

Web Scraping : Mubawab.ma

```
In [11]: # install chromium, its driver, and selenium
!apt-get update
!apt install chromium-chromedriver
!cp /usr/lib/chromium-browser/chromedriver /usr/bin
!pip install selenium
# set options to be headless, ..
from selenium import webdriver
options = webdriver.ChromeOptions()
options.add_argument('--headless')
options.add_argument('--no-sandbox')
options.add_argument('--disable-dev-shm-usage')

In [5]: import numpy as np
import pandas as pd

In [3]: # open it, go to a website, and get results
wd = webdriver.Chrome('chromedriver',options=options)
wd.get("https://www.mubawab.ma/fr/plr/casablanca-settat/listing-promotion")
# print(wd.page_source) # results

In [4]: url_ard = wd.current_url
```

Get the city's number districts to navigate

```
In [6]: n_elems = len(wd.find_elements_by_xpath('//*[@id="listingUl"]/li'))
n_elems
```

Get Prices , quartiers , Standing , IDs and Dates

```
In [4]: prices = []
quartiers = []
standings = []
ids = []
dates = []

for i in range(n_elems):
    # price
    prix = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{i+1}]/div[2]/span') #Locate the price
    prix_text = prix.text #Get the price value
    # print(prix_text)
    prices.append(prix_text)

    # quartier
    quartier = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{i+1}]/div[2]/h2/a') #Locate the nq'th neigh
    qrt_text = quartier.text #Get the neighbourhood name
    # print(qrt_text)
    quartiers.append(qrt_text)

    # standing
    standing = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{i+1}]/div[2]/div[1]/h4') #Locate the nq'th s
    standing_text = standing.text #Get the standing_text name
    # print(standing_text)
    standings.append(standing_text)

    # id annonceur
    id_annonceur = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{i+1}]')
    promotion_id = id_annonceur.get_attribute("promotion-id")
    # print(promotion_id)
    ids.append(promotion_id)

    # date livraison
    date_livraison = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{i+1}]/div[2]/div[2]/span/span') #Locat
    date_livraison_text = date_livraison.text
    # print(date_livraison_text)
    dates.append(date_livraison_text)
```

Show results in DataFrame

```
In [7]: all_data = [prices , quartiers , standings , ids , dates ]
df = pd.concat([pd.DataFrame(lst) for lst in all_data] , axis=1).reset_index(drop = True)
#Set the dataframe's columns
df.columns = ['prices','quartiers','standings','ids','dates']
df
```

	prices	quartiers	standings	ids	dates
0	À partir de 773 000 DH	Central Park	Haut standing, Finalisé	2621	Publié il y a plus de 6 mois
1	À partir de 2 800 000 DH	Atlantic Eden	Haut standing, Finalisé	2854	Publié il y a plus de 6 mois
2	À partir de 1 233 000 DH	QUEENS	Haut standing, Finalisé	2925	Publié il y a plus de 6 mois
3	Prix à consulter	Ahl Loughlam- Lots de terrain	Haut standing, Finalisé	2734	Publié il y a plus de 6 mois
4	À partir de 1 600 000 DH	Sea View	Haut standing, Finalisé	3294	Publié il y a plus de 6 mois
5	À partir de 726 000 DH	Résidence Assafoi - Terrains et loca...	Moyen standing, Finalisé	3221	Publié il y a plus de 6 mois
6	À partir de 2 900 000 DH	Duplex 3 chambres Front Mer	Haut standing, Finalisé	3357	Publié il y a plus de 6 mois
7	À partir de 720 000 DH	Rosa Parc Magasins	Haut standing, En cours de construction	2928	Publié il y a plus de 6 mois
8	À partir de 540 000 DH	Shems Al Madina - Lots de terrain	Haut standing, Finalisé	1729	Publié il y a plus de 6 mois
9	À partir de 519 000 DH	Ocean Park	Haut standing, En cours de construction	3262	Publié il y a plus de 6 mois
10	À partir de 978 500 DH	Fdala Business Center	Finalisé	3241	Publié il y a plus de 6 mois
11	À partir de 3 500 000 DH	Villas des Prés	Haut standing, Finalisé	2254	Publié il y a plus de 6 mois
12	À partir de 475 000 DH	Shems Al Madina - Lots de Villas	Haut standing, En cours de construction	1730	Publié il y a plus de 6 mois
13	À partir de 2 200 000 DH	Bouskoura Golf City	Haut standing, Finalisé	474	Publié il y a plus de 6 mois
14	À partir de 1 600 000 DH	Vert Marine 3	Haut standing, En cours de construction	3512	Publié il y a 2 mois
15	À partir de 880 000 DH	Les Perles Ain Sebaâ	Moyen standing, Finalisé	2772	Publié il y a plus de 6 mois
16	À partir de 890 000 DH	Casabay Sidi Rahal	Haut standing, Finalisé	863	Publié il y a plus de 6 mois
17	À partir de 848 000 DH	Joyau du Parc Magasins	Haut standing, Finalisé	2927	Publié il y a plus de 6 mois
18	À partir de 1 850 000 DH	Evasion Bouznika	Haut standing, Finalisé	2716	Publié il y a plus de 6 mois
19	À partir de 860 000 DH	Résidence DIAZ	Haut standing, Finalisé	3274	Publié il y a plus de 6 mois
20	À partir de 1 000 000 DH	Résidence Jacinth	Haut standing, Finalisé	3364	Publié il y a 5 mois
21	À partir de 823 000 DH	Résidence Les Jasmins I	Moyen standing, En cours de construction	3446	Publié il y a 3 mois
22	À partir de 820 000 DH	Next House Zenata Eco City	Haut standing, En cours de construction	2585	Publié il y a plus de 6 mois
23	À partir de 3 980 000 DH	Vert Marine - Triplex	Haut standing, Finalisé	3517	Publié il y a 1 mois
24	À partir de 1 070 000 DH	Les Allées de Californie lance sa 2è...	Haut standing, En cours de construction	3429	Publié il y a 3 mois
25	Prix à consulter	Ahl Loughlam - Moyen standing-Tit Mel...	Moyen standing, Finalisé	1732	Publié il y a plus de 6 mois
26	À partir de 519 303 DH	Zenata Tower	Haut standing, Finalisé	2563	Publié il y a plus de 6 mois
27	À partir de 770 000 DH	LIV'IN Garden	Haut standing, En cours de construction	3289	Publié il y a plus de 6 mois
28	Prix à consulter	Famille D'amour	Haut standing, Finalisé	3479	Publié il y a 2 mois
29	À partir de 716 000 DH	Les Jardins de Ain Sebâa lance sa de...	Moyen standing, En cours de construction	2227	Publié il y a plus de 6 mois

Get Latitude , Longitude and concatenate all columns

```
In [7]: lats = []
longs = []

import time
t = 5

for j in range(n_elems):
    # long and lat
    # wd.get(url_ard) #Go back to last page
    page_list = wd.find_element_by_xpath(f'//*[@id="listingUl"]/li[{j+1}]') #Locate the price
    page_list.click()
    time.sleep(t)
    mapOpen = wd.find_element_by_id('mapOpen')
    lat = mapOpen.get_attribute("lat")
    lon = mapOpen.get_attribute("lon")
    lats.append(lat)
    longs.append(lon)
    # wd.execute_script("window.history.go(-1)")
    wd.back()
    time.sleep(t)

all_data = [prices , quartiers , standings , ids , dates , lats , longs]
df = pd.concat([pd.DataFrame(lst) for lst in all_data], axis=1).reset_index(drop = True)
#Set the dataframe's columns
df.columns = ['prices','quartiers','standings','ids','dates','lats','longs']
```

Show all results in a Dataframe

```
In [9]: df
```

	prices	quartiers	standings	ids	dates	lats	longs
0	À partir de 773 000 DH	Central Park	Haut standing, Finalisé	2621	Publié il y a plus de 6 mois	33.70259958074599	-7.399420738220243
1	À partir de 2 800 000 DH	Atlantic Eden	Haut standing, Finalisé	2854	Publié il y a plus de 6 mois	33.51788561426908	-7.795518636703463
2	À partir de 1 233 000 DH	QUEENS	Haut standing, Finalisé	2925	Publié il y a plus de 6 mois	33.57905668978428	-7.647800104393838
3	Prix à consulter	Ahl Loughlam- Lots de terrain	Haut standing, Finalisé	2734	Publié il y a plus de 6 mois	33.57340232956212	-7.478427886963118
4	À partir de 1 600 000 DH	Sea View	Haut standing, Finalisé	3294	Publié il y a plus de 6 mois	33.53037726442764	-7.8294328124397055
5	À partir de 726 000 DH	Résidence Assafoi - Terrains et loca...	Moyen standing, Finalisé	3221	Publié il y a plus de 6 mois	33.54574410154525	-7.739739418029757
6	À partir de 2 900 000 DH	Duplex 3 chambres Front Mer	Haut standing, Finalisé	3357	Publié il y a plus de 6 mois	33.52996746536189	-7.829604148864917
7	À partir de 720 000 DH	Rosa Parc Magasins	Haut standing, En cours de construction	2928	Publié il y a plus de 6 mois	33.703081563375555	-7.391266822814885
8	À partir de 540 000 DH	Shems Al Madina - Lots de terrain	Haut standing, Finalisé	1729	Publié il y a plus de 6 mois	33.613046760187586	-7.128281593322384
9	À partir de 519 000 DH	Ocean Park	Haut standing, En cours de construction	3262	Publié il y a plus de 6 mois	33.65439484793648	-7.477188643157888
10	À partir de 978 500 DH	Fdala Business Center	Finalisé	3241	Publié il y a plus de 6 mois	33.708775894156645	-7.397832870483342
11	À partir de 3 500 000 DH	Villas des Prés	Haut standing, Finalisé	2254	Publié il y a plus de 6 mois	33.51741153987773	-7.819862365722656
12	À partir de 475 000 DH	Shems Al Madina - Lots de Villas	Haut standing, En cours de construction	1730	Publié il y a plus de 6 mois	33.612153263231825	-7.128067016601619
13	À partir de 2 200 000 DH	Bouskoura Golf City	Haut standing, Finalisé	474	Publié il y a plus de 6 mois	33.468909499547884	-7.605187939053451
14	À partir de 1 600 000 DH	Vert Marine 3	Haut standing, En cours de construction	3512	Publié il y a 2 mois	33.525617486869486	-7.779886722564612
15	À partir de 880 000 DH	Les Perles Ain Sebaâ	Moyen standing, Finalisé	2772	Publié il y a plus de 6 mois	33.47727218776036	-7.9520416259765625
16	À partir de 890 000 DH	Casabay Sidi Rahal	Haut standing, Finalisé	863	Publié il y a plus de 6 mois	33.613904508551194	-7.537994384765739
17	À partir de 848 000 DH	Joyau du Parc Magasins	Haut standing, Finalisé	2927	Publié il y a plus de 6 mois	33.79291733259281	-7.159635275602341
18	À partir de 1 850 000 DH	Evasion Bouznika	Haut standing, Finalisé	2716	Publié il y a plus de 6 mois	33.82720675529752	-7.1224145045069065
19	À partir de 860 000 DH	Résidence DIAZ	Haut standing, Finalisé	3274	Publié il y a plus de 6 mois	33.57851904580247	-7.643845358811291
20	À partir de 1 000 000 DH	Résidence Jacinth	Haut standing, Finalisé	3364	Publié il y a 5 mois	33.57204358621036	-7.4776983261109535
21	À partir de 823 000 DH	Résidence Les Jasmins I	Moyen standing, En cours de construction	3446	Publié il y a 3 mois	33.61252853308126	-7.128324508666964
22	À partir de 820 000 DH	Next House Zenata Eco City	Haut standing, En cours de construction	2585	Publié il y a plus de 6 mois	33.6115456800117	-7.535612583160315
23	À partir de 3 980 000 DH	Vert Marine - Triplex	Haut standing, Finalisé	3517	Publié il y a 1 mois	33.703081563375555	-7.391266822814885
24	À partir de 1 070 000 DH	Les Allées de Californie lance sa 2è...	Haut standing, En cours de construction	3429	Publié il y a 3 mois	33.612153263231825	-7.128067016601619
25	Prix à consulter	Ahl Loughlam - Moyen standing-Tit Mel...	Moyen standing, Finalisé	1732	Publié il y a plus de 6 mois	33.57905668978428	-7.647800104393838
26	À partir de 519 303 DH	Zenata Tower	Haut standing, Finalisé	2563	Publié il y a plus de 6 mois	33.547117579470296	-7.675495147705078
27	À partir de 770 000 DH	LIV'IN Garden	Haut standing, En cours de construction	3289	Publié il y a plus de 6 mois	33.70650177897868	-7.3963093757629395
28	Prix à consulter	Famille D'amour	Haut standing, Finalisé	3479	Publié il y a 2 mois	33.5201665019058	-7.757291793823242
29	À partir de 716 000 DH	Les Jardins de Ain Sebâa lance sa de...	Moyen standing, En cours de construction	2227	Publié il y a plus de 6 mois	33.61040882133904	-7.480355685654416

Save Data in xlsx file

```
In [10]: df.to_excel(r'mubawab_data.xlsx', index = False)
```