

TIPOLOGÍA Y CICLO DE VIDA DE LOS DATOS

PRÁCTICA 1: WEB SCRAPPING

ALEJANDRO MARTÍNEZ OTAL| WEB SCRAPPING | 29/10/2019

# Contexto

Con este proyecto, a partir del portal de alquiler de pisos [www.pisos.com](http://www.pisos.com), se puede extraer información sobre el precio de alquiler de habitaciones junto otros parámetros como número de habitaciones, superficie útil, etc.

# Título dataset

El título elegido para el dataset es : Alquiler de habitaciones por provincia(s).

# Descripción dataset

El dataset en concreto para este ejemplo contiene todas las habitaciones disponibles de alquilar en la provincia de Barcelona para la fecha 29/10/2019. Este contiene algunos campos que se explican más adelante en el documento como precio, superficie….

# Representación gráfica



Imagen 1 El alquiler de piso siempre ha sido un tema polémico que preocupa, especialmente a jóvenes.

# Contenido

El dataset contiene para cada habitación en alquiler la siguiente información:

* Id: Identificador único del piso proporcionado por la propia web.
* Location: Localidad dentro de la provincia elegida dónde se encuentra el piso.
* Precio: Cantidad en € mensual para el alquiler de la habitación.
* Descripción: Se almacena una breve descripción.
* Superficieutil: Número de m2 de la habitación.
* Numbanos: Número de baños con los que cuenta el inmueble.
* Estadoconservación: Estado en el que se encuentra el inmueble.
* Gastosincluidosalquiler: Si el precio de la columna ‘Precio’ incluye los gastos.
* Numhabitaciones: Número de habitaciones con los que cuenta el inmueble.

# Motivación

El precio de las viviendas siempre ha sido un tema polémico durante los últimos años desde la crisis mobiliaria. Obviamente, el precio del alquiler de las habitaciones está ligado al precio de las viviendas y nuevas formas de alquiler proporcionadas por plataformas como Airbnb han aparecido como alternativas al alquiler tradicional.

Los resultados de la extracción se pueden utilizar directamente para encontrar una habitación según ciertas preferencias o utilizar como un juego de datos para estudiar el mercado.

# Agradecimientos

Todos los datos han sido extraídos de la página web [www.pisos.com](http://www.pisos.com) mediante una herramienta de Web Scrapping programada en Python.

# Licencia

TODO

# Código

El código utilizado se puede encontrar en <link github>

# Dataset

El dataset se puede encontrar en <link github> bajo el nombre <nombre csv>