web-scrapping

April 17, 2023

[75]:
```
import requests
import pandas
from bs4 import BeautifulSoup

resp=requests.get("https://www.bikewale.com/new-bike-search/
 ↪best-bikes-above-2-lakh/") #requesting permission from website
print(resp)                          #printing response given by website(200␣
 ↪means granted)

sp=BeautifulSoup(resp.content,"html.parser")   #BeautifulSoup library, that␣
 ↪scraps data from website
names=sp.find_all('a',class_="bw-ga")
name=[]
for i in names[0:60]:
    temp=i['title']
    name.append(temp)

Name=[]
for i in name[0:60:2]:
    Name.append(i)
#Extracted names

prices=sp.find_all('span',class_="font18")
price=[]
for i in prices[0:30]:
    temp=i.text+" Onwards"
    price.append(temp)
#Extracted prices

engine_spec=sp.find_all('div',class_="text-xt-light-grey font14␣
 ↪margin-bottom15")
engine_specs=[]
for i in engine_spec[0:30]:
    temp=i.text
    engine_specs.append(temp)
#Extracted specifications
```

```
launch_dates=sp.find_all('span',class_="badge")
launch_date=[]
Launch_Date=[]
for i in launch_dates[0:60]:
    temp=i.text
    launch_date.append(temp)

for i in launch_date[1:60:2]:
    Launch_Date.append(i)
#Extracted launch dates

ratings=sp.find_all("span",class_="font11 text-xt-light-grey inline-block␣
 ↪padding-left3")
reviews=[]
for i in ratings[0:30]:
    temp=i.get_text()
    reviews.append(temp.strip())
#Extracted reviews

Rating=sp.find_all('span',class_="rate-star-color")
Ratings=[]
for i in Rating[0:30]:
    temp=i.get_text()
    Ratings.append(temp)
#Extracted ratings

Links=sp.find_all('a',class_="bw-ga")
Link=[]
for i in Links[0:30]:
    temp="https://www.bikewale.com"+i["href"]
    Link.append(temp)
#Extracted link of vehicle

data={"Model"    :    Name,
      "Price"    :    price,
      "Engine Specifications"  :  engine_specs,
      "Launch Date" :  Launch_Date,
      "Reviews"    :   reviews,
      "Rating"    :   Ratings,
      "Check out link"   :  Link
}                                #Assigning key for each type of data
Dframe=pandas.DataFrame(data)    #converting data into organized tabular format␣
 ↪using PANDAS
Dframe.to_csv("Latest_Bikes.csv")     # Data that is collected is stored in a␣
 ↪csv format.
```

```
<Response [200]>
```

[ ]: