

**DEÜ Fen Fakültesi
Bilgisayar Bilimleri Bölümü**

BİL 3013 Veri Madenciliğine Giriş

Ödev 1. Web kazıma tekniği ile veri toplama

Kodların ve Veri setlerinin Kaggle linki:

<https://www.kaggle.com/code/ozgurd5/bil3013-data-mining-assignment-1-code>
<https://www.kaggle.com/datasets/ozgurd5/bil3013-datamining-assignment-1-data>

Grup 2

Öğrenci: Özgür Dalbeler - 2022280084

Öğretim Üyesi: Prof. Dr. Efendi NASİBOĞLU

İzmir 2024

1- Ödevin Tanımı

hepsiemlak.com sitesinden **İzmir, Buca'daki** kiralık dairelerin **İlan tarihi, İlan metni, Mahallesi, Yerleştiği kat, Yaşı, Oda sayısı, Genişliği, Eşyalı mı?, Doğalgaz var/yok, Kira bedeli** gibi özelliklerini web kazıma yöntemleriyle çekerek bir **Excel** dosyasında depolayan **Python** uygulamasını geliştirin. Veriler çekilen sayfalarda **sayfalama** veya **scrolling** işlemleri **otomatik** yapılacaktır. **En az 500** örnek veri toplanacaktır.

Yapılması gerekenler:

- 1.Her öğrenci **kaggle.com** sitesinde kişisel hesap açmalı.
- 2.Geliştirdikleri **program kodlarını ve oluşturdukları veri setini** bu sitede public olarak diğer kullanıcılarla veya isterlerse dersin öğretim üyesi olarak sadece benimle **paylaşmalı**.
- 3.Herkes **kişisel** Ödev Raporunu **Word** dosyası olarak **Sakai** sisteminde **Ödev** sekmesine yüklemeli. Raporun kapak sayfasında erişim için öğrencinin Kaggle linkini de yazmalı.

2- Kullanılan Yöntemler ve Teknolojiler

Selenium: Selenium, Python ile tarayıcıda otomatik gezinme ve veri çıkarma işlemleri yapmak için kullanılır.

Kurulumu:

```
pip install selenium
```

Kullanımı:

```
from selenium import webdriver
from selenium.webdriver.common.by import By
driver = webdriver.Chrome()
driver.get("https://www.hepsiemlak.com")
```

Pandas: Pandas, Python'da veri analizi ve manipülasyonu için kullanılan bir kütüphanedir. DataFrame, tablo yapısında verileri saklar ve CSV, XLSX, JSON gibi formatlardan veri okuma ve yazma imkanı sunar.

Kurulumu:

```
pip install pandas
```

Kullanımı:

```
import pandas as pd
df = pd.DataFrame(ilanlar)
```

Openpyxl: Pandas kütüphanesinin Excel dosyasına veri yazabilmesi için openpyxl kütüphanesinin yüklenmesi gereklidir.

Kurulumu:

```
pip install openpyxl
```

Kullanımı:

```
df.to_excel('ilanlar.xlsx', index=False)
```

3- Uygulama

Bu projede **hepsiemlak.com** sitesindeki Buca ilçesindeki kiralık dairelerin ilanları toplanmıştır. Selenium ile web sayfasına giriş yapılmış ve ilan sayfalarından gerekli bilgiler alınmıştır. Veriler toplandıktan sonra dinamik olarak birxlsx dosyasına kaydedilmiştir.

Uygulama sayfaları dolaşmaktadır. Sayfaların içinde ilanları dolaşıp fiyat, başlık, konum, oda sayısı, alan, bina yaşı, kat sayısı ve ilan sayfası linkini kaydetmekte daha sonra ilan linklerini dolaşıp eşya durumu, yakıt tipi ve ısinma tipi verilerini elde etmektedir.

Uygulamanın önemli adımları aşağıdaki gibidir:

3.1- Modüller

```
# Pandas modülü

import pandas as pd


# Selenium modülü

from selenium import webdriver


# Sayfada filtreleme işlemi yapabilmek için gerekli modüller

from selenium.webdriver.common.by import By


# Sayfanın yüklenliğini algılamamız için gerekli modüller

from selenium.webdriver.support import expected_conditions

from selenium.webdriver.support.wait import WebDriverWait
```

3.2- Sayfayı yüklemek

```
# WebDriver başlatılıyor - Chrome tarayıcı

print("Tarayıcı başlatılıyor...")

driver = webdriver.Chrome()


# ilanları tutmak için liste

ilan_listesi = []
```

```
# Excel dosyası oluşturuluyor

excel_file = 'ilanlar.xlsx'

(pd.DataFrame(columns=[

    'Başlık',
    'Konum',
    'Fiyat',
    'Oda Sayısı',
    'Alan',
    'Bina Yaşı',
    'Kat',
    'Tarihi',
    'Link',
    'Eşya Durumu',
    'Isınma Tipi',
    'Yakıt Tipi']))

.to_excel(excel_file, index=False))

# Toplam sayfa sayısı. Her sayfada 24 ilan var. 21 sayfa için 504 ilan var.

sayfa_sayisi = 21

for page in range(1, sayfa_sayisi + 1):

    # Sayfa URL'si oluşturuluyor

    url = f'https://www.hepsiemlak.com/buca-kiralik?page={page}'

    # Sayfa yükleniyor

    print(f"Sayfa {page} yükleniyor: {url}")

    driver.get(url)
```

```
# Sayfa yüklenene kadar bekleniyor

WebDriverWait(driver,
10).until(expected_conditions.presence_of_element_located((By.CLASS_NAME,
'listing-item')))

print("Sayfa yüklendi.")
```

3-3 İlanların Bilgilerini Kazımak

İlanların '`listing-item`' adındaki bir class içinde tanımlandığı gözlemlendi. İlanlara erişmek ve elementlerini dolaşmak için gerekli kod yazıldı.

```
# İlanlar bulunuyor
ilanlar = driver.find_elements(By.CLASS_NAME, 'listing-item')
print(f"{len(ilanlar)} ilan bulundu.")

# İlanlar işleniyor
for ilan in ilanlar:
```

İlan başlığının '`card-link`' elementinin '`title`' değişkeninde yer aldığı gözlemlendi ve gerekli kod yazıldı.

The screenshot shows a search results page for 'Buca Kiralık' on the hepsiemlak.com website. The first listing is highlighted with a red box. The developer tools are open, showing the DOM structure of the highlighted listing. The title of the listing is 'ENDERES MH.DE 2+1 ÇIFT CEPHE ARAKAT 85 M2 ARAKAT KIRALIK DAIRE'.

```
# Başlık çıkarılıyor
baslik = ilan.find_element(By.CLASS_NAME, 'card-link').get_attribute('title')
```

İlan tarihinin '`list-view-date`' adındaki bir elementte tutulduğu gözlemlendi ve gerekli kod yazıldı.

The screenshot shows the same search results page for 'Buca Kiralık' on the hepsiemlak.com website. The date element of the first listing is highlighted with a red box. The developer tools show the DOM structure of the date element, which is a `span.list-view-date` element containing the text '19-10-2024'.

```
# Tarih çıkarılıyor
tarih = ilan.find_element(By.CLASS_NAME, 'list-view-date').text
```

İlan fiyatının '`list-view-price`' adındaki bir elementte tutulduğu gözlemlendi ve gerekli kod yazıldı.

```

Öğeler Konsol Kaynaklar Ağ Performans Bellek Uygulama Güvenlik Lighthouse
Kayıdıcı >> Stiller Hesapla Düzen Etkinlik Dinleyiciler >>
Y Filte Ejlenen sepiyi veya stil yok

```

```

14.000 TL

```

```

# Fiyat çıkarılıyor
fiyat = ilan.find_element(By.CLASS_NAME, 'list-view-price').text

```

Oda sayısı, alan, bina yaşı ve kat sayısı bilgilerinin '`short-property`' adındaki bir elementte tutulduğu gözlemlendi ve gerekli kod bloğu yazıldı. Sitedeki olası bir eksikliğin kod çalışırken hata vermemesi için elementlerin olmama ihtimali de göz önünde bulunduruldu.

```

Öğeler Konsol Kaynaklar Ağ Performans Bellek Uygulama Güvenlik Lighthouse
Kayıdıcı >> Stiller Hesapla Düzen Etkinlik Dinleyiciler >>
Y Filte Ejlenen sepiyi veya stil yok

```

```

2+1 | 84 m2 | 25 Yaşında | 2. Kat

```

```

# Özellikler çıkarılıyor
Özellikler = ilan.find_element(By.CLASS_NAME,
'short-property').text.split("\n")

```

```

# Özellikler listesi kontrol edilerek gerekli alanlar çıkarılıyor
# Eğer karşılık gelen index boş ise, ilgili alanlar boş olarak atanıyor
# Sıfırıncı indexte Kiralık Daire yazısı yer almaktır, o yüzden birinci indexten başlıyoruz

```

```

oda_sayisi = özellikler[1] if len(özellikler) > 1 else ''
alan = özellikler[2] if len(özellikler) > 2 else ''
bina_yası = özellikler[3] if len(özellikler) > 3 else ''
kat_sayısı = özellikler[4] if len(özellikler) > 4 else ''

```

İlan konumunun '`list-view-location`' adındaki bir elementte tutulduğu gözlemlendi ve gerekli kod bloğu yazıldı.

```

# Konum çıkarılıyor
konum = ilan.find_element(By.CLASS_NAME, 'list-view-location').text

```

Eşya durumu, yakıt tipi ve ısınma tipine erişmek için ilanların kendi sayfasına gidilmesi gerekmektedir. İlan linkinin '`card-link`' elementinin '`href`' değişkeninde yer aldığı gözlemlendi ve gerekli kod yazıldı.

```

# İlan linki alınıyor ve ilan_linkleri listesine ekleniyor

```

```
ilan_linki = ilan.find_element(By.CLASS_NAME,  
'card-link').get_attribute('href')  
sayfadaki_ilan_linkleri.append(ilan_linki)
```

İlan verileri listeye eklendi.

```
# İlan verileri listeye ekleniyor  
ilan_listesi.append({  
    'Başlık': baslik,  
    'Konum': konum,  
    'Fiyat': fiyat,  
    'Oda Sayısı': oda_sayisi,  
    'Alan': alan,  
    'Bina Yaşı': bina_yasi,  
    'Kat': kat_sayisi,  
    'Tarihi': tarih,  
    'Link': ilan_linki,  
})  
  
print(f"İlan eklendi: {baslik}, {konum}, {fiyat}, {oda_sayisi}, {alan},  
{bina_yasi}, {kat_sayisi}, {tarih}, {ilan_linki}")
```

3.4- İlan Sayfalarını Dolaşarak Diğer Verileri Kazımak

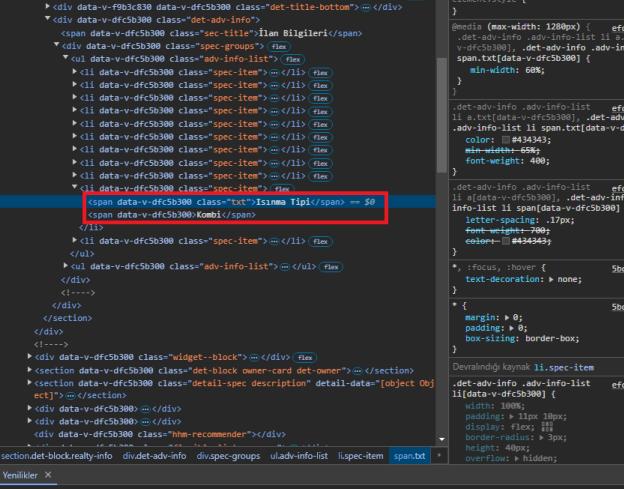
Eşya durumu, yakıt tipi ve ısınma tipi bilgilerine ulaşmak için ilan linkleri dolaşıldı.

```
# İlan linklerinde dolaşılıyor  
for ilan_linki in sayfadaki_ilan_linkleri:  
    # İlan linkine gidiliyor  
    print("İlan linki yükleniyor: ", ilan_linki)  
    driver.get(ilan_linki)  
  
    # Sayfa yüklenene kadar bekleniyor  
    WebDriverWait(driver,  
10).until(expected_conditions.presence_of_element_located((By.CLASS_NAME,  
'txt')))  
    print("İlan linki yüklendi.")
```

İlan bilgilerinin span olarak tutulduğu gözlemlendi ve gerekli kod yazıldı.

Menderes Mh. Öğrenciye Kiralık 4+1 130 m ² Teraslı Doğalgazlı Daire		20.000 TL
 Favori  Konumu  Yol Tarifi  Çevre		 Paylaş
İzmir Buca Menderes Kiralık Daire		 4 + 1  130 m ²
İlan Bilgileri		
İlan no	24511-9074	
Son Güncelleme Tarihi	26-09-2024	
İlan Durumu	Kiralık	
Konut Tipi	Daire	
Konut Şekli	Daire	
Oda + Salon Sayısı	4 + 1	
Brüt / Net M2	130 m ² / 125 m ²	
Bulunduğu Kat	2. Kat	
Bina Yaşı	25 Yaşında	
Isıtma Tipi	Kombi	
Kat Sayısı	2 Katlı	
		span.txt 217.8 x 16
Eşya Durumu		Eşyalı Değer
Banyo Sayısı	1	
Yapı Tipi	Betonarme	
Yapının Durumu	İkinci El	
Kullanım Durumu	Boş	
Tapu Durumu	Kat İrtifası	
Aidat	50 TL	
Cephe	Doğu, Batı	
Depozito	20.000 TL	
Yakit Tipi	Doğalgaz	
Yetkili Ofis	Evet	

Menderes Mh. Öğrenciye Kiralık 4+1 130 M2 Teraslı Doğalgazlı Daire				20.000 TL
İzmir	Buca	4 + 1	130 m ²	Favorilerime Ekle
Telefon Numarasını Göster	Mesaj Gönder	Whatsapp		
Konut Tipi	Daire	Yapının Durumu	İkinci El	
Konut Şekli	Daire	Kullanım Durumu	Boş	
Oda + Salon Sayısı	4 + 1	Tapu Durumu	Kat İrtifası	
Brüt / Net M2	130 m ² / 125 m ²	Aldat	50 TL	
Bulunduğu Kat	2. Kat	Cephe	Doğu, Batı	
span.txt 217x8 25 Yaşında  Sıfır 1 Tipi Kombi		Depozito	20.000 TL	
Kat Sayısı	2 Katlı	Yakit Tipi	Doğalgaz	
		Yetkili Ofis	Evet	
Firmanız Diğer İlanları				
 Buc Yıldızlar Mh. 2+1 Kapılı Mutfak Doğalgazlı Kiralik Daire 18.500 TL Buca, Yiğitler Mah.		Hatalı İlan Bildir Güvenlik Hatırlatmaları		
 Buca Nikah Sarayı Karşısında 3+1 125m ² İç Full Yapılı Arakat		Telefon Numarasını Göster Mesaj Gönder Whatsapp		
 FIRAT MINTAŞ TURYAP SİRİNVER Danışmannan Diğer İlanları Taşınmaz Ticareti Yetki Belgesi 3502430				



```
graph TD; A[.det-ad-info .adv-info-list li a.txt[data-v-dfc5b300]] --> B[span data-v-dfc5b300 class="spec-item"]; B --> C[span data-v-dfc5b300 class="spec-item"]; C --> D[span data-v-dfc5b300 class="spec-item"]; D --> E[span data-v-dfc5b300 class="spec-item"]; E --> F[span data-v-dfc5b300 class="spec-item"]; F --> G[span data-v-dfc5b300 class="spec-item"]; G --> H[span data-v-dfc5b300 class="spec-item"]; H --> I[span data-v-dfc5b300 class="spec-item"]; I --> J[span data-v-dfc5b300 class="spec-item"]; J --> K[span data-v-dfc5b300 class="spec-item"]; K --> L[span data-v-dfc5b300 class="spec-item"]; L --> M[span data-v-dfc5b300 class="spec-item"]; M --> N[span data-v-dfc5b300 class="spec-item"]; N --> O[span data-v-dfc5b300 class="spec-item"]; O --> P[span data-v-dfc5b300 class="spec-item"]; P --> Q[span data-v-dfc5b300 class="spec-item"]; Q --> R[span data-v-dfc5b300 class="spec-item"]; R --> S[span data-v-dfc5b300 class="spec-item"]; S --> T[span data-v-dfc5b300 class="spec-item"]; T --> U[span data-v-dfc5b300 class="spec-item"]; U --> V[span data-v-dfc5b300 class="spec-item"]; V --> W[span data-v-dfc5b300 class="spec-item"]; W --> X[span data-v-dfc5b300 class="spec-item"]; X --> Y[span data-v-dfc5b300 class="spec-item"]; Y --> Z[span data-v-dfc5b300 class="spec-item"]
```

Menderes Mh. Öğrenciye Kiralık 4+1 130 m² Teraslı
Doğalgazlı Daire

20.000 TL ⚡

Favori Konumu Yol Tarifi Çevre Paylaş

Izmir | Buca | Menderes | Kiralık | Daire | 4 + 1 | 130 m²

İlan Bilgileri

İlan no	24511-9074
Son Güncelleme Tarihi	26-09-2024
İlan Durumu	Kiralık
Konut Tipi	Daire
Konut Şekli	Daire
Oda + Salon Sayısı	4 + 1
Büyük / Net M2	130 m ² / 125 m ²
Bulunduğu Kat	2. Kat
Bina Yaşı	25 Yaşında
Isınma Tipi	Kombi
Kat Sayısı	2 Katlı
Eşya Durumu	Eşya Değil
Banyo Sayısı	1
Yapı Tipi	Betonarme
Yapının Durumu	İkinci El
Kullanım Durumu	Bos
Tapu Durumu	Kat İrtifağı
Aüstü	50 TL
Cephe	Doğu, Batı
Yerli Ofis	Evet
Yakıt Tipi	Doğalgaz
span.txt	217.8 × 16
20.000 TL	

FIRAT MİNTAŞ
TURYAP SİHİNYER
Danışmanın Diğer İlanları
Taşınmaz Ticareti Yetki Belgesi 3502430

Telefon Numarasını Göster

Mesaj Gönder WhatsApp

Developer Tools (Chrome DevTools) showing the DOM structure for the highlighted 'Yakıt Tipi' field.

```
# Tüm spanlar alınıyor
spans = driver.find_elements(By.CLASS_NAME, 'txt')

esya_durumu = ''
isinma_tipi = ''
yakit_tipi = ''

for span in spans:
    if span.text == 'Eşya Durumu': # Eğer span 'Eşya Durumu' metnini içeriyorsa
        # Kardeş ögesi olan diğer span'dan Eşya Durumu'nu al
        esya_durumu = span.find_element(By.XPATH,
"following-sibling::span").text

    elif span.text == 'Isınma Tipi': # Eğer span 'Isınma Tipi' metnini içeriyorsa
        # Kardeş ögesi olan diğer span'dan Isınma Tipi'ni al
        isinma_tipi = span.find_element(By.XPATH,
"following-sibling::span").text

    elif span.text == 'Yakıt Tipi': # Eğer span 'Yakıt Tipi' metnini içeriyorsa
        # Kardeş ögesi olan diğer span'dan Yakıt Tipi'ni al
        yakit_tipi = span.find_element(By.XPATH, "following-sibling::span").text

    if esya_durumu and yakit_tipi and isinma_tipi != '': # Eğer tüm bilgiler
alındıysa
        break # Döngüden çıkış

# İlan listesine 'Eşya Durumu', 'Yakıt Tipi' ve 'Isınma Tipi' ekleniyor
ilan_index = ilan_linkleri.index(ilan_linki)
ilan_listesi[ilan_index]['Eşya Durumu'] = esya_durumu
ilan_listesi[ilan_index]['Isınma Tipi'] = isinma_tipi
ilan_listesi[ilan_index]['Yakıt Tipi'] = yakit_tipi

print(f"İlan verileri eklendi: {esya_durumu}, {isinma_tipi}, {yakit_tipi}")
```

3.5- Verileri XSLX Olarak Dinamik Kaydetmek

```
# İlanı dinamik olarak Excel'e yazıyoruz

# Mevcut ilanı DataFrame'e dönüştürüyoruz ve liste haline getiriyoruz
df = pd.DataFrame([ilan_listesi[ilan_index]])
rows = df.values.tolist()

# Excel dosyasını yükliyoruz
workbook = load_workbook('ilanlar.xlsx')
sheet = workbook.active

# İlanı Excel dosyasına ekliyoruz ve dosyayı kaydediyoruz
for row in rows: sheet.append(row)
workbook.save('ilanlar.xlsx')

print(f"İlan Excel dosyasına eklendi: {ilan_linki}")

print(f"Sayfa {page} işlendi.")

# Tarayıcı kapatılıyor
print("Tarayıcı kapatılıyor...")
driver.quit()
```

3.6- Ekran Görüntüleri

```
İlan eklendi: Menderes Mh. Öğrenciye Kiralık 4+1 130 M2 Teraslı Doğalgazlı Daire, İzmir / Buca / Menderes Mah., 20.000 TL, 4 + 1, 130 m2, 25 Yaşında, 2. Kat, 26-09-2024, https://www.hepsiemlak.com/izmir-buca-menderes-kiralik/daire/24511-9074
İlan eklendi: GENİŞ, 2 ODA 1 SALON BAHÇELİ YÜKSEK ZEMİN, İzmir / Buca / Efeler Mah., 15.000 TL, 2 + 1, 109 m2, 29 Yaşında, Yüksek Giriş, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-efeler-kiralik/daire/13470-2483
İlan eklendi: Buca Göksu Mah İnkılap İzban Dibi 3+1 120 M2 D.gazlı, İzmir / Buca / Göksu Mah., 14.000 TL, 3 + 1, 120 m2, 21 Yaşında, Giriş Katı, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-goksu-kiralik/daire/83461-581
İlan eklendi: GÜNEŞ TEN ŞİRİNYER İZBAN 9DK 1+1 50M ASANSÖR KLİMA GÜNEY CEPHE, İzmir / Buca / İnkılap Mah., 16.500 TL, 1 + 1, 49 m2, 2 Yaşında, 4. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/130324-802
İlan eklendi: SİTE İÇİ KİRALIK 3+1 145m2 ARAKAT DAİRE, İzmir / Buca / Mustafa Kemal Mah., 34.000 TL, 3 + 1, 145 m2, 3 Yaşında, 12. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-mustafa-kemal-kiralik/daire/145608-91
İlan eklendi: Deniz emlak dan Buca'da Yenigün mahallesinde Kiralık daire, İzmir / Buca / Kozağaç Mah., 15.000 TL, 3 + 1, 130 m2, 16 Yaşında, 4. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-kozagac-kiralik/daire/120009-3
İlan eklendi: BUCADA EŞYALI, ASANSÖRLÜ, DOĞALGAZLI,KLİMALI, DOKUZ EYLÜL ÜNİ YANI, 1+1 KİRALIK, İzmir / Buca / Kurucusme Mah., 25.000 TL, 1 + 1, 50 m2, 4 Yaşında, 8. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-kurucesme-kiralik/daire/77173-1070
İlan eklendi: Kiralık Buca da Yiğitler Mah.3+1 110 m2 Doğalgazlı açık otoparkı olan merkez de bir daire, İzmir / Buca / Yiğitler Mah., 20.000 TL, 3 + 1, 110 m2, 7 Yaşında, Yarı Bodrum, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-yigitler-kiralik/daire/122602-100
İlan eklendi: FOR SALE & FOR RENTTEN 3+1 VALİ RAHMİ BEY MAH 150 M2 ARA KAT DAİRE, İzmir / Buca / Vali Rahmi Bey Mah., 25.000 TL, 3 + 1, 155 m2, 25 Yaşında, 3. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-valirahmi-bey-kiralik/daire/32982-4460
İlan eklendi: Buca Yedigöller 2+1 85 M2 D.gazlı Arakat Kiralık Daire, İzmir / Buca / Çağdaş Mah., 17.000 TL, 2 + 1, 85 m2, 25 Yaşında, 2. Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-cagdas-kiralik/daire/83461-473
İlan eklendi: FOR SALE & FOR RENTTEN 3+1 İNKILAPTA 150M2 ŞİRİNYER METRO YAKINI SON KAT DAİRE, İzmir / Buca / İnkılap Mah., 21.000 TL, 3 + 1, 160 m2, 25 Yasında, En Üst Kat, 19-10-2024, https://www.hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/32982-4450
```

Project DataMining main

```

Project
  - DataMining C:\Projects\Datav
    - .venv library root
    - Assignment_1
      -ilanlar.xlsx 19/10/2024 20
      -ilanlar1.xlsx 19/10/2024 2
      -ilanlar2.xlsx 18/10/2024 3
      -mainRequests.py 19/10/0
      -mainSelenium.py 19/10/0
      -Rapor örneği.pdf 17/10/2024
      -Ödev1.2.pdf 17/10/2024
    - External Libraries
Run mainSelenium

```

ilan linki yükleniyor: <https://www.hepsiemlak.com/izmir-buca-kozagac-kiralik/daire/120009-3>
ilan linki yüklenindi.
ilan verileri eklenindi: Eşyalı, Kombi, Doğalgaz
ilan Excel dosyasına eklenindi: <https://www.hepsiemlak.com/izmir-buca-kozagac-kiralik/daire/120009-3>
ilan linki yükleniyor: <https://www.hepsiemlak.com/izmir-buca-kurucesme-kiralik/daire/77173-1070>
ilan linki yüklenindi.
ilan verileri eklenindi: Eşyalı, Kombi, Doğalgaz
ilan Excel dosyasına eklenindi: <https://www.hepsiemlak.com/izmir-buca-kurucesme-kiralik/daire/77173-1070>
ilan linki yükleniyor: <https://www.hepsiemlak.com/izmir-buca-yigitler-kiralik/daire/122602-100>
ilan linki yüklenindi.
ilan verileri eklenindi: Eşyalı Değil, Kombi, Doğalgaz
ilan Excel dosyasına eklenindi: <https://www.hepsiemlak.com/izmir-buca-vali-rahmi-bey-kiralik/daire/32982-4460>
ilan linki yüklenindi.
ilan verileri eklenindi: Eşyalı Değil, Klima, Doğalgaz
ilan Excel dosyasına eklenindi: <https://www.hepsiemlak.com/izmir-buca-caqqas-kiralik/daire/83461-473>
ilan linki yükleniyor: <https://www.hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/32982-4450>

FOR SALE & FOR RENT 3+1 hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/32982-4450

Chrome is being controlled by automated test software.

Tam istediğiniz evi, arsayı, ofisi Hepsiemlak'ta kolayca ara, rahatça bul! Reklam filmimizi i

Konum, ilan no ya da emlak ofisi adıyla Giriş Yap veya üye ol Ücretsiz İlan EN - RU

İlan Kiralik > Izmir Kiralik > Buca Kiralik > İnkılâp Kiralik > Daire > 32982-4450



DİLEK BAŞER FOR SALE & FOR RENT Firmann Diğer İlanları Taşımmaz Ticaret Yetki Belgesi 500492 Cerezler ile toplanan kişisel verileriniz, gerek politikamızda belirtilen amaçlarla sınırlı ve mevzuata uygun şekilde kullanılacaktır Telefon Numarasını Göster Mesaj Gönder WhatsApp

Project DataMining main

```

Project
  - DataMining C:\Projects\Datav
    - .venv library root
    - Assignment_1
      -ilanlar.xlsx 19/10/2024 20
      -ilanlar1.xlsx 19/10/2024 2
      -ilanlar2.xlsx 18/10/2024 3
      -mainRequests.py 19/10/0
      -mainSelenium.py 19/10/0
      -Rapor örneği.pdf 17/10/2024
      -Ödev1.2.pdf 17/10/2024
    - External Libraries
Run mainSelenium

```

ilan eklendi: Menderes Mh. Öğrenciye Kiralık 4+1 130 M2 Teraslı Doğalgazlı Daire, İzmir / Buca / Menderes Mah. 4 + 1, 130 m², 25 Yaşında, 2. Kat, 26-09-2024, <https://www.hepsiemlak.com/izmir-buca-menderes-kiralik/daire/142478-123>
ilan eklendi: GENİŞ, 2 ODA 1 SALON BAĞELİ YÜKSEK ZEMİN, İzmir / Buca / Efeler Mah., 15.000
ilan eklendi: TL, 2 + 1, 109 m², 29 Yaşında, Yüksek Giriş, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-efeler-kiralik/daire/142478-123>
ilan eklendi: Buca Gökkuş Mah İnkılâp İZBAN Dibi 3+1 120 M2 D.Gazlı, İzmir / Buca / Gökkuş Mah., 14.000
ilan eklendi: TL, 3 + 1, 120 m², 21 yaşında, Giriş Katı, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-gokkus-kiralik/daire/142478-123>
ilan eklendi: GÜNEŞ TEN ŞİRİNVER İZBAN 9DK 1+1 50M ASANSÖR KLİMİ GÜNEŞ CEPHE, İzmir / Buca / İnkılâp Mah., 1 TL, 1 + 1, 49 m², 2 Yaşında, 4. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/142478-123>
ilan eklendi: SİTE İÇİ KİRALIK 3+1 145m² ARAKAT DAIRE, İzmir / Buca / Mustafa Kemal Mah., 34.000
ilan eklendi: TL, 3 + 1, 145 m², 3 Yaşında, 12. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-mustafa-kemal-kiralik/daire/142478-123>
ilan eklendi: Deniz emlak dan Buca'da Yenigün mahallesinde kiralık daire, İzmir / Buca / Kozağac Mah., 15.000
ilan eklendi: TL, 3 + 1, 130 m², 16 yaşında, 4. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-kozagac-kiralik/daire/142478-123>
ilan eklendi: BUCADA EŞYALI, ASANSÖRLÜ, DOĞALGAZLI, KLİMALI, DOKUZ EYLÜL ÜNİ YANI, 1+1 KİRALIK, İzmir / Buca TL, 1 + 1, 50 m², 4 Yaşında, 8. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-kurucesme-kiralik/daire/142478-123>
ilan eklendi: Kiralık Buca da Yigitler Mah.3+1 110 m² Doğalgazlı açık otoparkı olan merkez de bir daire, İzmir / Buca / Vali Rahmi Mah., 17.000
ilan eklendi: FOR SALE & FOR RENTEN 3+1 VALİ RAHMI BEY MAH 150 M2 ARR KAT DAIRE, İzmir / Buca / Vali Rahmi Mah., 3 + 1, 155 m², 25 Yaşında, 3. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-vali-rahmi-bey-kiralik/daire/142478-123>
ilan eklendi: Buca Yedigöller 2+1 85 M2 D.gazlı Arakat Kiralık Daire, İzmir / Buca / Çağdaş Mah., 17.000
ilan eklendi: TL, 2 + 1, 85 m², 25 Yaşında, 2. Kat, 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-caqqas-kiralik/daire/142478-123>
ilan eklendi: FOR SALE & FOR RENTEN 3+1 İNKILAPTA 150M2 ŞİRİNVER METRO YAKINI SON KAT DAIRE, İzmir / Buca / Vali Rahmi Mah., 3 + 1, 160 m², 25 Yaşında, En Üst Kat. 19-10-2024, <https://www.hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/142478-123>

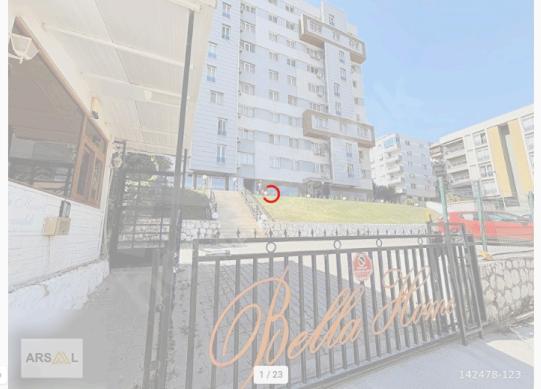
FOR SALE & FOR RENT 3+1 hepsiemlak.com/izmir-buca-inkilap-kiralik/daire/142478-123

Chrome is being controlled by automated test software.

Tam istediğiniz evi, arsayı, ofisi Hepsiemlak'ta kolayca ara, rahatça bul! Reklam filmimizi i

Konum, ilan no ya da emlak ofisi adıyla Giriş Yap veya üye ol Ücretsiz İlan EN - RU

İlan Kiralik > Izmir Kiralik > Buca Kiralik > Atatürk Kiralik > Daire > 142478-123



ARS L Dev Foto 1 / 23 142478-123

SUM: Not enough values Python 3.12 (DataMining) main

C0	C1	C2	C3	C4	C5	C6	C7	C8
	Konum	Fiyat	Oda Say..	Alan	Bina Ya..	Kat	Tarihi	Li
1	Başlık							
2	MENDERES MH.DE 2+1 ÇİFT CEPHE ARAKAT 85 M2 ARAKAT Kİ.. İzmir / Buca / Menderes Mah.	14.000	2 + 1	84 m ²	25 Yaşı..	2. Kat	19-10-2..	ht
3	Menderes Mh. Öğrenciye Kiralık 4+1 130 M2 Teraslı Do.. İzmir / Buca / Menderes Mah.	20.000	4 + 1	130 m ²	25 Yaşı..	2. Kat	26-09-2..	ht
4	GENİŞ, 2 ODA 1 SALON BAHÇELİ YÜKSEK ZEMİN	15.000	2 + 1	109 m ²	29 Yaşı..	Yüksek Gi..	19-10-2..	ht
5	Bucas Göksu Mah İnkılap İzban Dibi 3+1 120 M2 D.gazlı	14.000	3 + 1	120 m ²	21 Yaşı..	Giriş Katı	19-10-2..	ht
6	GÜNEŞ'TEN ŞİRİNYER İZBAN 9DK 1+1 50M ASANSÖR KLİMA G.. İzmir / Buca / İnkılap Mah.	16.500	1 + 1	49 m ²	2 Yaşın..	4. Kat	19-10-2..	ht
7	SİTE İÇİ KIRALIK 3+1 145m2 ARAKAT DAİRE	34.000	3 + 1	145 m ²	3 Yaşın..	12. Kat	19-10-2..	ht
8	Deniz emlak dan Buca'da Yenigün mahallesinde kiralık.. İzmir / Buca / Kozağaç Mah.	15.000	3 + 1	130 m ²	16 Yaşı..	4. Kat	19-10-2..	ht
9	BUÇADA EŞYALI, ASANSÖRLÜ, DOĞALGAZLI, KLİMALI, DOKUZ .. İzmir / Buca / Kuruçeşme Ma..	25.000	1 + 1	50 m ²	4 Yaşın..	8. Kat	19-10-2..	ht
10	Kiralık Buca da Yigitler Mah. 3+1 110 m2 Doğalgazlı a.. İzmir / Buca / Yigitler Mah.	20.000	3 + 1	110 m ²	7 Yaşın..	Yarı Bodrum	19-10-2..	ht
11	FOR SALE & FOR RENTEN 3+1 VALİ RAHİMİ BEY MAH 150 M2.. İzmir / Buca / Vali Rahmi B..	25.000	3 + 1	155 m ²	25 Yaşı..	3. Kat	19-10-2..	ht
12	Buca Yedigöller 2+1 85 M2 D.gazlı Arakat Kiralık Da.. İzmir / Buca / Çağdaş Mah.	17.000	2 + 1	85 m ²	25 Yaşı..	2. Kat	19-10-2..	ht
13	FOR SALE & FOR RENTEN 3+1 İNKILAPTA 150M2 ŞİRİNYER.. İzmir / Buca / İnkılap Mah.	21.000	3 + 1	160 m ²	25 Yaşı..	En Üst Kat	19-10-2..	ht
14	MARKA'DAN H.AĞA BAHÇESİ YAKINI 1+0 EŞYALI KIRALIK DA.. İzmir / Buca / Adatepe Mah.	11.000		Stüdyo	45 m ²	2 Yaşın..	Yüksek Gi..	19-10-2..
15	--DOĞALGAZLI-ARAKAT-Buca-dojum hastanesi YAKINI - 3+.. İzmir / Buca / Atatürk Mah.	19.000	3 + 1	125 m ²	13 Yaşı..	Ara Kat	19-10-2..	ht
16	Buca Çevik-1 Pazaryeri Çift Cepheli 2+1 Ayrı Mutfakl.. İzmir / Buca / Kozağaç Mah.	18.000	2 + 1	120 m ²	20 Yaşı..	2. Kat	19-10-2..	ht
17	MARKA'DAN ANACADDE&KAMPÜS YAKINI 1+1 EŞYALI ARAKAT K.. İzmir / Buca / Kuruçeşme Ma..	12.500	1 + 1	55 m ²	21 Yaşı..	2. Kat	19-10-2..	ht
18	GOLDİA YAPI'DAN KAMPÜS YAKINI KIRALIK 1+1 DAİRE !!!	16.000	1 + 1	50 m ²	3 Yaşın..	3. Kat	19-10-2..	ht
19	MARKA'DAN ANACADDE&KAMPÜS YAKINI 1+1 KLİMALI KIRALIK.. İzmir / Buca / Kuruçeşme Ma..	13.000	1 + 1	50 m ²	4 Yaşın..	Yüksek Gi..	19-10-2..	ht
20	MARKA'DAN ANACADDE&KAMPÜS YAKINI EŞYALI KLİMALI KİRA.. İzmir / Buca / Kuruçeşme Ma..	15.000	1 + 1	55 m ²	2 Yaşın..	2. Kat	19-10-2..	ht
21	ATATÜRK MAH'DE SİTE İÇİ ARA KAT D. GAZLI FULL EŞYALI.. İzmir / Buca / Atatürk Mah.	17.000	1 + 1	55 m ²	5 Yaşın..	5. Kat	18-10-2..	ht
22	BUCA ATATÜRK MAHALLESİNDE 3+1 FULL TADİLATLI MERKEZİ.. İzmir / Buca / Atatürk Mah.	20.000	3 + 1	130 m ²	15 Yaşı..	En Üst Kat	17-10-2..	ht
23	Şirinyer Güven Mh. 3+1 145 M2 Doğalgazlı Arakat Kira.. İzmir / Buca / Güven Mah.	21.500	3 + 1	145 m ²	20 Yaşı..	3. Kat	17-10-2..	ht
	ADACALDAM BUCA ŞİRİNYERDE 7+1 ESVAİLTİ DOKALCAZIT VİDA .. İzmir / Buca / VİZYON Mah	22.000	7 + 1	175 m ²	27 Yaş..	Ara Kat	17-10-2..	ht

SUM: Not enough values Python 3.12 (DataMining) main

C3	C4	C5	C6	C7	C8	C9	C10	C11
	Alan	Bina Ya..	Kat	Tarifi	Link	Eşya Duru..	İsimma Tipi	Yakıt T..
1	Oda Say..	Alan	Bina Ya..	Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-menderes-kiral..	Eşyali De..	İsimma Yok
2	2 + 1	84 m ²	25 Yaşı..	2. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-menderes-kiral..	Eşyali De..	Elektrik
3	3 + 4 + 1	130 m ²	25 Yaşı..	2. Kat	26-09-2..	https://www.hepsiemlak.com/izmir-buca-menderes-kiral..	Eşyali De..	Doğalgaz
4	2 + 1	109 m ²	29 Yaşı..	Yüksek Gi..	19-10-2..	https://www.hepsiemlak.com/izmir-buca-efeler-kiralik..	Eşyali De..	İsimma Yok
5	3 + 1	120 m ²	21 Yaşı..	Giriş Katı	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kiralik/..	Eşyali De..	Kombi
6	1 + 1	49 m ²	2 Yaşın..	4. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-inkilap-kiral..	Eşyali De..	Doğalgaz
7	3 + 1	145 m ²	3 Yaşın..	12. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-mustafa-kemal..	Eşyali De..	Merkezi (Pay Öl..)
8	3 + 1	130 m ²	16 Yaşı..	4. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kozagac-kiral..	Eşyali De..	Doğalgaz
9	1 + 1	50 m ²	4 Yaşın..	8. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kurucusme-kira..	Eşyali	Kombi
10	3 + 1	110 m ²	7 Yaşı..	Yarı Bodrum	19-10-2..	https://www.hepsiemlak.com/izmir-buca-yigitler-kiral..	Eşyali De..	Doğalgaz
11	3 + 1	155 m ²	25 Yaşı..	3. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-vali-rahmi-bey..	Eşyali De..	Doğalgaz Sobası
12	2 + 1	85 m ²	25 Yaşı..	2. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-cagdas-kiralik..	Eşyali De..	Klima
13	3 + 1	160 m ²	25 Yaşı..	En Üst Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-inkilap-kiral..	Eşyali De..	Doğalgaz
14	Stüdyo	45 m ²	2 Yaşın..	Yüksek Gi..	19-10-2..	https://www.hepsiemlak.com/izmir-buca-adatepe-kiral..	Eşyali	İsimma Yok
15	3 + 1	125 m ²	13 Yaşı..	Ara Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-ataeturk-kiral..	<null>	Doğalgaz
16	2 + 1	120 m ²	20 Yaşı..	2. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kozagac-kiral..	Eşyali	Kombi
17	1 + 1	55 m ²	21 Yaşı..	2. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kurucusme-kira..	Eşyali	Klima
18	1 + 1	50 m ²	3 Yaşın..	3. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-adatepe-kiral..	Eşyali	Doğalgaz
19	1 + 1	50 m ²	4 Yaşın..	Yüksek Gi..	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kurucusme-kira..	Eşyali	Klima
20	1 + 1	55 m ²	2 Yaşın..	2. Kat	19-10-2..	https://www.hepsiemlak.com/izmir-buca-kurucusme-kira..	Eşyali	<unset>
21	1 + 1	55 m ²	5 Yaşın..	5. Kat	18-10-2..	https://www.hepsiemlak.com/izmir-buca-ataturk-kiral..	Eşyali	Merkezi (Pay Öl..)
22	3 + 1	130 m ²	15 Yaşı..	En Üst Kat	17-10-2..	https://www.hepsiemlak.com/izmir-buca-ataturk-kiral..	Eşyali	Doğalgaz
23	3 + 1	145 m ²	20 Yaşı..	3. Kat	17-10-2..	https://www.hepsiemlak.com/izmir-buca-guven-kiralik/..	Eşyali De..	Kombi
	2 + 1	175 m ²	27 Yaş..	Ara Kat	17-10-2..	https://www.hepsiemlak.com/izmir-buca-vizyon-kiral..	Kombi	Doğalgaz

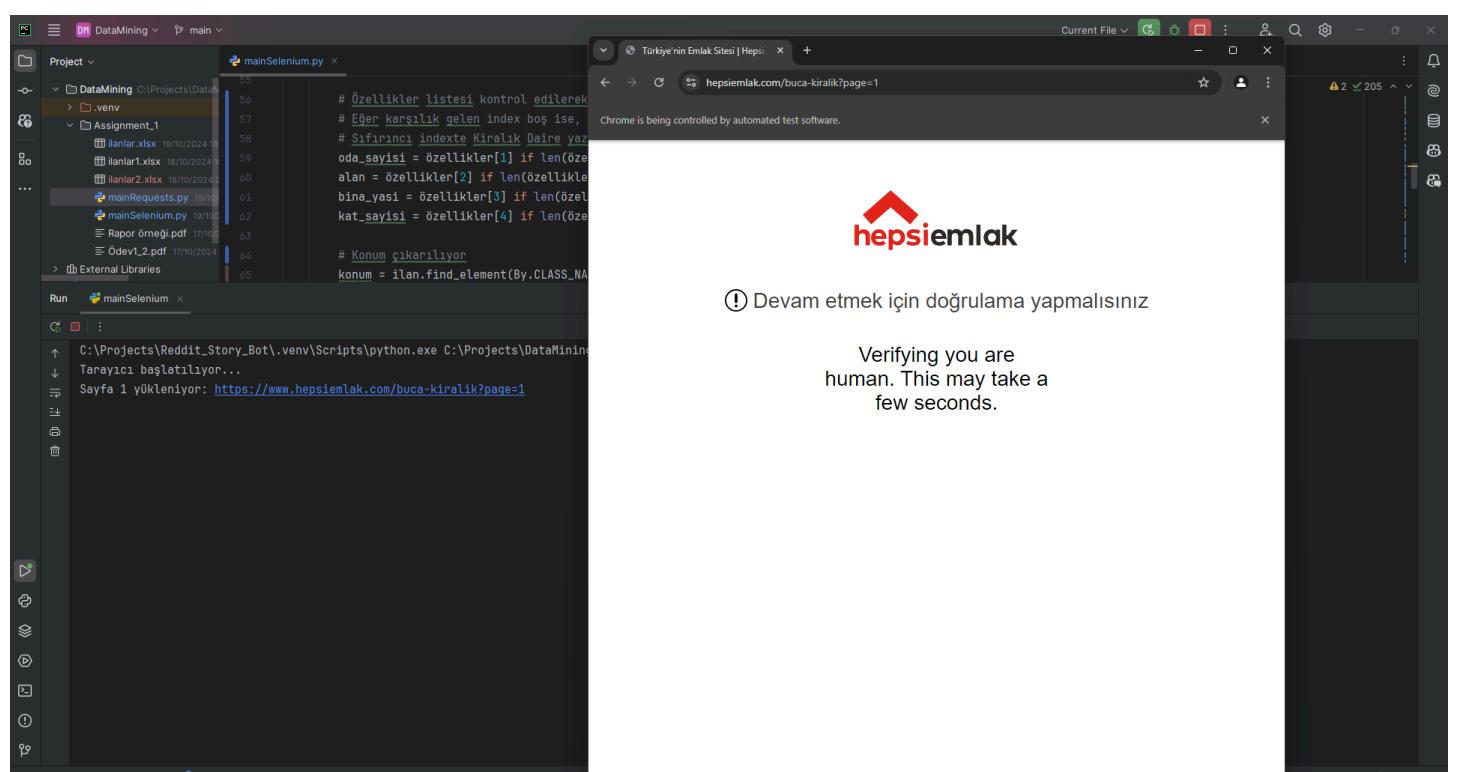
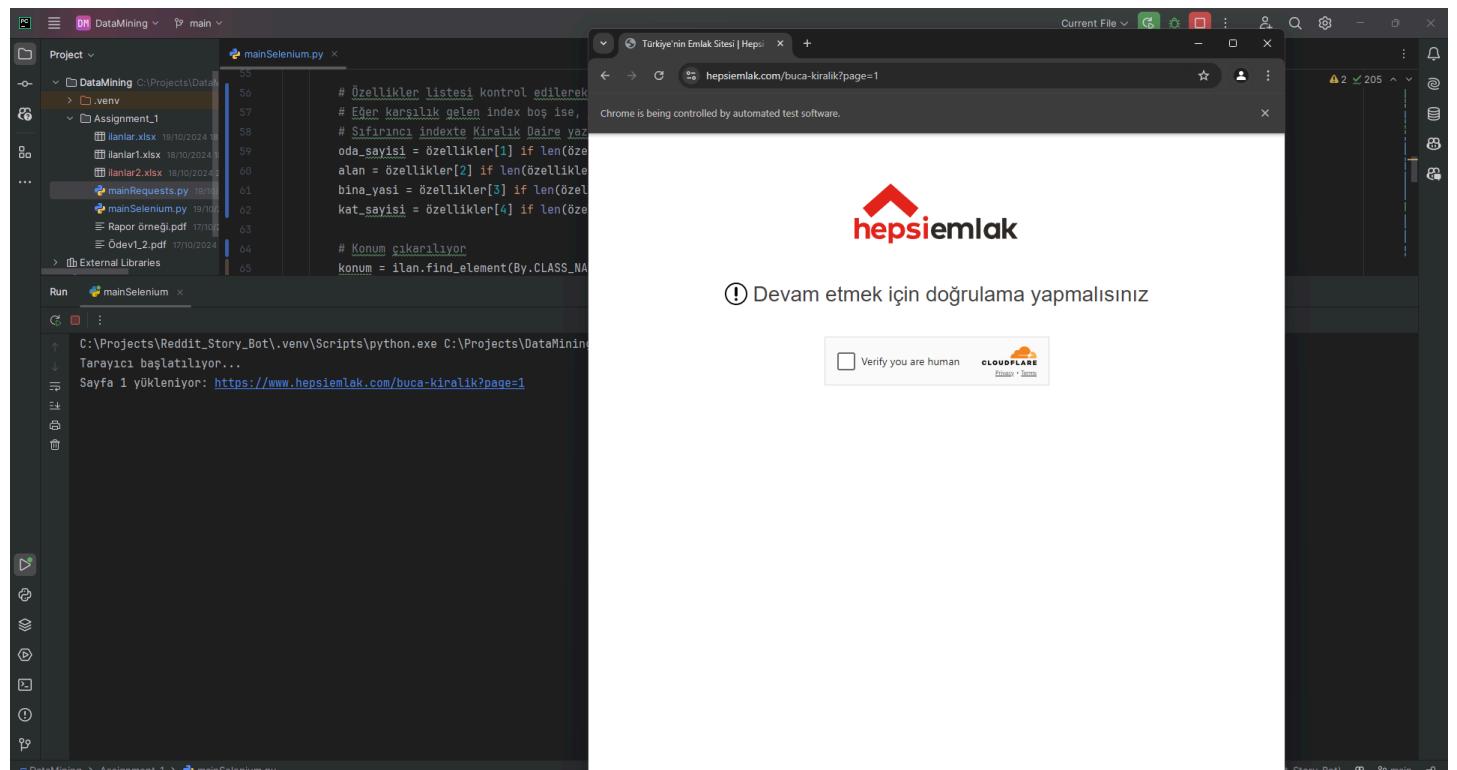
4- Değerlendirme

Öncelikle oyun geliştirme ile ilgilenen birisi olarak web geliştirme konusundaki tecrübesizliğim ve verilen ödevin herhangi bir derste de görmemişim bir şey olmasından ötürü kullanmam gereken yöntemler konusunda web development ile ilgilenen arkadaşlarımdan tavsiye aldım.

Çok ama çok fazla zorlukla karşılaştım ancak bu zorlukların hiçbiri veri kazıma ile ilgili değildi. Tüm sorunlarım doğrulama servisleri ile ilgiliydi.

Hangi verinin hangi elementte tutulduğunu keşfetmek, bunun nasıl çekileceği, nasıl tutulacağı, xlsx dosyasına nasıl kaydedileceği çok basit işlemlerdi ancak Cloudflare doğrulamasını aşmak veya devre dışı bırakmak inanılmaz zordu. Bunun için günlerce uğraştım ve en sonunda başaramadan pes ettim.

İlk önce Selenium kullanan bir proje hazırladım. Yukarıda tanıttığım ve ödev dosyasına eklediğim proje de bu. Bu proje ile 57 ilanın istenen tüm bilgilerine ulaşabildim. Eğer sitenin içine girip eşya ve doğalgaz durumunu çekmezsem 12 sayfa dolaşabiliyorum. Bu da 288 ilan ediyordu. Hiçbir koşulda en az 500 ilana ulaşamadım ve bu yüzden tüm verileri içeren 57 ilanı yollamakta karar kıldım.



Cloudflare doğrulamasını aşmak için web development ile ilgilenen arkadaşlarından yardım istedim ve bana request atarak yapmamı istediler. Attığım requestler işe yaramadı ve 403 Forbidden hatası

aldım. Bu sitede yasaklandığım anlamına geliyordu. Ardından yine aynı arkadaşlarımı danıştım ve bana aşağıdaki repoyu yolladılar, kendilerinin bununla uğraştıklarını söylediler. Bu repoyu kullanmak bir arkadaşımda işe yarasa da bende ve başka bir arkadaşımda işe yaramadı.

<https://github.com/zfcsoftware/cf-clearance-scrap>

Repodaki uygulama yolladığımız isteklere ekleyeceğimiz header ve cookie sağlayarak Cloudflare'i kandırmayı amaçlıyor.

Uygulamayı çalıştmak için bilgisayarıma docker kurdum. Projeyi selenium ile değil Requests ve BeautifulSoup ile hazırlamam gerekiyordu. Bunu da yaptım. Deneme amaçlı sadece fiyat listesini çeken ve kaydeden bir kod yazdım.

The screenshot shows the Docker Desktop interface. On the left, a sidebar has 'Containers' selected. The main area is titled 'Containers' with a search bar and a filter for 'Only show running containers'. A single container is listed:

Name	Image	Status	Port(s)	CPU (%)	Last started	Actions
jovial_dijik	zfcsoftware/cf-clearance-scrap:cc272db9	Running	3000:3000	1.07%	7 minutes ago	[Edit, More, Delete]

At the bottom, there are 'Walkthroughs' for 'Multi-container applications' (8 mins) and '\$ docker init' (3 mins). The status bar at the bottom shows 'Engine running', system resources (RAM 1.98 GB, CPU 0.06%), and a terminal tab.

```
# WAF oturumu olusturma fonksiyonu
def create_waf_session(url): 3 usages  ozgurd5
    try:
        print(f"WAF oturumu olusturuluyor: {url}")
        response = requests.post(
            CF_CLEARANCE_URL,
            headers={'content-type': 'application/json'},
            json={'url': url, 'mode': 'waf-session'},
            timeout=20
        )

        if response.status_code == 200:
            print("WAF oturumu olusturuldu.")
            data = response.json()
            return data.text

    else:
        print(f"WAF oturumu baglanti hatali: {response.status_code}")
        create_waf_session(url)

except Exception as e:
    print(f"Hata olustu, tekrar deneniyor: {e}")
    create_waf_session(url)
```

```
response = requests.get(url)

while response.status_code != 200:
    print(f"Ba\u111glant\u0111 hatas\u0111: {response.status_code}")
    header = create_waf_session(url)
    cookie_str = ""
    for cookie in header["cookies"]:
        cookie_str += f"{cookie['name']}={cookie['value']};"
    cookie_str = cookie_str[:-1]
    header["headers"]["cookie"] = cookie_str
    header["headers"]['user-agent'] = 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4453.89 Safari/537.36'
    response = requests.get(url, headers=header["headers"])

# Sayfa kayna\u111\u0111ini BeautifulSoup ile \u0131sleme
soup = BeautifulSoup(response.content, features="html.parser")
ilanlar = soup.find_all(name="li", class_="listing-item")
print(f"{len(ilanlar)} ilan bulundu.")

# \u0131lanları \u0131sleme
for ilan in ilanlar:
    # Fiyat c\u0111kariliyor
    fiyat = ilan.find('span', class_='list-view-price').get_text(strip=True)
    print("Fiyat: ", fiyat)
```

```
C:\Projects\DataMining\.venv\Scripts\python.exe C:\Projects\DataMining\Assignment_1\mainRequests.py
Ba\u111glant\u0111 hatası: 403
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
Hata oldu\u0111, tekrar deneniyor: HTTPConnectionPool(host='localhost', port=3000): Read timed out. (read timeout=20)
WAF oturumu olu\u111turu\u0111uyor: https://www.hepsiemlak.com/buca-kiralik?page=1
.
```

Bu yol da işe yaramadı. Modemimi resetledikten sonra Selenium ile tekrar ile tekrar başladım ancak request atma durumunda kalıcı olarak ban yemiştim. Ne yaparsam yapayım 403 Forbidden hatası alıyordu ve almaya devam ediyorum.

Tüm bu denemelerim sonucu modemimi bir kez daha resetledim ve Selenium projemi olduğu kadariyla yolladım. Modemimi tekrar tekrar resetleyip uyqlamamı en son kaldığım sayfadan başlatıp

verileri farklı bir dosyaya kaydedip daha sonra excel dosyalarını birbirlerine ekleyebilirdim ancak yapmadım.

Ayrıca scrape yapmanın yasal olarak gri bölgede olduğu ve sitelerin scrape edilmeyi engellemesine rağmen scrape etmeye çalışmanın suç olduğunu öğrendim. Bunun üzerine Google'da "Hepsiemlak robots.txt" şeklinde arama yaparak bu linke ulaştım.

<https://www.hepsiemlak.com/en/kullanim-kosullari>

Without the prior written permission of hepsiemlak the use of automatic programs, robots, web crawlers, spiders, data mining (data mining) data scanning (data crawling) etc. "screen scraping" software or systems on the site;

It is illegal to use the Site and the content on the Site outside the usage limits set by these Terms of Use; Hepsiemlak's related claims, lawsuits and follow-up rights are reserved.

Yaptığım şeyin Hepsiemlak tarafından hoş karşılanmadığını ve yasadışı olduğunun belirtildiğini de görünce daha fazla zorlamadım ve ödevimi ne yazık ki eksik yaptım. Bu olay yasal olarak sıkıntılı olmasaydı bile ödevimi eksik yapacaktım çünkü doğrulama sistemlerinin nasıl geçileceğini bilmiyorum. Özellikle de bu konuda dünyada 1 numara olan Cloudflare'i hackleyebileceğimi hiç zannetmiyorum.

Son olarak ödev hakkında geri dönüş vermek isterim. Pek çok kişi ödevdeki şeyleri derste görmemiş olmamızı bir sorun olarak görse de ben bunu "öğrenmeyi öğrenmek" olarak görüyorum. Ayrıca bu ödev güzel bir "gerçek hayat problemi". Ama tüm bunlara rağmen ders kazanımlarında doğrulama sistemlerini bypass etmek veya ilişkili bir şey olmadığı için bu ödevi kapsam dışı buldum. Doğrulama sistemi kullanmayan bir dummy website scrape etmek çok daha doğru olurdu. Çünkü bu ödevi yapan kimse sayfaları otomatik dolaşmak, otomatik scroll yapmak, HTML elementleri arasında bilgiyi bulmak, bunları tablolamak gibi şeylerle uğraşmadı, bu çok basit bir şey. Herkes zamanının neredeyse tamamını Cloudflare'i geçmek için bir yol bulmaya harcadı.

Son ve son olarak da bir sonraki ödevlerde veri madenciliğine izin veren veya public api sunan bir siteyi scrape etmek herkes için çok daha güvenli olur.

5- Kaynaklar

Chat GPT

<https://pandas.pydata.org/docs/reference/api/pandas.ExcelWriter.html>

<https://stackoverflow.com/questions/47737220/append-dataframe-to-excel-with-pandas>

<https://www.selenium.dev/documentation/>

https://www.selenium.dev/documentation/webdriver/getting_started/,

<https://requests.readthedocs.io/en/latest/user/quickstart/#make-a-request>

<https://www.docker.com/>

<https://github.com/zfcsoftware/cf-clearance-scrapers>