	5
Loft	Barcelona	NA	El Poble Sec-Montjuic	1210	90	1	1