

In [ ]:

```
#EN ESTE SCRIP SE DEBERÁ INGRESAR DE MANERA MANUAL CADA UNA DE LAS COMUNAS EN LAS QUE SE DESEAN HACER WEBS  
CRAPING EN LA PAGINA WEB, ESTO SE DEBE HACER EN LA VARIABLE "comuna"  
#Y LUEGO AL FINAL DE LA PROGRAMACION EN "with open" SE DEBE CAMBIAR EL NOMBRE DEL ARCHIVO A EXPORTAR, EL C  
UAL EN ESTE CASO ES UN CSV.
```

```
import time  
import traceback  
import csv  
from selenium import webdriver  
from selenium.webdriver.chrome.service import Service as ChromeService  
from webdriver_manager.chrome import ChromeDriverManager  
from selenium.webdriver.common.by import By  
from selenium.common.exceptions import NoSuchElementException  
from selenium.webdriver.support.ui import WebDriverWait  
from selenium.webdriver.support import expected_conditions as EC  
  
driver = webdriver.Chrome(service=ChromeService(ChromeDriverManager().install()))  
# Lista de comunas  
comuna = ["Concepción, Biobío"]  
  
all_data = []  
  
# Crear instancia de driver  
try:  
    driver.get("https://www.portalinmobiliario.com/")  
  
    # Escribir la ubicación de donde quieres buscar  
    time.sleep(5)  
    buscadorlugar = driver.find_element(By.XPATH, '//*[@id="Rml5r:"]')  
    buscadorlugar.send_keys(comuna)  
  
    wait = WebDriverWait(driver, 10)  
    opciones = wait.until(EC.visibility_of_element_located((By.XPATH, '//button[contains(@class, "andes-l  
ist__item-action")]')))  
    opciones.click()  
  
    # Seleccionar el tipo de vivienda que quieres buscar  
    time.sleep(5)  
    buscadorvivienda = driver.find_element(By.XPATH, '//*[@id="R4l5r:-trigger"]')  
    buscadorvivienda.click()  
  
    opcion_a_seleccionar = driver.find_element(By.XPATH, '//span[text()="Casas"]')  
    opcion_a_seleccionar.click()  
  
    # Botón buscar  
    botonbuscar = driver.find_element(By.XPATH, '//*[@id="R355r:"]')  
    botonbuscar.click()  
  
    time.sleep(5)  
  
    list_datos_casas = []  
    existe_boton_siguiente = True  
  
    propiedad_casa = "Casa"  
  
    # Bucle para recopilar datos de las casas  
    while existe_boton_siguiente:  
        list_url_casas = []  
        list_casas = driver.find_elements(By.XPATH, '//li[@class="ui-search-layout__item"]')  
  
        for casas in list_casas:  
            url_casa = casas.find_element(By.XPATH, './a[@class="ui-search-result__content-wrapper ui-se  
arch-link"]').get_attribute("href")  
            list_url_casas.append(url_casa)  
  
            for url_casa in list_url_casas:  
                driver.get(url_casa)  
                #time.sleep(5)  
                WebDriverWait(driver, 5).until(EC.presence_of_element_located((By.XPATH, '//div[@class="ui-pd  
p-price__second-line"]//span[@class="andes-money-amount__fraction"]')))  
  
                try:  
                    precio = driver.find_element(By.XPATH, '//div[@class="ui-pdp-price__second-line"]//span[@  
class="andes-money-amount__fraction"]').text  
                except NoSuchElementException:  
                    precio = "No disponible"
```

```

        try:
            direccion = driver.find_element(By.XPATH, '//div[@class="ui-pdp-media ui-vip-location__subtitle ui-pdp-color--BLACK"]//p[@class="ui-pdp-color--BLACK ui-pdp-size--SMALL ui-pdp-family--REGULAR ui-pdp-media__title"]').text
        except NoSuchElementException:
            direccion = "No disponible"

        detalles = driver.find_elements(By.XPATH, '//div[@class="ui-pdp-highlighted-specs-res__icon-label"]//span[@class="ui-pdp-color--BLACK ui-pdp-size--SMALL ui-pdp-family--REGULAR ui-pdp-label"]')

        tamanios = []
        dormitorios = []
        banios = []

        for i in range(0, len(detalles), 3):
            if i < len(detalles):
                tamanios.append(detalles[i].text.strip())
            if i + 1 < len(detalles):
                dormitorios.append(detalles[i + 1].text.strip())
            if i + 2 < len(detalles):
                banios.append(detalles[i + 2].text.strip())

        datos_casas = [comuna, propiedad_casa, precio, direccion, tamanios, dormitorios, banios]

        list_datos_casas.append(datos_casas)

        driver.back()

    try:
        boton_aceptar_cookies = driver.find_element(By.XPATH, '//*[@id="newCookieDisclaimerButton"]')
        # Reemplaza con el ID real del botón
        boton_aceptar_cookies.click()
        time.sleep(2) # Espera un momento después de cerrar el banner
    except Exception as e:
        pass

    try:
        wait = WebDriverWait(driver, 20)
        elemento_siguiente = wait.until(EC.element_to_be_clickable((By.XPATH, '//span[contains(@class, "andes-pagination__arrow-title") and text()="Siguiente"]')))
        elemento_siguiente.click()
        print("hace click")
        time.sleep(10)
    except Exception as e:
        existe_boton_siguiente = False
        print("no hace click")
        traceback.print_exc()

    with open('casas_concepcion.csv', 'w', newline='', encoding='utf-8') as file: #CAMBIAR NOMBRE DEL ARCHIVO
        writer = csv.writer(file)
        writer.writerow(["Comuna", "Propiedad", "Precio", "Direccion", "Tamanios", "Dormitorios", "Baños"])
        writer.writerows(list_datos_casas)

    driver.get("https://www.portalinmobiliario.com/")
    time.sleep(5)

except Exception as e:
    print(f"Error al procesar {comuna}: {e}")
    traceback.print_exc()

driver.quit()

```