

```
In [ ]: import requests
import urllib.request
from bs4 import BeautifulSoup
import pandas as pd
```

```
In [ ]: class Cars:
    def __init__(self, mark_, price_, produced_year_, imported_year_, distance_, motor_volume_, color_, type_, hurd, hodolguur_):
        self.mark = mark_
        self.price = price_
        self.produced_year = produced_year_
        self.imported_year = imported_year_
        self.distance = distance_
        self.motor_volume = motor_volume_
        self.color = color_
        self.type = type_
        self.hurd = hurd
        self.hodolguur = hodolguur_
```

```
In [ ]: baseurl = "https://www.unegui.mn/avto-mashin/-avtomashin-zarna/?page="
car_list = []
for i in range(1, 200):
    url = baseurl + str(i)

    response = requests.get(url)
    if response.status_code != 200:
        print(response.status_code)
        print('error', url)
        continue
    soup = BeautifulSoup(response.text, "html.parser")

    li_list = soup.find_all("div", {"class": "swiper-wrapper"})
    for li in li_list:
        a = li.find('a')
        car_url = "https://www.unegui.mn" + a['href']
        #print(car_url)
        car_list.append(car_url)
```

```
In [ ]: def findFeature(li_list, header):
        ref = None
        for li in li_list:
            text = li.text.strip()
            if text.startswith(header):
                ref = text[len(header):].strip()
                break
        return ref
```

```
In [ ]: print(len(car_list))
        car_set = set(car_list)
        print(len(car_set))
```

11940

11940

```
In [ ]: it = 0
        cars_data = []
        for url in car_set:
            it += 1
            response = requests.get(url)
            if response.status_code != 200:
                print(response.status_code)
                print('error', url)
                continue
            soup = BeautifulSoup(response.text, "html.parser")

            mark = soup.find("h1", {"class": "title-announcement"}).text.strip()
            mark = str(mark.split(',')[0])
            price = soup.find("div", {"class": "announcement-price__cost"}).text.strip()
            price = float(price.split('сая')[0])
            li_class = soup.find_all("li")
            produced_year = findFeature(li_class, "Үйлдвэрлэсэн он:")
            imported_year = findFeature(li_class, "Орж ирсэн он:")
            distance = findFeature(li_class, "Явсан:")
            motor_volume = findFeature(li_class, "Мотор багтаамж:")
            color = findFeature(li_class, "Дотор өнгө:")
            type = findFeature(li_class, "Төрөл:")
            hurd = findFeature(li_class, "Хүрд:")
            hodolguur = findFeature(li_class, "Хөдөлгүүр:")
```

```
#print(title, price, p_year, i_year, distance, motor, color)
car = Cars(mark, price, produced_year, imported_year, distance, motor_volume, color, type, hurd, hodoiguur)
cars_data.append(car.__dict__)
#print(car)
```

```
In [ ]: df = pd.DataFrame(cars_data)
df.to_csv('car_info.tsv', sep="\t", index=False)
```