

Análisis de viviendas en Madrid mediante Web Scraping-Python

Xinyuan Zheng



ÍNDICE

- 1** INTRODUCCIÓN
- 2** EXTRACCIÓN DE DATOS
- 3** BBDD - MYSQL
- 4** VISUALIZACIÓN DE DATOS



1. INTRODUCCIÓN



¿Quién es Fotocasa?

Fotocasa es una de las plataformas inmobiliarias más grandes de España.

Ofrece información sobre alquiler y compra-venta tanto viviendas nuevas y de segunda mano.

Entre sus filtros, incluyen datos sobre la ubicación, el precio y otras características esenciales para los arrendatarios.



fotocasa

Todos tenemos un sitio |

2. EXTRACCIÓN DE DATOS



2.1 Configurar parámetros iniciales

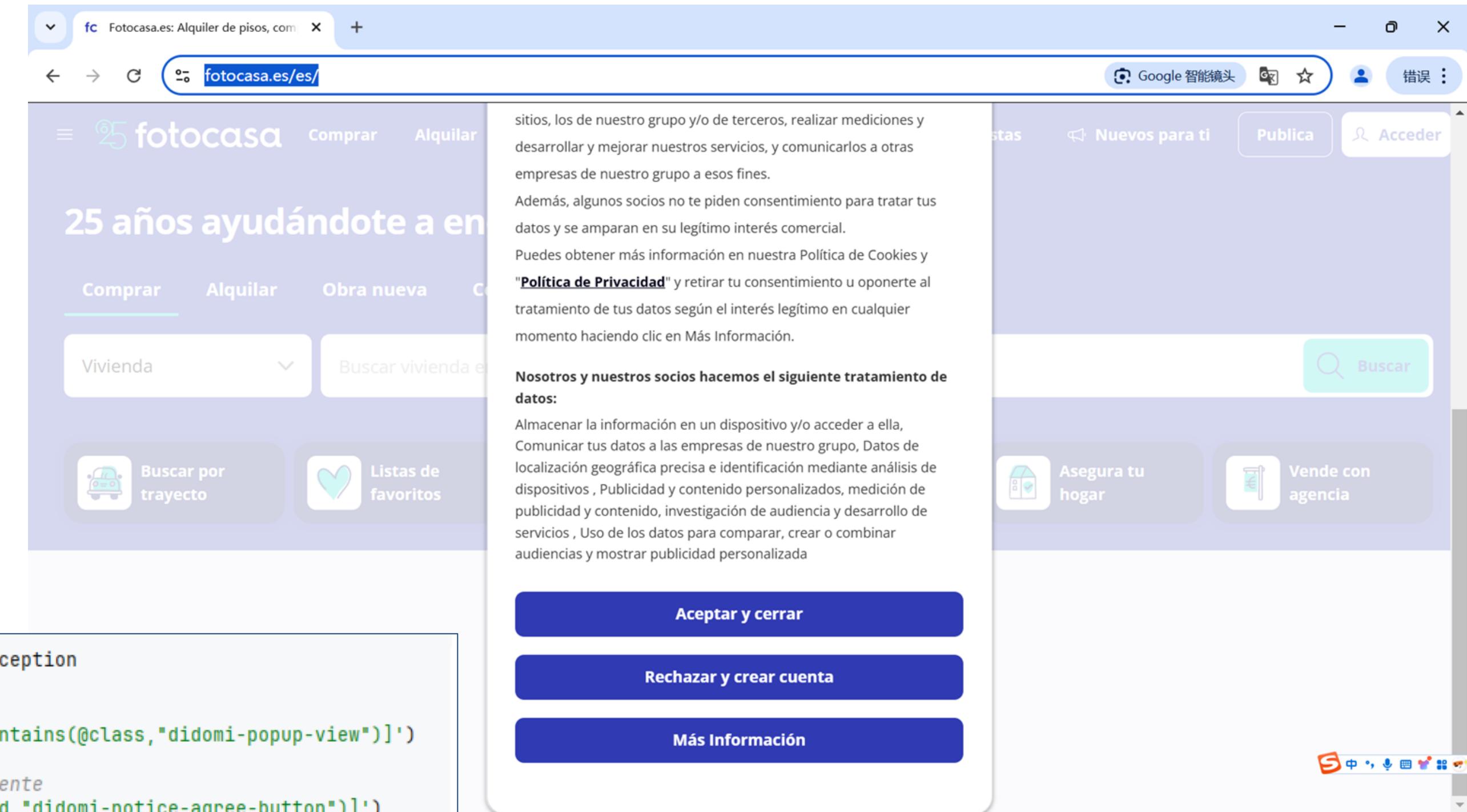
Paso 1: Acceder a la página web de “Fotocasa” con Selenium WebDriver



```
1 # pip install webdriver-manager
2 from selenium import webdriver
3 from selenium.webdriver.chrome.service import Service
4 from selenium.webdriver.common.by import By
5 from bs4 import BeautifulSoup
6 import undetected_chromedriver as uc
7
8
9 service = Service("D:\\\\XinYuan Zheng\\\\Downloads\\\\APP\\\\chromedriver-win64\\\\chromedriver-win64
10 \\\chromedriver.exe")
11 options = webdriver.ChromeOptions()
12 options.add_argument("--incognito")
13
14 driver = uc.Chrome()
15 driver.maximize_window()
16 url = "https://www.fotocasa.es/es/"
17
18 driver.get(url)
19 [2] 8s 11ms
```

2.1 Configurar parámetros iniciales

Paso 2: Gestión de la Ventana de Cookies



```
1 from selenium.common.exceptions import NoSuchElementException
2
3 # Verifica si aparece la ventana emergente de cookies
4 cookie_popup = driver.find_element(By.XPATH, '//div[contains(@class, "didomi-popup-view")]')
5
6 if cookie_popup: # Si la ventana de cookies está presente
7     driver.find_element(By.XPATH, '//button[contains(@id, "didomi-notice-agree-button")]')
8         .click()
9     print("Ventana de cookies encontrada y aceptada.")
10
11 else:
12     print('No se encontró la ventana de cookies')
13
14 [3] 4s 705ms
15
16 Ventana de cookies encontrada y aceptada.
```

2.1 Configurar parámetros iniciales

Paso 3: Selección de la Opción "Alquilar"



```
1 # Seleccionar el botón de "Alquilar" dentro de la sección de búsqueda
2 alquiler_button = driver.find_element(By.XPATH, '//div[contains(@class, "re-HomeSearchSelector-item re-HomeSearchSelector-item--rent")]')
3 alquiler_button.click()
```

2.1 Configurar parámetros iniciales

The screenshot shows a search interface with a blue header bar containing tabs: 'Comprar', 'Alquilar' (which is selected), 'Obra nueva', and 'Compartir'. Below the header is a dropdown menu set to 'Vivienda'. A large button labeled 'Buscar por trayecto' with a car icon is visible. On the right, a list of cities and their capitals is shown, with 'Madrid Capital, Madrid' being the top result. A QR code is present at the bottom.



The screenshot shows the fotocasa website with a search bar set to 'Madrid Capital'. It features filters for 'Filtros (0)', 'Distrito', 'Precio', and 'Habitaciones'. A banner at the top says 'El precio de tu casa en menos de 2 minutos. ¡Totalmente gratis!' and includes a 'Crear alerta' button. The main section displays 'Viviendas y casas de alquiler en Madrid Capital' with 7.056 results. One listing is highlighted: 'Piso con balcón en Sol' with a price of 1.200 €, 2 bedrooms, 1 bathroom, and 1 m². The listing includes icons for balcony, heating, and air conditioning. To the right is a map of the Madrid area with a callout showing options like 'Dibujar en mapa', 'Buscar por trayecto', and 'Vista satélite'.

```
1  from selenium.webdriver.common.keys import Keys
2  from selenium.webdriver.common.keys import Keys
3
4  search_box = driver.find_element(By.XPATH, '//div[contains(@class, "re-SharedSearch-content")]
5  ]//input')
6
7  # Escribir "Madrid Provincia" en la barra de búsqueda
8  search_box.send_keys("Madrid Capital, Madrid")
9
10 time.sleep(4) # Esperar para que se carguen las sugerencias
11
12 # Presionar ENTER para buscar
13 search_box.send_keys(Keys.RETURN)
time.sleep(4)
[40]
```

Paso 4: Búsqueda por Ciudad

2.1 Configurar parámetros iniciales

Paso 5: Aplicación de Filtros Avanzados

- . **Precio máx:** 1000 €
- . **Tipo de vivienda:** "Plantas intermedias"
- . **Fecha de publicación:** "última semana"

The screenshot shows a 'Filters' modal window with several sections:

- Tipo de construcción:** Includes checkboxes for 'Segunda mano' and 'Obra nueva'.
- Precio:** A range slider with 'Mínimo' set to 'Indiferente' and 'Máximo' set to '1.000 €'.
- Tipo de vivienda:** A dropdown menu with 'Todos los pisos' selected, and checkboxes for 'Plantas intermedias' (which is checked) and 'Apartamento'.
- Filtros:** A grid of icons and labels: Terraza (Terrace), Balcón (Balcony), Calefacción (Heating), Trastero (Storage), Jardín (Garden), No amueblado (Unfurnished), Aire acondicionado (Air conditioning), and Electrodomésticos (Household appliances).
- Fecha de publicación:** A section with radio buttons for 'Indiferente', 'Últimas 48 horas', and 'Última semana' (which is selected). Below it is a button labeled 'Mostrar 438 anuncios'.

```
1 # 1. Hacer clic en el botón de Filtros
2 filtros_button = driver.find_element(By.XPATH, '//*[contains(@class, "re-SearchFiltersTop-filtersButton")]')
3 filtros_button.click()
4 time.sleep(2) # Esperar a que se abra el menú de filtros
5
6 # 2. Encontrar y hacer clic en el selector de precio máximo
7 precio_max_box = driver.find_element(By.XPATH, '//div[contains(@class, "re-FiltersFilterPrice-selectors-rightSelector")]/div[contains(@class, "sui-MoleculeSelect")]')
8 precio_max_box.click()
9 time.sleep(2) # Esperar para que la lista se despliegue
10
11 # 3. Establecer precio max =1000
12 precio_max_1000 = driver.find_element(By.XPATH, '//*[@id="FiltersModal"]/div/div/div[2]/div/div[3]/div[2]/div[2]/ul/li[20]')
13 precio_max_1000.click()
14 time.sleep(2)
15
16 # 4. Establecer tipo de viviendas como "Plantas intermedias"
17 type_house = driver.find_element(By.XPATH, '//div[contains(@class, "re-FiltersFilterPropertySubtype")]/label[contains(@for, "1")]')
18 type_house.click()
19 time.sleep(2)
20
21 # 5. Establecer fecha de publicación como "Última semana"
22 date_publication = driver.find_element(By.XPATH, '//div[contains(@class, "re-FilterFilterPublicationDate")]/label[contains(@for, "ONE_WEEK")]')
23 date_publication.click()
24
25 time.sleep(2)
26
27 # 6. Clicar en el botón "Mostrar los anuncios resultantes"
28 boton_mostrar_anuncios = driver.find_element(By.XPATH, '//button[contains(@class, "re-FiltersButton re-FiltersButton--isFiltersModal")]')
29 boton_mostrar_anuncios.click()
```

2.1 Configurar parámetros iniciales

Paso 6: Manejo de Pop-ups



¿Quieres ser el primero en enterarte de las novedades?

Guárdate una alerta de esta búsqueda:

- ✓ Madrid Capital
- ✓ Alquiler, Viviendas
- ✓ Plantas intermedias
- ✓ Hasta 1.000€

No, gracias

Sí, crear alerta

```
1 try:
2     # Verificar si aparece la ventana emergente de "Crear Alerta"
3     time.sleep(4) # Esperar un poco para que la ventana cargue si aparece
4     ventana_alerta = driver.find_element(By.XPATH, '//*[@id="modal-react-portal"]/div/div')
5
6     if ventana_alerta:
7         boton_no_gracias = driver.find_element(By.XPATH, '//*[@id="modal-react-portal"]/div/div/div[2]/div/button[2]')
8         boton_no_gracias.click()
9         print("Ventana de 'Crear Alerta' detectada y cerrada con 'No, gracias'.")
10        time.sleep(1) # Esperar para que se cierre correctamente
11    except NoSuchElementException:
12        print("No apareció la ventana de 'Crear Alerta', continuando con el proceso.")
13
14 [42]
15
16 Ventana de 'Crear Alerta' detectada y cerrada con 'No, gracias'.
```

The screenshot shows a search results page for 'Plantas intermedias de alquiler en Madrid Capital'. It displays 52 second-hand units. A listing for a 41m² apartment in Calle Valverde is highlighted, priced at 990 €. The page also features a map with a search overlay and a sidebar with alert configuration options.

Filtros (3)

Distrito

hasta 1.000 €

Habitaciones

Crear alerta

600.000 personas buscan un inmueble como el tuyo

Publicar tu anuncio GRATIS

Plantas intermedias de alquiler en Madrid Capital

52 de segunda mano

Ordenar: Puntuación Fotocasa

Calidad Fotocasa · AGENCIA ...

990 €

Hace 5 días

Piso en Calle Valverde, Universidad -...

1 hab. 1 baño 41 m²

Expandir

Dibujar en mapa

Buscar por trayecto

Vista satélite

N-110

Buitrago del Lozoya

Colmenar Viejo

Tres Cantos

San Sebastián de los Reyes

Alcalá de Henares

Las Rozas de Madrid

Boadilla del Monte

Móstoles

Getafe

Arganda del Rey

2.2 Retos y soluciones

- Lazy Loading: Carga diferida
- Banner de anuncio

```
1 def scroll_page(driver, step=100, pause_time=0.2):  
2     """Hace scroll para cargar todas las ofertas visibles en la página."""  
3     last_position = driver.execute_script("return window.pageYOffset;")  
4     while True:  
5         driver.execute_script("window.scrollBy(0, arguments[0]);", step)  
6         time.sleep(pause_time)  
7         new_position = driver.execute_script("return window.pageYOffset;")  
8         if new_position == last_position: # Si no cambia la posición, hemos llegado al final  
9             break  
10        last_position = new_position  
11    time.sleep(2) # Espera extra para que cargue el contenido
```

Filtros (3) Distrito hasta 1.000 € Habitaciones Crear alerta

article.re-cardpackminimal 937.47 x 263.99

970 € Piso en ANDRES TORREJON, Jerónimos 1 hab. · 1 baño · 45 m² · Ascensor · Calefacción Redpiso Niño Jesús Jerónimos pone a su disposición este apartamento frente al Retiro. Sus 45 m² se distribuyen en : hall de entrada, amplio salón, cocina , baño... Leer más

Contactar Llamar

950 € Piso en Pueblo Nuevo 1 hab. · 1 baño · 38 m² · Bajos · Calefacción

Hace 3 días

1/17

URBAMAR Hace 4 días

DevTools is now available in Spanish!

Always match Chrome's language Switch DevTools to Spanish Don't show again

Elements Console Sources Network > 15 ▲ 11 326

...

> <div class="re-SearchResult-header"> ... </div> flex

><section class="re-SearchResult"> grid

><article class="re-CardPackMinimal"> ... </article> flex

><div class="re-SearchResult-advertisingWrapper re-SearchResult-advertisingReserveSpace--banner"> ... </div> grid

><div class="sui-PerfDynamicRendering-placeholder"> ... </div>

><div class="re-SearchResult-advertisingWrapper re-SearchResult-advertisingReserveSpace--native"> ... </div> grid

><div class="sui-PerfDynamicRendering-placeholder"> ... </div>

><div class="re-SearchResult-advertisingWrapper re-SearchResult-advertisingReserveSpace--native"> ... </div> grid

><div class="sui-PerfDynamicRendering-placeholder"> ... </div>

2.2 Retos y soluciones

- Falta de ID individual extraible
- ID presente en el link individual: “href”

```
13 def extract_articles(driver):  
14     """ Extrae los datos de los artículos presentes en la página de Fotocasa. """  
15     soup = BeautifulSoup(driver.page_source, 'lxml')  
16     articles = soup.find("section", class_="re-SearchResult").find_all("article")  
17  
18     data = []  
19     for article in articles:  
20         a_tag = article.find("a", href=True)
```

The screenshot shows a search results page for apartments in Madrid. At the top, there are filters for 'Filtros (3)', 'Distrito' (set to 'Chamartín'), 'Habitaciones' (set to '1'), and price range 'hasta 1.000 €'. A 'Crear alerta' button is also visible. Below the filters, the results are ordered by 'Puntuación Fotocasa'. The first listing is for a 1-bedroom apartment in 'ANDRES TORREJON, Jerónimos' at 970 €, featuring a purple sofa and a white coffee table. The second listing is for a studio in 'Pueblo Nuevo' at 950 €, featuring a blue sofa and a desk. Both listings include contact and call buttons.

The screenshot shows the browser's developer tools with the 'Elements' tab selected. The page is a search results page for 'Fotocasa'. The DOM tree is displayed, highlighting the HTML structure of the search results. Two specific links are highlighted with green and blue boxes: one for a 1-bedroom apartment in 'ANDRES TORREJON, Jerónimos' and another for a studio in 'Pueblo Nuevo'. These highlighted links correspond to the ones shown in the screenshot above. The bottom of the developer tools shows various tabs like 'Styles', 'Computed', 'Layout', etc.

2.2 Retos y soluciones

- Ejemplo de página web similares: Idealista

Nuevos anuncios en tu email

Guardar búsqueda

Tipo de alquiler

Residencial de larga estancia

De temporada

Tamaño

Mín Máx

Tipo de vivienda

Pisos, áticos y dúplex

Pisos

Áticos

Dúplex

Casas y chalets

Otras denominaciones

Apartamentos

Piso en calle del Marqués de Santa Ana, Malasaña-Universidad, Madrid
970 €/mes
1 hab. 40 m² Bajo exterior con ascensor 7 horas
Contactar SÓLO POR MENSAJE enviando datos con perfil laboral, ingresos, etc. Nos pondremos en...
Apartamento
Contactar Ver teléfono

Piso en calle de Quiñones, 5, Malasaña-Universidad,...
NIK INVERSIÓN Y GESTIÓN
980 €/mes
1 hab. 32 m² Bajo interior 14 mar
Nik Inversión y Gestión, ofrece en alquiler este estudio de 32 m², recién reformado en pleno centr...
Contactar Ver teléfono

Estudio en calle de Amaniel, Malasaña-Universidad, Madrid
850 €/mes

DevTools is now available in Spanish!
Always match Chrome's language Switch DevTools to Spanish
Don't show again

Elements Console Sources Network ▲ 5 □ 5 ⚙

```
<div id="listing-aside">...</div>
<a id="map-view-button" href="/alquiler-viviendas/madrid/centro/malasaña-universidad/con-precio-hasta_1000.pisos,publicado_ultima-semana/ma...</a>
data-element-id="10760691" data-online-booking="false"></a>
<div id="order-by" class="order-by" data-param="ordenado-por">...</div>
<div id="favorites-lists-root"></div>
<main class="listing-items core-vitals-listing-map" id="main-content">...</main>
<section class="items-container items-list">
  <section id="auto-disclaimer" data-disclaimer-id="RENTINGBDMI">...</section>
  <article class="item_shared-flat">...</article>
  <article class="item extended-item item-multimedia-container" data-element-id="107636570" data-online-booking="false">...</article>
  <article class="item extended-item item-multimedia-container" data-element-id="107663966" data-online-booking="false">...</article>
  <article class="adv noHover">...</article>
  <article class="item item_contains_branding extended-item item-multimedia-container" data-element-id="104619822" data-online-booking="false">...</article>
  <article class="item extended-item item-multimedia-container" data-element-id="34313517" data-online-booking="false">...</article>
  <article class="save-search-card-module">...</article>
  <article class="item extended-item item-multimedia-container" data-element-id="107649378" data-online-booking="false">...</article>
```

2.2 Retos y soluciones

```
13 def extract_articles(driver):
14     """ Extrae los datos de los artículos presentes en la página de Fotocasa. """
15     soup = BeautifulSoup(driver.page_source, 'lxml')
16     articles = soup.find("section", class_="re-SearchResult").find_all("article")
17
18     data = []
19     for article in articles:
20         a_tag = article.find("a", href=True)
21
22         # Extraer título
23         title = a_tag.get("title", "").strip() if a_tag else None
24         # Extraer link de la oferta
25         href = f"https://www.fotocasa.es{a_tag['href']}" if a_tag else None
26
27         # Extracción y limpieza del precio
28         price_tag = article.find("span", class_="re-CardPrice")
29         if price_tag:
30             price_text = price_tag.text.replace(".", "").replace(",", "").strip().split("€")[0]
31             try:
32                 price = int(price_text)
33             except ValueError:
34                 price = np.nan
35         else:
36             price = np.nan
37
38         # Extraer localización a partir del título
39         location = title.split(" en ")[-1] if title else None
```

A screenshot of a real estate listing on the Fotocasa website. At the top, there are filters for 'Filtros (3)', 'Distrito' (set to 'hasta 1.000 €'), and a price slider set to '970 €'. Below the filters is a blurred image of an apartment interior. To the right of the image, the price '970 €' is circled in red. Below the price, the location 'Piso en ANDRES TORREJON, Jerónimos' is highlighted with a green box. Further down, the text '1 hab. · 1 baño · 45 m² · Ascensor · Calefacción' is visible. At the bottom, there are buttons for 'Contactar' and 'Llamar'.

```
<main class="re-searchPage-wrapper"> </main>
::before
<div class="re-SearchPage-container-and-banner-wrapper">
    <div class="re-SearchPage-container">
        <div class="re-SearchDominance">...</div>
        <div class="re-SearchPage-title-section">...</div> <flex>
            <div class="re-SearchResult-wrapper">
                <div class="re-SearchResult-header">...</div> <flex>
                    <section class="re-SearchResult"> <grid>
                        ...
                            <article class="re-CardPackMinimal"> <flex> == $0
                                <a title="Piso de alquiler en ANDRES TORREJON, Jerónimos" href="/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-amueblado/185941379/d?from=list">...</a>
                                    <div class="re-CardPackMinimal-info">...</div> <flex>
                                </article>
                            <article class="re-CardPackMinimal">...</article> <flex>
                            <article class="re-CardPackMinimal">...</article> <flex>
                            <article class="re-CardPackMinimal">...</article> <flex>
                        ...
                    </section>
                </div>
            </div>
        </div>
    </div>
</div>
```

2.2 Retos y soluciones

```
40      # Extracción y limpieza del resto de características
41      features = article.find_all("span", class_="re-CardFeaturesWithIcons-feature-icon") or \
42                  article.find_all("li", class_="re-CardFeatures-item re-CardFeatures-feature")
43
44      # Tamaño en m2
45      try:
46          size = next(int(f.text.split()[0]) for f in features if "m2" in f.text)
47      except StopIteration:
48          size = np.nan
49
50      # Número de habitaciones
51      try:
52          room = next(int(f.text.split()[0]) for f in features if "hab" in f.text)
53      except StopIteration:
54          room = np.nan
55
56      # Número de baños
57      try:
58          bathroom = next(int(f.text.split()[0]) for f in features if "baño" in f.text)
59      except StopIteration:
60          bathroom = np.nan
61
62      # Calefacción
63      heating = next((f.text.strip() for f in features if "Calefacción" in f.text), "")
64
65      # Aire acondicionado
66      air_conditioning = next((f.text.strip() for f in features if "Aire acondicionado" in f.text), "")
67
68      # Planta
69      floor = next((f.text.strip() for f in features if "Planta" in f.text or "Bajos" in f.text), "")
70
71      data.append({
72          "title": title,
73          "location": location,
74          "price": price,
75          "size_m2": size,
76          "rooms": room,
77          "bathrooms": bathroom,
78          "heating": heating,
79          "air_conditioning": air_conditioning,
80          "floor": floor,
81          "href": href
82      })
83
84      return pd.DataFrame(data)
85
86
```

- Contenido presente en diferentes clases

The screenshot displays two real estate listings from 'Viancasa'. Each listing includes a summary, contact buttons, and a detailed description with its corresponding HTML structure.

Top Listing:

- Summary:** 68.25 x 23.99 m², 1 habitación, 1 baño, 50 m².
- Contact:** Contactar, Llamar.
- Description:** Viancasa Grupo Inmobiliario Ofrece en alquiler piso recien ... [Leer más](#)
- HTML Structure:** ``

Bottom Listing:

- Summary:** 59.93 x 20.8 m², 1 habitación, 1 baño, 45 m², Ascensor, Calefacción.
- Contact:** Contactar, Llamar.
- Description:** Redpiso Niño Jesús Jerónimos pone a su disposición este apartamento frente al Retiro. Sus 45 m² se distribuyen en : hall de entrada, amplio salón, cocina , baño... [Leer más](#)
- HTML Structure:** `<li class="re-CardFeaturesWithIcons-feature"> == $0 ::before "1 hab." `

2.3. Extracción de datos

Extracción de datos de la primera página

```
1 scroll_page(driver)
[44]

1 pd.set_option('display.max_rows', None)
2 pd.set_option('display.max_columns', None)
3 pd.set_option('display.max_colwidth', None)
4
5 df_1 = extract_articles(driver)
6 print(df_1)
[45]

23      55    2    1 Calefacción
24      40    1    1 Calefacción Aire acondicionado
25      50    1    1
26      45    1    1 Calefacción
27      47    1    1
28      35    1    1 Calefacción
29      40    1    1 Calefacción

0
1
2
3
4
5
6
7
8
9
10
11
```

href
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-amueblado/185941379/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion/185926186/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-no-amueblado/185923603/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-ascensor-amueblado/185894818/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-parking-jardin-terraza-zona-comunitaria-ascensor/185945023/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-terraza-ascensor/185945012/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/terraza-amueblado/185909313/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-parking-ascensor-piscina-amueblado/185952846/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-no-amueblado/185947223/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-terraza-ascensor-amueblado/185930534/d?from=list>
<https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-amueblado-internet/185908568/d?from=list>

Una vez extraído la información, se desplaza a la siguiente página

A partir de las funciones anteriores se procede a extraer los datos de la página 1

```
1 # Para poder ejecutar este comando, se requiere desplazar el mouse hasta la sección de pasar página de forma manual, dejándolo
2 visible en la pantalla
3 next_button = driver.find_element(By.XPATH,
4                                     '//a[contains(@class,"sui-AtomButton sui-AtomButton--primary sui-AtomButton--outline
5                                     sui-AtomButton--center sui-AtomButton--small sui-AtomButton--link sui-AtomButton--empty
6                                     sui-AtomButton--rounded")]')
7
8 next_button.click()
9 time.sleep(3)
10
11 [46]
```



995 €

Piso en de Hermosilla, Goya

2 habs. · 1 baño · 65 m² · Ascensor · Calefacción

Vivienda recién pintada de 65m² distribuida en 2 habitaciones, amplio salón, cocina amueblada y equipada...

Contactar

Llamar



...

2.3. Extracción de datos

Extracción de datos de la segunda página

```
1 scroll_page(driver)
[47]

1 df_2 = extract_articles(driver)
2 print(df_2)
[48]

43 1.0 1.0
44 2.0 1.0
45 2.0 1.0
Aire acondicionado 1* Planta

          href
0 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-amueblado/185923594/d?from=list
1 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor/185925652/d?from=list
2 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/patio/185827524/d?from=list
3 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-terraza-ascensor-amueblado-television-piscina/185891451/d?from=list
4 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-amueblado/185896846/d?from=list
5 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-ascensor-no-amueblado/185911529/d?from=list
6 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/ascensor-amueblado/185339150/d?from=list
7 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/pueblo-nuevo/185950153/d?from=list
8 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/terraza/185573908/d?from=list
9 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/jardin-ascensor-piscina-amueblado/185929900/d?from=list
10 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-internet/185950951/d?from=list
11 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/jardin-terraza-ascensor-amueblado/185944148/d?from=list
12 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-amueblado/185954235/d?from=list
13 https://www.fotocasa.es/es/
14 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/
15 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado
```

Extracción de datos completo en excel

```
1 df_total = pd.concat([df_1, df_2], ignore_index=True)
2 print(df_total)
[49]

26 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-ascensor-amueblado/185931068/d?from=list
27 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-parking-trastero-ascensor-piscina-amueblado/185902392/d?from=list
28 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-amueblado/185943293/d?from=list
29 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-no-amueblado/185919467/d?from=list
30 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-amueblado/185923594/d?from=list
31 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor/185925652/d?from=list
32 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/patio/185827524/d?from=list
33 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-terraza-ascensor-amueblado-television-piscina/185891451/d?from=list
34 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-amueblado/185896846/d?from=list
35 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-ascensor-no-amueblado/185911529/d?from=list
36 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/ascensor-amueblado/185339150/d?from=list
37 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/pueblo-nuevo/185950153/d?from=list
38 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/terraza/185573908/d?from=list
39 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/jardin-ascensor-piscina-amueblado/185929900/d?from=list
40 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-internet/185950951/d?from=list
41 https://www.Fotocasa.es/es/alquiler/vivienda/madrid-capital/jardin-terraza-ascensor-amueblado/185944148/d?from=list
42 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/calefaccion-ascensor-amueblado/185954235/d?from=list
43 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/ascensor-amueblado/185892116/d?from=list
44 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-ascensor-amueblado/185921010/d?from=list
45 https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-ascensor-amueblado-television-internet/185911939/d?from=list
46 https://www.fotocasa.es/es/alquiler/vivienda/pozuelo-de-alarcon/aire-acondicionado-calefaccion-amueblado-internet/183295513/d?from=list
```

Se extrae los datos de la página 2

Y se concatena ambas “DataFrames” en una para el trabajo posterior

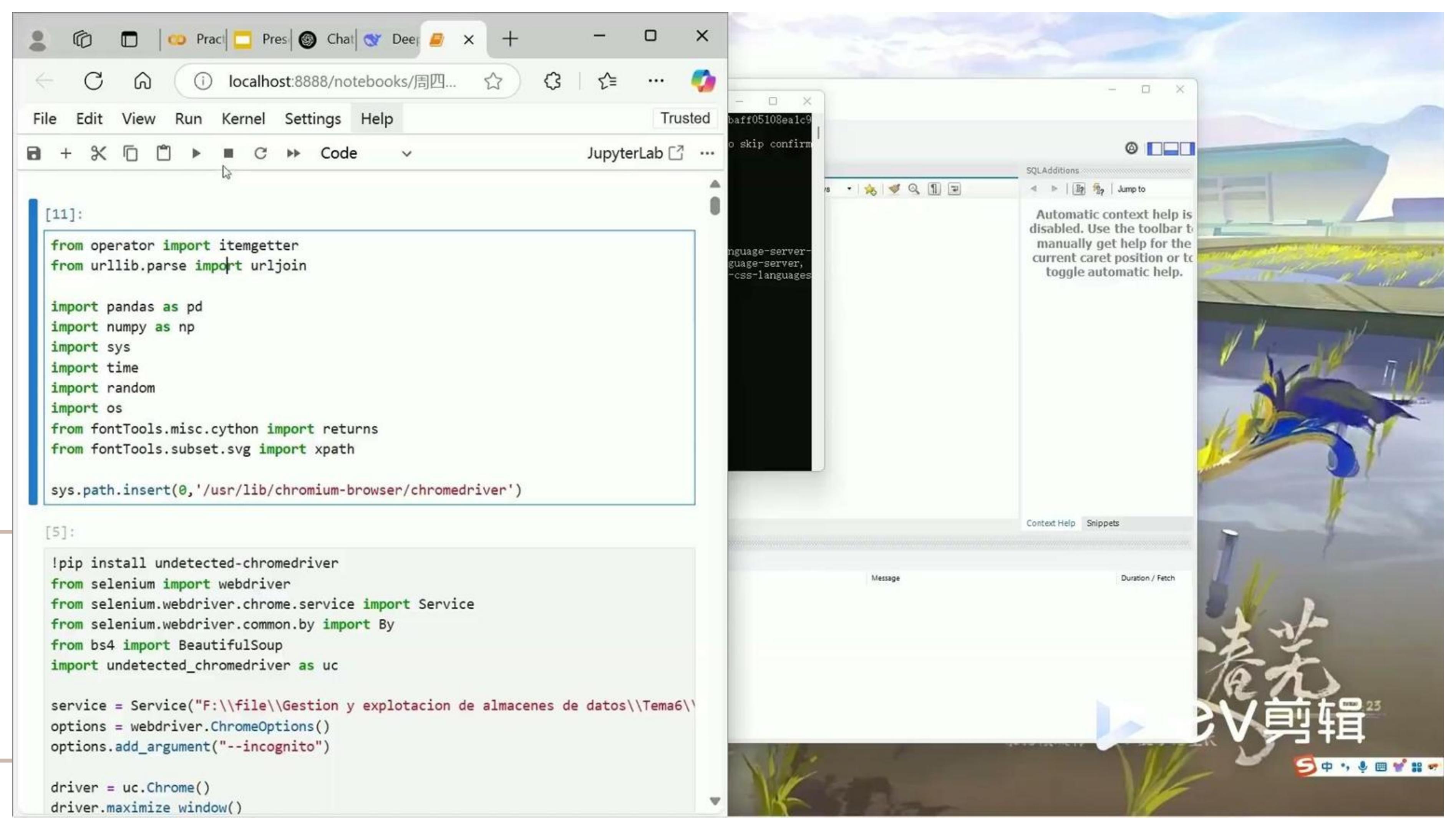
2.3 Extracción de datos

- Guardar el dataframe en excel (*.xlsx)

```
1 # Establecer el directorio de descarga
2 download_directory = "D:\\XinYuan Zheng\\OneDrive\\UCM\\2º CUATRIMESTRE\\Gestión y Explotación de Almacenes de Datos\\Tema 6\\Practica_1" # Cambia esta ruta a la que prefieras
3
4 # Ruta completa del archivo
5 file_path = os.path.join(download_directory, "viviendas_fotocasa.xlsx")
6
7 # Guardar como Excel
8 df_total.to_excel(file_path, index=False, engine='openpyxl')
[50]
```

 Practica_1		19/03/2025 3:00	Archivo IPYNB	188 KB
 viviendas_fotocasa		19/03/2025 3:18	Hoja de cálculo de Microsoft Excel	14 KB

1	title	location	price	size_m2	rooms	bathrooms	heating	air_conditioning	floor	href
2	Piso de alquiler en ANDRES TORREJON, Jerónimos	ANDRES TORREJON, Jerónimos	970	45	1	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000000
3	Piso de alquiler en Pueblo Nuevo	Pueblo Nuevo	950	38	1	1	Calefacción		Bajos	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000001
4	Piso de alquiler en San Eudaldo, Ambroz	San Eudaldo, Ambroz	1000	63	3	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000002
5	Piso de alquiler en de Méndez Álvaro, Palos de Moguer	de Méndez Álvaro, Palos de Moguer	950	35	1	1		Aire acondicionado		https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000003
6	Piso de alquiler en San Diego	San Diego	850	59	1	1			1º Planta	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000004
7	Piso de alquiler en Avenida de manoteras, Virgen del Cc	Avenida de manoteras, Virgen del Cortijo	900	56	1	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000005
8	Piso de alquiler en Santa Susana, Pinar del Rey	Santa Susana, Pinar del Rey	850	61	1	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000006
9	Piso de alquiler en de Virgen de Lluc, Quintana	de Virgen de Lluc, Quintana	850	42	1	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000007
10	Piso de alquiler en Pablo Neruda, Palomeras Bajas	Pablo Neruda, Palomeras Bajas	995	64	2	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000008
11	Piso de alquiler en Pico de los Artilleros, Fontarrón	Pico de los Artilleros, Fontarrón	990	65	3	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000009
12	Piso de alquiler en Salado, Orcasur - 12 de Octubre	Salado, Orcasur - 12 de Octubre	950	70	3	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000010
13	Piso de alquiler en Calle del Espejo, 8, Palacio	Calle del Espejo, 8, Palacio	850	25	1	1			Bajos	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000011
14	Piso de alquiler en Calle Laguna del Marquesado, 3, Villa	Calle Laguna del Marquesado, 3, Villaver	990	72	1	2				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000012
15	Piso de alquiler en Calle RODRIGUEZ ESPINOSA, Palome	Calle RODRIGUEZ ESPINOSA, Palomeras	975	45	1	1			Bajos	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000013
16	Piso de alquiler en Calle Embalse de Navacerrada, 47, E	Calle Embalse de Navacerrada, 47, Ensa	895	50	1	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000014
17	Piso de alquiler en Calle San Eustaquio, 26, Villaverde Al	Calle San Eustaquio, 26, Villaverde Alto	800	32	1	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000015
18	Piso de alquiler en Calle de Fernando el Católico, 28, Ar	Calle de Fernando el Católico, 28, Arapil	980	40	1	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000016
19	Piso de alquiler en Calle Valverde, Universidad - Malasa	Calle Valverde, Universidad - Malasaña	990	41	1	1			3º Planta	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000017
20	Piso de alquiler en C. de Sebastián Elcano, Palos de Mo	C. de Sebastián Elcano, Palos de Moguer	1000	40	1	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000018
21	Piso de alquiler en Camino de Perales, San Fermín	Camino de Perales, San Fermín	950	64	3	1			1º Planta	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000019
22	Piso de alquiler en Calle Alfareros, Arcos	Calle Alfareros, Arcos	825	56	3	1			Bajos	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000020
23	Piso de alquiler en Calle JAIME VERA, Puerta del Ángel	Calle JAIME VERA, Puerta del Ángel	950	84	3	1			1º Planta	https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000021
24	Piso de alquiler en Puerto de Alazores, Casco Histórico	Puerto de Alazores, Casco Histórico de V	900	65	2	1				https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000022
25	Piso de alquiler en de La Seo de Urgel, Pinar del Rey	de La Seo de Urgel, Pinar del Rey	900	55	2	1	Calefacción			https://www.fotocasa.es/es/alquiler/vivienda/1000000000000000023



3. BBDD - MYSQL



Navigator

SCHEMAS

Filter objects

- db_fotocasa
 - Tables
 - ofertas_inmobiliarias
 - Columns
 - id
 - title
 - location
 - price
 - size_m2
 - rooms
 - bathrooms
 - heating
 - air_conditioning
 - floor
 - href
 - Indexes
 - Foreign Keys
 - Triggers
 - Views
 - Stored Procedures
 - Functions

Query 1

File Edit View Insert Tools Help

1

```

1 import mysql.connector
2 import pandas as pd
3
4 # Conexión a MySQL
5 conexión = mysql.connector.connect(
6     host="localhost",
7     user="root",
8     password="20010714"
9 )
10
11 cursor = conexión.cursor()
[53]

1 cursor.execute("CREATE DATABASE IF NOT EXISTS db_fotocasa;") # Crear la base de datos
2 cursor.execute("USE db_fotocasa;")
[54]

1 # Crear la tabla en my_sql
2 cursor.execute("""
3 CREATE TABLE IF NOT EXISTS ofertas_inmobiliarias (
4     id INT AUTO_INCREMENT PRIMARY KEY,
5     title VARCHAR(255),
6     location VARCHAR(255),
7     price INT,
8     size_m2 INT,
9     rooms INT,
10    bathrooms INT,
11    heating VARCHAR(255),
12    air_conditioning VARCHAR(255),
13    floor VARCHAR(255),
14    href TEXT
15 );
16""")
17 conexión.commit() # Guardar cambios
[55]

```

Creación del base de datos en MySQL a partir del archivo excel creado

```

1 insert_query = """
2 INSERT INTO ofertas_inmobiliarias
3 (title, location, price, size_m2, rooms, bathrooms, heating, air_conditioning, floor, href)
4 VALUES (%s, %s, %s, %s, %s, %s, %s, %s, %s, %s);
5 """
[56]

1 data = [
2     (
3         str(row["title"]) if row["title"] else None,
4         str(row["location"]) if row["location"] else None,
5         int(row["price"]) if row["price"] else None,
6         int(row["size_m2"]) if pd.notna(row["size_m2"]) else None,
7         int(row["rooms"]) if pd.notna(row["rooms"]) else None,
8         int(row["bathrooms"]) if pd.notna(row["bathrooms"]) else None,
9         str(row["heating"]) if row["heating"] else None,
10        str(row["air_conditioning"]) if row["air_conditioning"] else None,
11        str(row["floor"]) if row["floor"] else None,
12        str(row["href"]) if row["href"] else None
13     )
14     for _, row in df.iterrows()
15 ]
[57]

1 cursor.executemany(insert_query, data)
2 conexión.commit() # Guardar cambios
[58]

1 cursor.close()
2 conexión.close()
[59]

```

4. Visualización de Datos



4.1 Importar datos

```
1 import pandas as pd
2 from sqlalchemy import create_engine
3
4 # Conexión usando SQLAlchemy
5 engine = create_engine('mysql+pymysql://root:20010714@localhost/db_fotocasa')
6
7 # Leer los datos usando la conexión de SQLAlchemy
8 query = "SELECT * FROM ofertas_inmobiliarias"
9 df = pd.read_sql(query, engine)
10
11 # Cerrar la conexión
12 engine.dispose()
13
14 # Verificar los datos
15 print(df.head())
16
```

Importar el conjunto de datos desde MySQL

	location	price	size_m2	rooms	bathrooms	\
0	ANDRES TORREJON, Jerónimos	970	45.0	1.0	1.0	
1	Pueblo Nuevo	950	38.0	1.0	1.0	
2	San Eudaldo, Ambroz	1000	63.0	3.0	1.0	
3	de Méndez Álvaro, Palos de Moguer	950	35.0	1.0	1.0	
4	San Diego	850	59.0	1.0	1.0	

	heating	air_conditioning	floor	\
0	Calefacción	None	None	
1	Calefacción	None	Bajos	
2	Calefacción	None	None	
3	None	Aire acondicionado	None	
4	None	None	1º Planta	

	href
0	https://www.fotocasa.es/es/alquiler/vivienda/m...
1	https://www.fotocasa.es/es/alquiler/vivienda/m...
2	https://www.fotocasa.es/es/alquiler/vivienda/m...
3	https://www.fotocasa.es/es/alquiler/vivienda/m...
4	https://www.fotocasa.es/es/alquiler/vivienda/m...

4.2 Muestra en tablas

- Parámetros estadísticos básicos (var.)

numéricas)

```
1 print("* Número total de ofertas:", len(df))
2 print("\n* Descripción de los datos numéricos:") # Datos estadísticos básicos
3 print(df.describe())
4
[3]
* Número total de ofertas: 76
* Descripción de los datos numéricos:
   id      price    size_m2   rooms  bathrooms
count  76.000000  76.000000  75.000000  67.000000  67.000000
mean   38.500000  897.881579  46.906667  1.373134  1.014925
std    22.083176  91.094086  16.995368  0.692729  0.122169
min    1.000000  450.000000  1.000000  1.000000  1.000000
25%   19.750000  850.000000  37.000000  1.000000  1.000000
50%   38.500000  900.000000  45.000000  1.000000  1.000000
75%   57.250000  975.000000  57.500000  1.500000  1.000000
max   76.000000  1000.000000 115.000000  3.000000  2.000000
```

- Frecuencia de la localización de la oferta

- Distribución del precio ascendente

```
1 df_sorted_price = df.sort_values(by="price", ascending=True).head(10)
2 print(df_sorted_price) # Ordenar por precios ascendentes
3
[5]
   id      title \
41  42  Piso de alquiler en Aluche
24  25  Piso de alquiler en de José del Pino Jiménez, ...
42  43  Piso de alquiler en San Valero, Numancia
61  62  Piso de alquiler en Cortes - Huertas
32  33  Piso de alquiler en Plaza de la Cebada, 1, Pal...
47  48  Piso de alquiler en Valdeacederas
48  49  Piso de alquiler en Canillejas
15  16  Piso de alquiler en Calle San Eustaquio, 26, V...
35  36  Piso de alquiler en Fuenlabrada, Opañel
37  38  Piso de alquiler en Pueblo Nuevo
               location  price  size_m2  rooms \
41          Aluche     450    115.0    3.0
24  de José del Pino Jiménez, Villaverde Alto     750    40.0    1.0
42          San Valero, Numancia     750    60.0    2.0
61          Cortes - Huertas     750    30.0    NaN
32          Plaza de la Cebada, 1, Palacio     760    25.0    1.0
47          Valdeacederas     780    33.0    NaN
48          Canillejas     799    25.0    NaN
15  Calle San Eustaquio, 26, Villaverde Alto     800    32.0    1.0
```

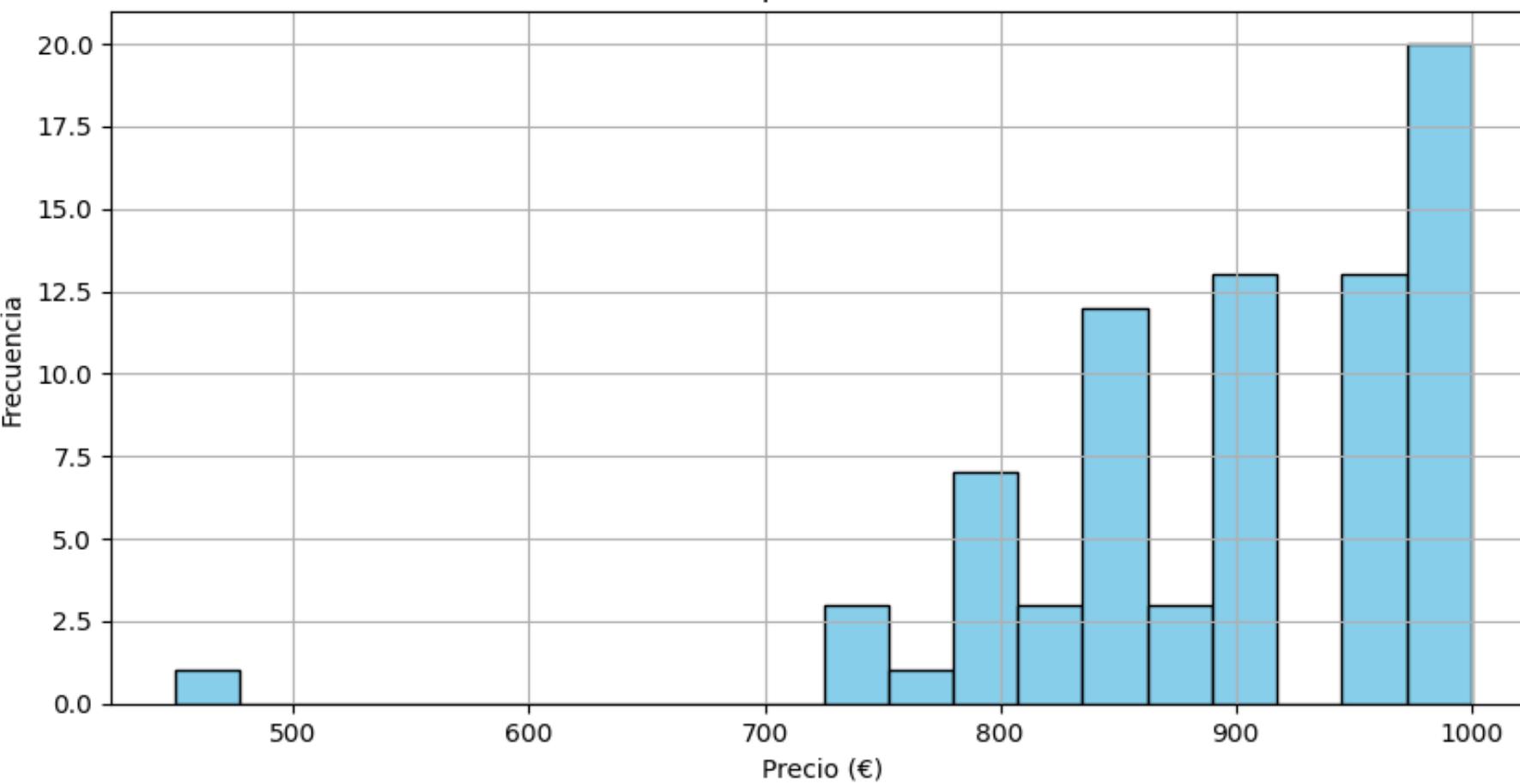
```
1 print("\n* Ubicaciones más populares:")
2 print(df["location"].value_counts().head(10)) # Top 10 zonas con más viviendas
3
```

[4]

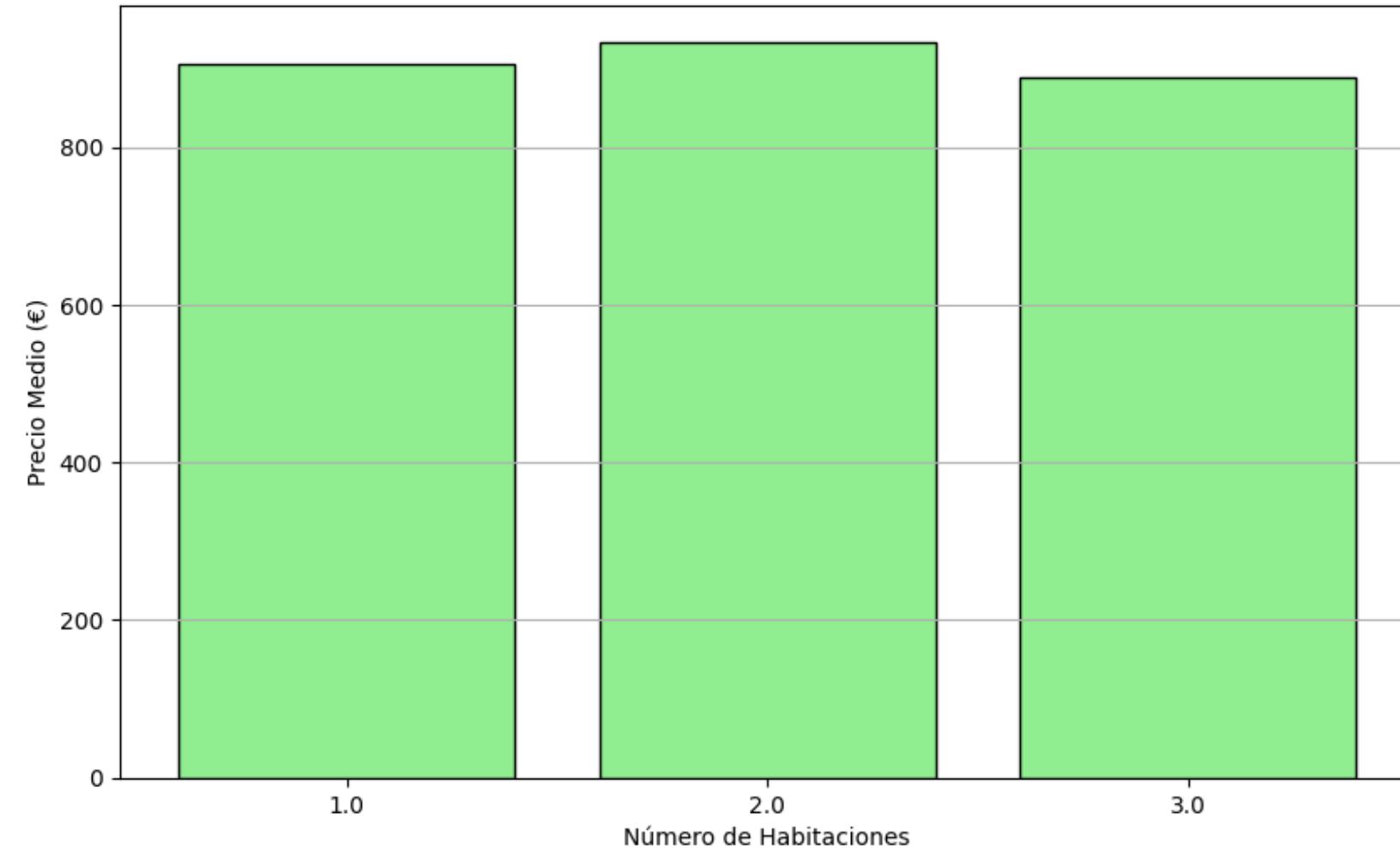
```
* Ubicaciones más populares:
location
Pueblo Nuevo
ANDRES TORREJON, Jerónimos
San Eudaldo, Ambroz
de Méndez Álvaro, Palos de Moguer
San Diego
Avenida de manoteras, Virgen del Cortijo - Manoteras
Santa Susana, Pinar del Rey
de Virgen de Lluc, Quintana
Pablo Neruda, Palomeras Bajas
N/A, Pueblo Nuevo
Name: count, dtype: int64
```

4.3 Muestra en gráficas

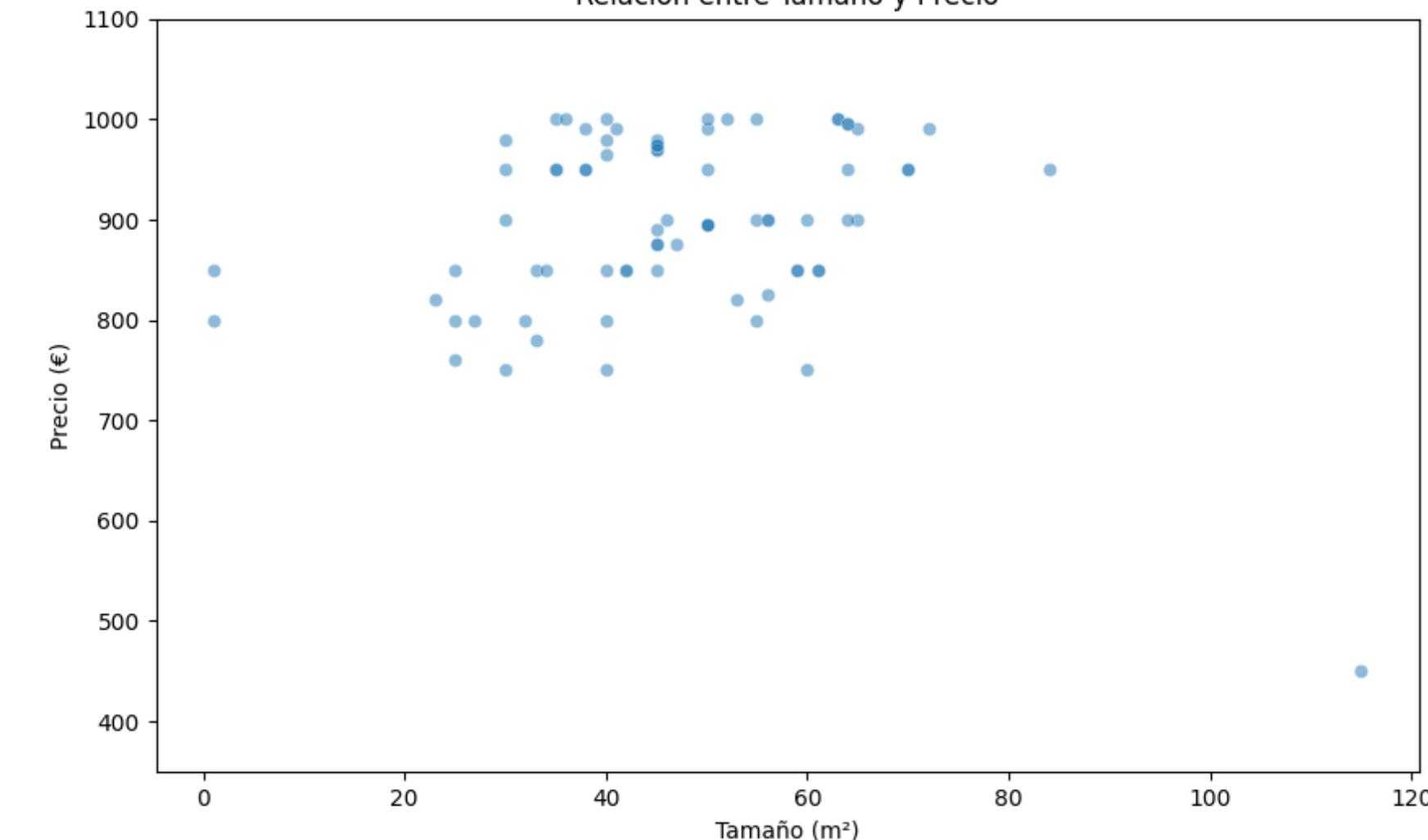
Distribución de precios de las viviendas



Precio Medio según Número de Habitaciones

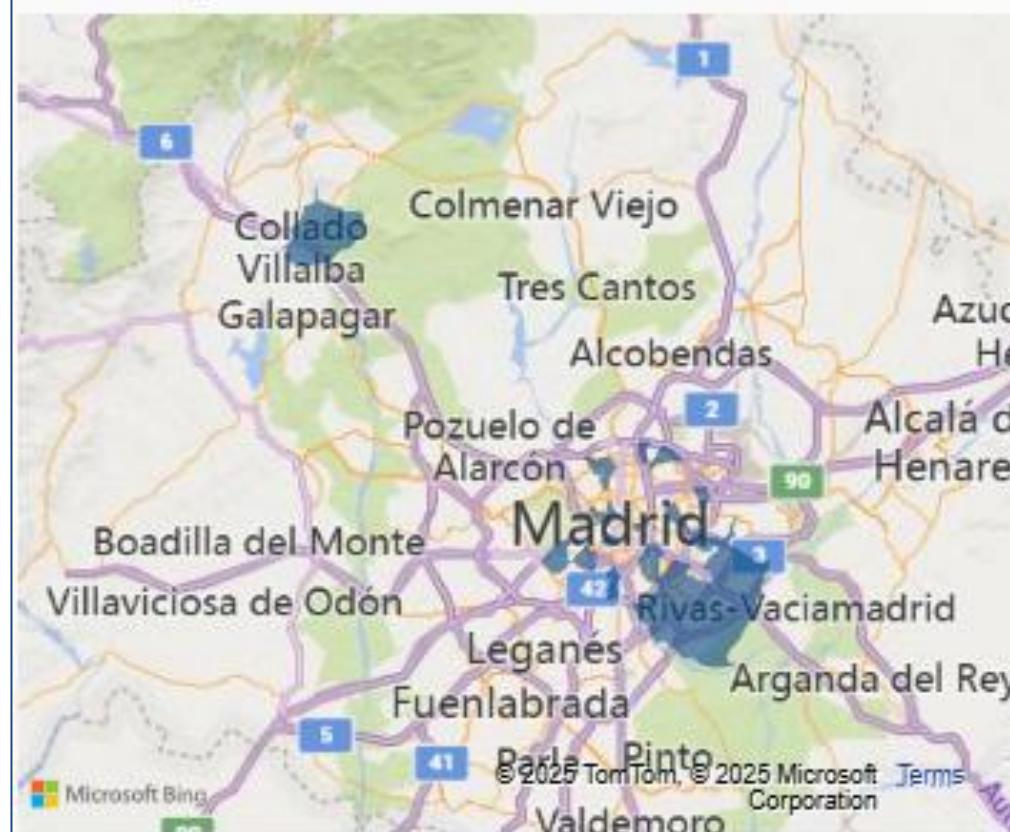


Relación entre Tamaño y Precio



4.4 Panel de control

location_full



- floor
- 1º Planta
 - 2º Planta
 - 3º Planta
 - 4º Planta
 - 5º Planta
 - 6º Planta
 - 7º Planta
 - Bajos

price



heating

 Calefacción

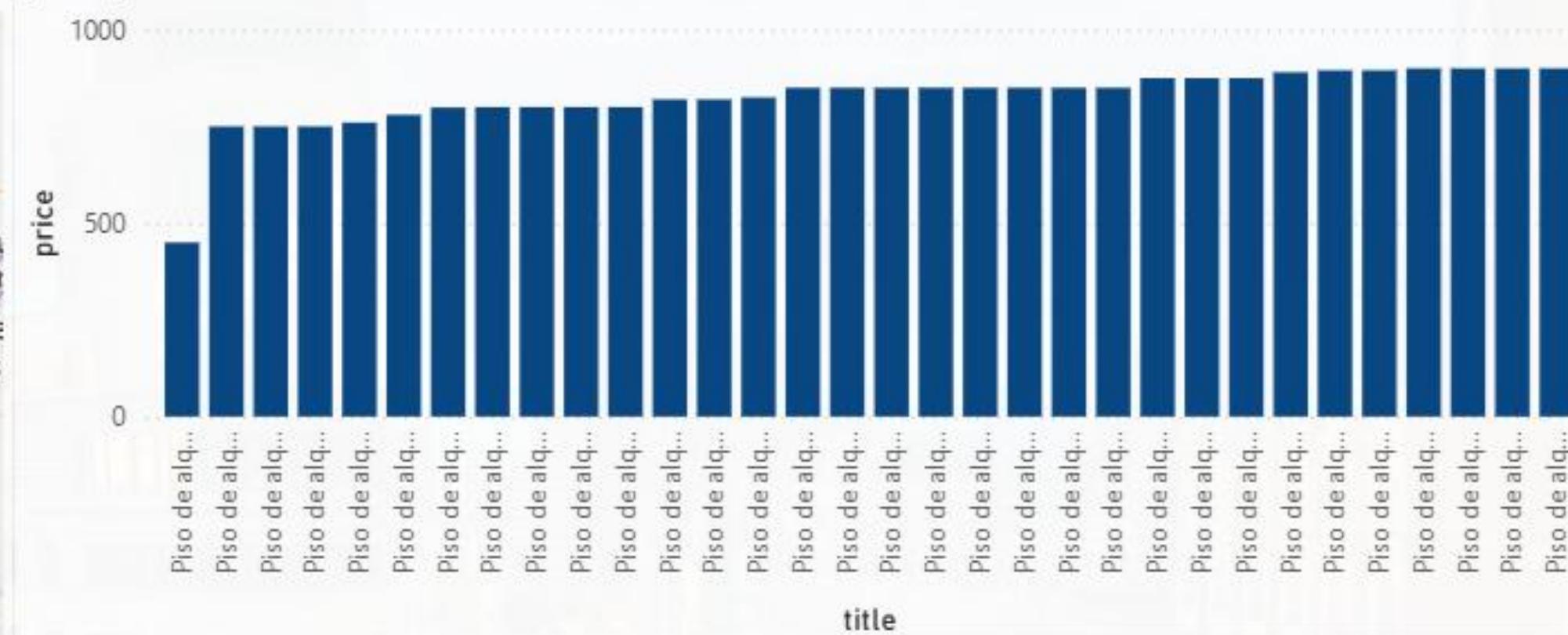
air_conditioning

 Aire acondicionado

title location price size_m2 rooms

title	location	price	size_m2	rooms	bathrooms	heating	air_conditioning	floor	href
Piso de alquiler en Bellas Vistas	Bellas Vistas	800	1		1	Calefacción	Aire acondicionado		https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-amueblado-internet/183338703/d?from=list
Piso de alquiler en Canillejas	Canillejas	799	25		1	Calefacción	Aire acondicionado		https://www.fotocasa.es/es/alquiler/vivienda/madrid-capital/aire-acondicionado-calefaccion-amueblado-internet/183295674/d?from=list

price por title



1000

price

1 a

Planta

floor

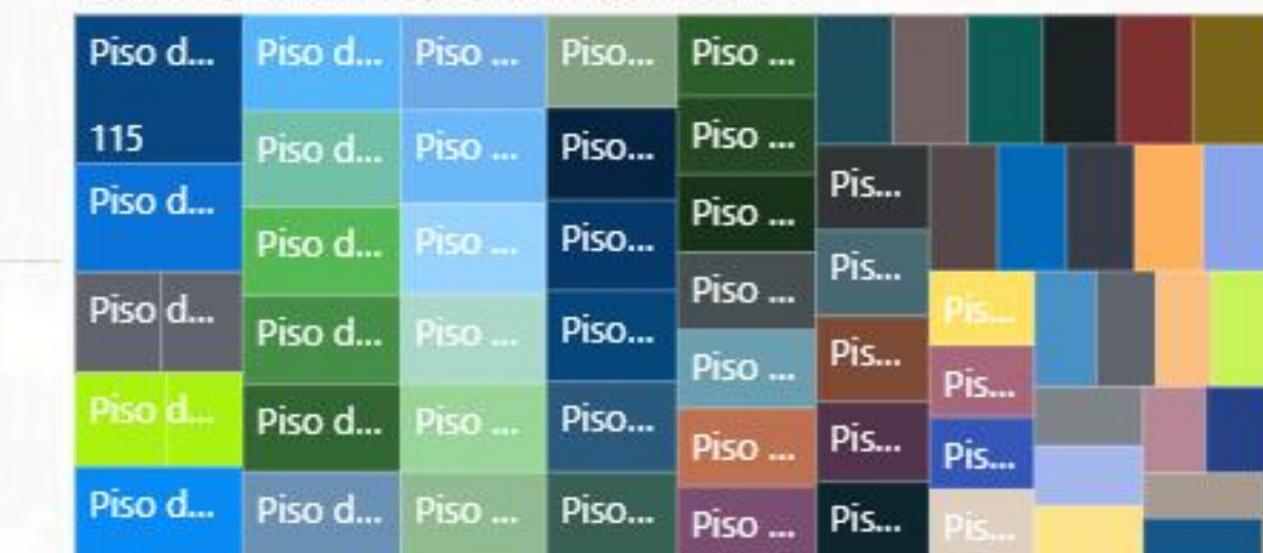
Máx. de rooms



Máx. de bathrooms



Máx. de size_m2 por title y size_m2





Muchas gracias