```
# Find the elements containing the details you want to scrape
restaurant_names = soup.find_all('h2', class_='restnt-name ellipsis')
cuisines = soup.find_all('span', class_='double-line-ellipsis')
locations = soup.find_all('span', class_='double-line-ellipsis')
ratings = soup.find_all('span', class_='rating-value')
image_urls = soup.find_all('img', class_='img-responsive')

# Create empty lists to store the scraped data
restaurant_list = []
cuisine_list = []
location_list = []
rating_list = []
image_url_list = []

# Extract the data from the elements and append them to the respective lists
for name in restaurant_names:
  restaurant_list.append(name.text.strip())

for cuisine in cuisines:
  cuisine_list.append(cuisine.text.strip())

for location in locations:
  location_list.append(location.text.strip())

for rating in ratings:
  rating_list.append(rating.text.strip())

for image in image_urls:
  image_url_list.append(image['src'])

# Create a dictionary from the lists
data = {
  'Restaurant Name': restaurant_list,
  'Cuisine': cuisine_list,
  'Location': location_list,
  'Ratings': rating_list,
  'Image URL': image_url_list
}

# Create a dataframe from the dictionary
df = pd.DataFrame(data)

# Print the dataframe
print(df)
```

```
Empty DataFrame
Columns: [Restaurant Name, Cuisine, Location, Ratings, Image URL]
Index: []
```

In [1]:

```
Import requests
From bs4 import BeautifulSoup
Import pandas as pd

# Step 1: Get the webpage
url = https://www.shine.com/
response = requests.get(url)

# Step 2: Enter search criteria and click the search button
Job_title = “Data Analyst”
Location = “Bangalore”
```

```
# Create a session to maintain cookies
Session = requests.Session()

# Get the search page
Search_url = https://www.shine.com/job-search/data-analyst-jobs-in-bangalore
Search_response = session.get(search_url)

# Step 3: Scrape the data for the first 10 jobs
Soup = BeautifulSoup(search_response.content, “html.parser”)
Job_listings = soup.find_all(“li”, class_=”search_listing”)

Data = []
For job_listing in job_listings[:10]:
  Job_title = job_listing.find(“h3”).text.strip()
  Job_location = job_listing.find(“span”, class_=”location”).text.strip()
  Company_name = job_listing.find(“span”, class_=”company_name”).text.strip()
  Experience_required = job_listing.find(“span”, class_=”exp”).text.strip()

  Data.append({
  “Job Title”: job_title,
  “Job Location”: job_location,
  “Company Name”: company_name,
  “Experience Required”: experience_required
  })

# Step 5: Create a dataframe of the scraped data
Df = pd.DataFrame(data)

# Print the dataframe
Print(df)
```

```
  Cell In[1], line 10
    Job_title = “Data Analyst”
                ^
SyntaxError: invalid character '“' (U+201C)
```

In [2]:

```
Import requests
From bs4 import BeautifulSoup
Import pandas as pd

# Step 1: Get the webpage
url = https://www.shine.com
response = requests.get(url)

# Step 2: Enter search criteria and click search button
Job_title = “Data Scientist”
Location = “Bangalore”
Payload = {
  “search_query”: job_title,
  “loc_query”: location
}
Response = requests.post(url, data=payload)

# Step 3: Scrape the data for the first 10 jobs
Soup = BeautifulSoup(response.content, “html.parser”)
Job_results = soup.find_all(“div”, class_=”result-display”)
Job_data = []
For result in job_results[:10]:
  Title = result.find(“h2”).text.strip()
```

```
  Company = result.find(“span”, class_=”company-name”).text.strip()
  Location = result.find(“span”, class_=”location”).text.strip()
  Job_data.append({“Job Title”: title, “Company Name”: company, “Location”: location})

# Step 4: Create a dataframe of the scraped data
Df = pd.DataFrame(job_data)

# Print the dataframe
Print(df)
```

```
  Cell In[2], line 10
    Job_title = “Data Scientist”
                ^
SyntaxError: invalid character '“' (U+201C)
```

In [3]:
```
Import requests
From bs4 import BeautifulSoup
Import pandas as pd

url = https://www.shine.com/
response = requests.get(url)

soup = BeautifulSoup(response.content, ‘html.parser’)

search_input = soup.find(‘input’, {‘id’: ‘txt_search’})
search_input[‘value’] = ‘Data Scientist’

search_button = soup.find(‘button’, {‘id’: ‘btn_search’})
response = requests.post(url, data={‘txt_search’: ‘Data Scientist’})

soup = BeautifulSoup(response.content, ‘html.parser’)
location_filter = soup.find(‘input’, {‘id’: ‘chk_location_1’})
location_filter[‘checked’] = True
salary_filter = soup.find(‘input’, {‘id’: ‘chk_salary_1’})
salary_filter[‘checked’] = True

job_listings = soup.find_all(‘div’, {‘class’: ‘w-100’})
data = []
for job in job_listings[:10]:
  title = job.find(‘h3’).text.strip()
  location = job.find(‘span’, {‘class’: ‘location’}).text.strip()
  company = job.find(‘span’, {‘class’: ‘company-name’}).text.strip()
  experience = job.find(‘span’, {‘class’: ‘exp’}).text.strip()
  data.append([title, location, company, experience])

df = pd.DataFrame(data, columns=[‘Job Title’, ‘Job Location’, ‘Company Name’, ‘Experie
```

```
  Cell In[3], line 8
    soup = BeautifulSoup(response.content, ‘html.parser’)
                                           ^
SyntaxError: invalid character '‘' (U+2018)
```

In [4]:
```
Import requests
From bs4 import BeautifulSoup

url = https://www.flipkart.com/
search_query = “sunglasses”
max_listings = 100
scraped_data = []
```

```
while len(scraped_data) < max_listings:
  response = requests.get(url)
  soup = BeautifulSoup(response.content, “html.parser”)

  # Find the search field and enter the search query
  Search_field = soup.find(“input”, attrs={“title”: “Search for products, brands and m
  Search_field[“value”] = search_query

  # Click the search icon
  Search_icon = soup.find(“button”, attrs={“type”: “submit”})
  Response = requests.post(url, data=search_icon.form)
  Soup = BeautifulSoup(response.content, “html.parser”)

  # Find the container holding the listings
  Listings_container = soup.find(“div”, attrs={“class”: “_1AtVbE”})

  # Extract the required attributes from each listing
  For listing in listings_container.find_all(“div”, attrs={“class”: “_2kHMtA”}):
  Brand = listing.find(“div”, attrs={“class”: “_2WkVRV”}).text
  Description = listing.find(“a”, attrs={“class”: “IRpwTa”}).text
  Price = listing.find(“div”, attrs={“class”: “_30jeq3 _1_WHN1”}).text

  Scraped_data.append({“Brand”: brand, “ProductDescription”: description, “Price”: pri

  If len(scraped_data) == max_listings:
  Break

  # Find the “Next” button and navigate to the next page
  Next_button = soup.find(“a”, attrs={“class”: “_1LKTO3”})
  If next_button:
  url = https://www.flipkart.com + next_button[“href”]
  else:
  break

# Print the scraped data
For data in scraped_data:
  Print(data)
```

```
  Cell In[4], line 5
    search_query = “sunglasses”
                   ^
SyntaxError: invalid character '“' (U+201C)
```

In [5]:

```
Import bs4
From bs4 import BeautifulSoup as bs
Import requests

Link=’https://www.flipkart.com/search?q=tv&as=on&as-show=on&otracker=AS_Query_Trending

Page = requests.get(link)

Page.content

Soup = bs(page.content, ‘html.parser’)
#it gives us the visual representation of data
Print(soup.prettify())

Name=soup.find(‘div’,class_=”_4rR01T”)
Print(name)
```

```
# to get just the name we will use the below code
Name.text

#get rating of a product
Rating=soup.find('div',class_="_3LWZlK")
Print(rating)
Rating.text

#get other details and specifications of the product
Specification=soup.find('div',class_="fMghEO")
Print(specification)
Specification.text

For each in specification:
    Spec=each.find_all('li',class_='rgWa7D')
    Print(spec[0].text)
    Print(spec[1].text)
    Print(spec[2].text)
    Print(spec[4].text)
    Print(spec[5].text)
    Print(spec[7].text)

#get price of the product
Price=soup.find('div',class_='_30jeq3 _1_WHN1')
Print(price)
Price.text

Products=[]             #List to store the name of the product
Prices=[]               #List to store price of the product
Ratings=[]              #List to store rating of the product
Apps = []               #List to store supported apps
Os = []                 #List to store operating system
Hd = []                 #List to store resolution
Sound = []              #List to sound output

For data in soup.findAll('div',class_='_3pLy-c row'):
        Names=data.find('div', attrs={'class':'_4rR01T'})
        Price=data.find('div', attrs={'class':'_30jeq3 _1_WHN1'})
        Rating=data.find('div', attrs={'class':'_3LWZlK'})
        Specification = data.find('div', attrs={'class':'fMghEO'})

        For each in specification:
            Col=each.find_all('li', attrs={'class':'rgWa7D'})
            App =col[0].text
            Os_ = col[1].text
            Hd_ = col[2].text
            Sound_ = col[3].text
        Products.append(names.text) # Add product name to list
        Prices.append(price.text) # Add price to list
        Apps.append(app)# Add supported apps specifications to list
        Os.append(os_) # Add operating system specifications to list
        Hd.append(hd_) # Add resolution specifications to list
        Sound.append(sound_) # Add sound specifications to list
        Ratings.append(rating)

#printing the length of list
Print(len(products))
Print(len(ratings))
Print(len(prices))
Print(len(apps))
```

```
Print(len(sound))
Print(len(os))
Print(len(had))

Import pandas as pd
Df=pd.DataFrame({'Product Name':products,'Supported_apps':apps,'sound_system':sound,'C
Df.head(10)
```

```
  Cell In[5], line 5
    Link='https://www.flipkart.com/search?q=tv&as=on&as-show=on&otracker=AS_Query_Tre
ndingAutoSuggest_8_0_na_na_na&otracker1=AS_Query_TrendingAutoSuggest_8_0_na_na_na&as-
pos=8&as-type=TRENDING&suggestionId=tv&requestId=9c9fa553-b7e5-454b-a65b-bbb7a9c74a29
          ^
SyntaxError: invalid character '’' (U+2019)
```

In [6]:

```
# let's first connect to the web driver
Driver = webdriver.Chrome

url=https://www.flipkart.com/
driver.get(url)

# finding element for job search bar
Search_g= driver.find_element_by_xpath("//input[@type='text']")
Search_g

<selenium.webdriver.remote.webelement.WebElement (session="0003341c3c278ca61eb98703d16

# write on search bar
Search_g.send_keys('sneakers')
Search_btn=driver.find_element_by_xpath("//button[@class='L0Z3Pu']")
Search_btn

<selenium.webdriver.remote.webelement.WebElement (session="0003341c3c278ca61eb98703d16

Search_btn=driver.find_element_by_class_name('L0Z3Pu')
Search_btn.click()

B_name=[]
Price=[]
P_desc=[]
Discount=[]

For I in range(3):
    B_name=driver.find_elements_by_xpath("//div[@class='_2WkVRV']")
    P_desc=driver.find_elements_by_xpath("//a[@class='IRpwTa']")
    Price =driver.find_elements_by_xpath("//div[@class='_25b18c']")
    Discount=driver.find_elements_by_xpath("//div[@class='_3Ay6Sb']")

    For j  in b_name:
        B_name.append(j.text)
    B_name[:100]

      For k in p_desc:
        P_desc.append(k.text)
    P_desc[:100]


    For l in price:
        Price.append(l.text)
    Price[:100]
```

```
    For t in discount:
        Discount.append(t.text)
    Discount[:100]
Print(len(B_name[:100])),print(len(Price[:100])),print(len(P_desc[:100])),print(len(Di
100
100
100
100
(None, None, None, None)

Sun_gl=pd.DataFrame({})
Sun_gl['Brand_name']=B_name[:100]
Sun_gl['P_price']=Price[:100]
Sun_gl['Pr_desc']=P_desc[:100]
Sun_gl['P_discount']=Discount[:100]
Sun_gl

Driver.close()
```

```
  Cell In[6], line 8
    Search_g= driver.find_element_by_xpath("//input[@type='text']")
                                           ^
SyntaxError: invalid character '"' (U+201C)
```

In [7]:

```
# let's first connect to the web driver
Driver = webdriver.Chrome

url=" https://www.amazon.in "
driver.get(url)

# finding element for job search bar
Search_g= driver.find_element_by_xpath("//input[@type='text']")
Search_g

<selenium.webdriver.remote.webelement.WebElement (session="a2ad6e875a1461b017453195fbc

# write on search bar
Search_g.send_keys('Laptop')

Search_btn=driver.find_element_by_xpath("//input[@id='nav-search-submit-button']")
Search_btn

<selenium.webdriver.remote.webelement.WebElement (session="a2ad6e875a1461b017453195fbc
Search_btn=driver.find_element_by_xpath("//input[@id='nav-search-submit-button']")
Search_btn.click()

Title=[]
Price=[]
Rating=[]

For I in range(3):
    B_name=driver.find_elements_by_xpath("//div[@class='_2WkVRV']")
    P_desc=driver.find_elements_by_xpath("//a[@class='IRpwTa']")
    Price =driver.find_elements_by_xpath("//div[@class='_25b18c']")

Title=[]
Price=[]
Rating=[]
```

```
For I in range(3):
    B_name=driver.find_elements_by_xpath(“//div[@class=’_2WkVRV’]”)
    P_desc=driver.find_elements_by_xpath(“//a[@class=’IRpwTa’]”)
    Price =driver.find_elements_by_xpath(“//div[@class=’_25b18c’]”)

    For j  in b_name:
        Title.append(j.text)
    Title[:100]

     For k in p_desc:
        P_desc.append(k.text)
    P_desc[:100]


    For l in price:
        Price.append(l.text)
    Price[:100]
```

```
  Cell In[7], line 4
    url=” https://www.amazon.in “
        ^
SyntaxError: invalid character '”' (U+201D)
```

In [8]:

```
From selenium import webdriver
From selenium.webdriver.common.by import By

Driver = webdriver.Chrome()

Driver.get(https://www.azquotes.com/)

Top_quotes_button = driver.find_element(By.LINK_TEXT, “Top Quotes”)
Top_quotes_button.click()

Quotes = driver.find_elements(By.CSS_SELECTOR, “.title a”)
Authors = driver.find_elements(By.CSS_SELECTOR, “.author a”)
Types = driver.find_elements(By.CSS_SELECTOR, “.kw-box a”)

For quote, author, quote_type in zip(quotes, authors, types):
Print(“Quote:”, quote.text)
Print(“Author:”, author.text)
Print(“Type of Quote:”, quote_type.text)
Print()

Driver.quit()
```

```
  Cell In[8], line 8
    Top_quotes_button = driver.find_element(By.LINK_TEXT, “Top Quotes”)
                                                           ^
SyntaxError: invalid character '“' (U+201C)
```

In [9]:

```
From selenium import webdriver
Import pandas as pd

Driver = webdriver.Chrome(‘path_to_chromedriver’)

Driver.get(‘https://www.jagranjosh.com/’)

Gk_option = driver.find_element_by_link_text(‘GK’)
Gk_option.click()
```

```
Pm_option = driver.find_element_by_link_text(‘List of all Prime Ministers of India’)
Pm_option.click()

Data = []
Table = driver.find_element_by_xpath(‘//table[@class=”table4”]’)
Rows = table.find_elements_by_tag_name(‘tr’)
For row in rows:
  Cols = row.find_elements_by_tag_name(‘td’)
  If len(cols) == 4:
  Name = cols[0].text
  Born_dead = cols[1].text
  Term_of_office = cols[2].text
  Remarks = cols[3].text
  Data.append([name, born_dead, term_of_office, remarks])

Df = pd.DataFrame(data, columns=[‘Name’, ‘Born-Dead’, ‘Term of Office’, ‘Remarks’])

Driver.quit()
```

```
  Cell In[9], line 4
    Driver = webdriver.Chrome(‘path_to_chromedriver’)
                              ^
SyntaxError: invalid character '‘' (U+2018)
```

In [10]:

```
From selenium import webdriver
Import pandas as pd

# Step 1: Get the webpage
Driver = webdriver.Chrome(‘path_to_chromedriver’)  # Replace ‘path_to_chromedriver’ wi
Driver.get(‘https://www.motor1.com/’)

# Step 2: Type in the search bar
Search_bar = driver.find_element_by_id(‘search-input’)
Search_bar.send_keys(’50 most expensive cars’)
Search_bar.submit()

# Step 3: Click on the link
Link = driver.find_element_by_link_text(’50 Most Expensive Cars in the World’)
Link.click()

# Step 4: Scrape the data and create a dataframe
Car_names = driver.find_elements_by_xpath(‘//div[@class=”article-content”]/h3’)
Car_prices = driver.find_elements_by_xpath(‘//div[@class=”article-content”]/p’)

Data = []
For name, price in zip(car_names, car_prices):
  Data.append([name.text, price.text])

Df = pd.DataFrame(data, columns=[‘Car Name’, ‘Price’])
Print(df)

Driver.quit()
```

```
  Cell In[10], line 5
    Driver = webdriver.Chrome(‘path_to_chromedriver’)  # Replace ‘path_to_chromedrive
r’ with the actual path to your ChromeDriver executable
                              ^
SyntaxError: invalid character '‘' (U+2018)
```

In [ ]: