In [1]:
```
!pip install bs4
!pip install requests
```

```
Requirement already satisfied: bs4 in c:\users\manisha singh\onedrive\documents\py
thon scripts\lib\site-packages (0.0.1)
Requirement already satisfied: beautifulsoup4 in c:\users\manisha singh\onedrive\d
ocuments\python scripts\lib\site-packages (from bs4) (4.12.2)
Requirement already satisfied: soupsieve>1.2 in c:\users\manisha singh\onedrive\do
cuments\python scripts\lib\site-packages (from beautifulsoup4->bs4) (2.4)
Requirement already satisfied: requests in c:\users\manisha singh\onedrive\documen
ts\python scripts\lib\site-packages (2.31.0)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\manisha singh
\onedrive\documents\python scripts\lib\site-packages (from requests) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\manisha singh\onedrive\doc
uments\python scripts\lib\site-packages (from requests) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\manisha singh\onedri
ve\documents\python scripts\lib\site-packages (from requests) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\manisha singh\onedri
ve\documents\python scripts\lib\site-packages (from requests) (2023.7.22)
```

In [10]:
```
import requests
from bs4 import BeautifulSoup
import pandas as pd
import bs4
```

In [11]:
```
headers = {'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537


Webpage = requests.get('https://www.nike.com/in/w/mens-shoes-nik1zy7ok', headers=he
```

In [12]:
```
Soup = BeautifulSoup (Webpage,'lxml')
```

In [13]:
```
def my_url(keyyword):
    keyword = keyyword.replace(' ', '%20')
    url = 'https://www.nike.com/w?q={}'.format(keyword)
    return url

def extract_data(obj):
    name = obj.find('div', 'product-card__title').text.strip()
    description = obj.find('div', 'product-card__subtitle').text.strip()
    color = obj.find('div','product-card__count-item').text
    url2 = obj.find('a', 'product-card__link-overlay').get('href')

    #use a try block incase because the price might be in the discount block
    try:
        price = obj.find('div', 'product-price us__styling is--current-price css-11
    except AttributeError:
        price = obj.find('div','product-price is--current-price css-1ydfahe').text

    try:
        old_price = obj.find('div','product-price us__styling is--striked-out css-0
    except AttributeError:
        old_price = ''

    data = requests.get(url2)
    soup2 = bs4.BeautifulSoup(data.text, 'html.parser')
    image = soup2.find('img', class_ = 'css-viwop1 u-full-width u-full-height css-m

    result = (name, description, color, price, old_price, image, url2)
    return result
```

```
def search(keyword):
    records = []
    url = my_url(keyword)

    response = requests.get(url)
    soup = bs4.BeautifulSoup(response.text, 'html.parser')
    soup_results = soup.findAll('div', class_ = 'product-card__body')
    for item in soup_results:
        record = extract_data(item)
        if record:
            records.append(record)

    # Convert list to dataframe
    columns_ = ['Name','Description','Available In', 'Price', 'Old Price', 'Image L
    df = pd.DataFrame(records, columns = columns_)

    return df

    # To store data in csv file, Run this
    # with open('Nike_Results.csv','w', newline = '', encoding = 'utf-8') as f:
    #     writer = csv.writer(f)
    #     writer.writerow(['Name','Description','Available In', 'Price', 'Old Price
    #     writer.writerows(records)
```

Testing

In [14]:
```
search('ball')
```

Out[14]:

	Name	Description	Available In	Price	Old Price	Image Li
0	Nike Skills	Kids' Basketball	1 Color	$15		https://static.nike.com/a/images/t_PDP_1280_v1
1	Premier League Academy	Soccer Ball	3 Colors	$27.97	$37	https://static.nike.com/a/images/t_PDP_1280_v1
2	Jordan Skills	Basketball	1 Color	$15		https://static.nike.com/a/images/t_PDP_1280_v1
3	Premier League Academy	Soccer Ball	1 Color	$37		https://static.nike.com/a/images/t_PDP_1280_v1
4	Jordan Playground 8P	Basketball	1 Color	$30		https://static.nike.com/a/images/t_PDP_1280_v1
5	Jordan Legacy 8P	Basketball	1 Color	$40		https://static.nike.com/a/images/t_PDP_1280_v1
6	Nike Elite All-Court 8P	Basketball	1 Color	$35		https://static.nike.com/a/images/t_PDP_1280_v1
7	Nike Academy	Soccer Ball	2 Colors	$27.97	$32	https://static.nike.com/a/images/t_PDP_1280_v1
8	Nike Elite Tournament	Basketball (Size 6 and 7)	1 Color	$55		https://static.nike.com/a/images/t_PDP_1280_v1
9	Premier League Flight	Soccer Ball	1 Color	$167		https://static.nike.com/a/images/t_PDP_1280_v1
10	NWSL Academy	Soccer Ball	1 Color	$27.97	$37	https://static.nike.com/a/images/t_PDP_1280_v1
11	Brazil Academy	Soccer Ball	1 Color	$30		https://static.nike.com/a/images/t_PDP_1280_v1
12	LeBron Playground 8P	Basketball	1 Color	$25.97	$30	https://static.nike.com/a/images/t_PDP_1280_v1
13	Nike Everyday Playground 8P	Graphic Basketball	1 Color	$22.97	$25	https://static.nike.com/a/images/t_PDP_1280_v1
14	Nike Club Elite Team	Soccer Ball	1 Color	$62		https://static.nike.com/a/images/t_PDP_1280_v1
15	Nike Skills	Soccer Ball	1 Color	$15.97	$20	https://static.nike.com/a/images/t_PDP_1280_v1
16	Nike All-Court	Volleyball	1 Color	$25.97	$30	https://static.nike.com/a/images/t_PDP_1280_v1
17	Paris Saint-Germain Strike	Soccer Ball	1 Color	$30		https://static.nike.com/a/images/t_PDP_1280_v1
18	Oklahoma	Nike College Mini Football	2 Colors	$18		https://static.nike.com/a/images/t_PDP_1280_v1

	Name	Description	Available In	Price	Old Price	Image Li
19	Nike Elite Championship 8P`	Basketball	1 Color	$75		https://static.nike.com/a/images/t_PDP_1280_v1
20	Liverpool FC Strike	Soccer Ball	1 Color	$32		https://static.nike.com/a/images/t_PDP_1280_v1
21	Nike Skills	Volleyball	1 Color	$15		https://static.nike.com/a/images/t_PDP_1280_v1
22	Nike Skills	Soccer Ball	1 Color	$14.97	$16	https://static.nike.com/a/images/t_PDP_1280_v1
23	Jordan Diamond	Basketball	1 Color	$50		https://static.nike.com/a/images/t_PDP_1280_v1

PROCESS AND EXPLANATIONS

In [16]:
```
# Import libraries
from selenium import webdriver
import bs4
import csv
import requests
```

In [17]:
```
# activate driver
driver = webdriver.Chrome()
```

In [18]:
```
def my_url(keyyword):
    keyword = keyyword.replace(' ', '%20')
    url = 'https://www.nike.com/w?q={}'.format(keyword)
    return url
```

In [19]:
```
url = my_url('blue shirt')
url
```

Out[19]:
```
'https://www.nike.com/w?q=blue%20shirt'
```

In [20]:
```
response = requests.get(url)
```

In [21]:
```
soup = bs4.BeautifulSoup(response.text, 'html.parser')
```

In [22]:
```
soup_results = soup.findAll('div', class_ = 'product-card__body')
len(soup_results)
```

Out[22]:
```
24
```

Getting the Name and Description

In [23]:
```
obj = soup_results[13]
name = obj.find('div', 'product-card__title').text.strip()
name
```

Out[23]:
```
'Nike Dri-FIT'
```

In [24]:
```
description = obj.find('div', 'product-card__subtitle').text.strip()
description
```

Out[24]:
```
"Big Kids' T-Shirt"
```

Getting the number of Color

In [26]:
```
color = obj.find('div','product-card__count-item').text
color
```

Out[26]:
```
'1 Color'
```

Getting the Price

In [27]:
```
try:
    price = obj.find('div', 'product-price us__styling is--current-price css-11s12a
except AttributeError:
    price = obj.find('div','product-price is--current-price css-1ydfahe').text
price
```

Out[27]:
```
'$21.97'
```

Getting prices with Discount

In [28]:
```
try:
    old_price = obj.find('div','product-price us__styling is--striked-out css-0').t
except AttributeError:
    old_price = ''
old_price
```

Out[28]:
```
'$28'
```

Getting the Urls and Images

In [29]:
```
url2 = obj.find('a', 'product-card__link-overlay').get('href')
url2
```

Out[29]:
```
'https://www.nike.com/t/dri-fit-big-kids-t-shirt-f0q1x5/FD0842-410'
```

In [30]:
```
data = requests.get(url2)
soup2 = bs4.BeautifulSoup(data.text, 'html.parser')
image = soup2.find('img', class_ = 'css-viwop1 u-full-width u-full-height css-m5dkr
```

In [31]:
```
image
```

Out[31]:
```
'https://static.nike.com/a/images/t_PDP_1280_v1/f_auto/d3b01368-0d28-4ca8-a1ea-415
3993892fc/dri-fit-big-kids-t-shirt-f0q1x5.png'
```

In [32]:
```
import csv
```

In [33]:
```python
# Change this to read data from the web scraped list
shoe_data = [{'name': 'Nike Air Force 1 Shadow', 'price': 'MRP : ₹ 8995.00'}]

def clean_price(price):
    return price.replace('MRP : ₹ ', '').strip()

filename = 'shoe_data.csv'

with open(filename, 'w', newline='', encoding='utf-8') as csvfile:
    writer = csv.writer(csvfile)
    for shoe in shoe_data:
        writer.writerow([shoe['name'], clean_price(shoe['price'])])

print(f"Data written to {filename}")
```

```
Data written to shoe_data.csv
```

In [ ]: