SAMSUNG SMARTPHONES WEB SCRAPING

Project for python module

Institute Name- Itvedant Education Pvt. Ltd
Name- Kafil Irfan Khan

Email Address- kafilk286@gmail.com

Date of Submission- 4 October 2023

Description:

Samsung smartphones web scraping project involves extracting data from websites by automating the process using scripts or software. This data can be used for various purposes, such as data analysis, research, reporting, or populating databases.

This database contains 4 tables:

1. phone name

2. Phone_price

3. discount

4. rating

Project Steps:

1. **Identify Target Websites**:

    • Determine the websites from which you want to scrape data. Ensure that the websites allow web scraping or have a robots.txt file that permits scraping.

2. **Define Data Requirements**:

    • Specify the exact data you need to scrape. This could include text, images, links, product details, prices, reviews, etc.

3. **Choose a Web Scraping Tool/Programming Language**:

    • Decide whether you will use a web scraping library like BeautifulSoup (Python), Scrapy, or a headless browser automation tool like Puppeteer (JavaScript).

4. **Develop Web Scraping Code**:

- Write code to navigate the target website(s), locate relevant HTML elements, and extract the desired data. Handle any pagination, authentication, or CAPTCHA challenges if needed.

5. **Data Storage**:

- Decide how and where you will store the scraped data. Common options include CSV files, databases (SQL or NoSQL), or cloud storage services.

Project Benefits:

- Access to valuable data for analysis or research.

- Automation of data collection, saving time and effort.

- Stay up-to-date with changes on the target websites.

Project Risks:

- Legal issues related to web scraping.

- Technical challenges due to website changes.

Boat smart watch web scraping coding command:

(1)Access boat web site by using import requests URL and BeautifulSoup library.

Output:

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
url="https://www.flipkart.com/search?q=samsung+phone&sid=tyy%2C4io&as=on&as-show=on&otracker=AS_QueryStore_OrganicAutoSuggest_1_1
phone_name=[]
r=requests.get(url)
htmlcontent=r.content
soup=BeautifulSoup(htmlcontent,"html.parser")
print(soup)
#soup.prettify
```

```
<!DOCTYPE html>
<html lang="en"><head><link href="https://rukminim2.flixcart.com" rel="preconnect"/><link href="//static-assets-web.flixcart.
com/fk-p-linchpin-web/fk-cp-zion/css/app_modules.chunk.905c37.css" rel="stylesheet"/><link href="//static-assets-web.flixcar
t.com/fk-p-linchpin-web/fk-cp-zion/css/app.chunk.ccbde3.css" rel="stylesheet"/><meta content="text/html; charset=utf-8" http-
equiv="Content-type"/><meta content="IE=Edge" http-equiv="X-UA-Compatible"/><meta content="102988293558" property="fb:page_i
d"/><meta content="658873552,624500995,100000233612389" property="fb:admins"/><meta content="noodp" name="robots"/><link href
="https:///www/promos/new/20150528-140547-favicon-retina.ico" rel="shortcut icon"/><link href="/osdd.xml?v=2" rel="search" ty
pe="application/opensearchdescription+xml"/><meta content="website" property="og:type"/><meta content="Flipkart.com" name="og
_site_name" property="og:site_name"/><link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes="57x57"/><link hre
f="/apple-touch-icon-72x72.png" rel="apple-touch-icon" sizes="72x72"/><link href="/apple-touch-icon-114x114.png" rel="apple-t
ouch-icon" sizes="114x114"/><link href="/apple-touch-icon-144x144.png" rel="apple-touch-icon" sizes="144x144"/><link href="/a
pple-touch-icon-57x57.png" rel="apple-touch-icon"/><meta content="app" name="twitter:card"/><meta content="@flipkart" name="t
witter:site"/><meta content="@flipkart" name="twitter:creator"/><meta content="Samsung Phone- Buy Products Online at Best Pri
ce in India - All Categories | Flipkart.com" name="twitter:title"/><meta content="Shop for electronics, apparels &amp; more u
sing our Flipkart app Free shipping &amp; COD." name="twitter:description"/><meta content="in" name="twitter:app:country"/><m
eta content="Flipkart" name="al:ios:app_name"/><meta content="742044692" name="al:ios:app_store_id"/><meta content="Flipkart"
name="twitter:app:name:iphone"/><meta content="742044692" name="twitter:app:id:iphone"/><meta content="http://dl.flipkart.co
m/dl/home?" name="twitter:app:url:iphone"/><meta content="Flipkart" name="twitter:app:name:ipad"/><meta content="742044692" n
ame="twitter:app:id:ipad"/><meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:ipad"/><meta content="Flipka
```

(2)Then access names of phones by using this class.

Output:

```
names=soup.find_all("div",class_="_4rR01T")
print(len(names))
for i in range (0,18):
    phone_name.append(names[i].get_text())
print(phone_name)
```

```
24
['SAMSUNG Galaxy F04 (Opal Green, 64 GB)', 'SAMSUNG Galaxy F04 (Jade Purple, 64 GB)', 'SAMSUNG Galaxy F13 (Waterfall Blue, 64 G
B)', 'SAMSUNG Galaxy F13 (Waterfall Blue, 128 GB)', 'SAMSUNG Galaxy F13 (Nightsky Green, 64 GB)', 'SAMSUNG Galaxy F13 (Sunrise
Copper, 64 GB)', 'SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)', 'SAMSUNG Galaxy F13 (Nightsky Green, 128 GB)', 'SAMSUNG Galax
y F14 5G (OMG Black, 128 GB)', 'SAMSUNG Galaxy M14 5G (Smoky Teal, 128 GB)', 'SAMSUNG Galaxy M14 5G (Icy Silver, 128 GB)', 'SAM
SUNG Galaxy F13 (Sunrise Copper, 128 GB)', 'SAMSUNG Galaxy F23 5G (Aqua Blue, 128 GB)', 'SAMSUNG Galaxy F34 5G (Electric Black,
128 GB)', 'SAMSUNG Galaxy F23 5G (Copper Blush, 128 GB)', 'SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)', 'SAMSUNG Galaxy M14
5G (ICY Silver, 128 GB)', 'SAMSUNG Galaxy F14 5G (GOAT Green, 128 GB)']
```

(3) Then access phone price detail by using this class.

Output:

```
price=[]
phone_price=soup.find_all(class_="_30jeq3 _1_WHN1")
for i in range (0,18):
    price.append(phone_price[i].get_text())
print(price)
```

```
['₹6,499', '₹6,499', '₹9,199', '₹10,199', '₹9,199', '₹9,199', '₹12,490', '₹10,199', '₹12,490', '₹13,865', '₹14,195', '₹10,199',
'₹15,499', '₹16,499', '₹15,499', '₹11,490', '₹13,722', '₹12,490']
```

(4) Then access on phone discount by using this class.

Output:

```
discount=[]
phone_discount=soup.find_all(class_="_3Ay6Sb")
for i in range (0,18):
    discount.append(phone_discount[i].get_text())
print(discount)
```

```
['43% off', '43% off', '38% off', '40% off', '38% off', '38% off', '32% off', '40% off', '32% off', '27% off', '25% off', '40%
off', '35% off', '32% off', '35% off', '34% off', '23% off', '32% off']
```

(5) Then access rating of phone by using this class.

Output:

```
rating=[]
phone_rating=soup.find_all(class_="_3LWZlK")
for i in range (0,18):
    rating.append(phone_rating[i].get_text())
print(rating)
```

```
['4.2', '4.2', '4.3', '4.3', '4.3', '4.3', '4.2', '4.3', '4.2', '4.2', '4.2', '4.3', '4.3', '4.2', '4.3', '4.2', '4.2', '4.2']
```

(6) Then arrange properly in column all above data by using import pandas library.

```
import pandas as pd
samsung_phone=pd.DataFrame({'phone_name':phone_name,'phone_price':price,'phone_discount':discount,'phone_rating':rating})
samsung_phone
```

Output:

	phone_name	phone_price	phone_discount	phone_rating
0	SAMSUNG Galaxy F04 (Opal Green, 64 GB)	₹6,499	43% off	4.2
1	SAMSUNG Galaxy F04 (Jade Purple, 64 GB)	₹6,499	43% off	4.2
2	SAMSUNG Galaxy F13 (Waterfall Blue, 64 GB)	₹9,199	38% off	4.3
3	SAMSUNG Galaxy F13 (Waterfall Blue, 128 GB)	₹10,199	40% off	4.3
4	SAMSUNG Galaxy F13 (Nightsky Green, 64 GB)	₹9,199	38% off	4.3
5	SAMSUNG Galaxy F13 (Sunrise Copper, 64 GB)	₹9,199	38% off	4.3
6	SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)	₹12,490	32% off	4.2
7	SAMSUNG Galaxy F13 (Nightsky Green, 128 GB)	₹10,199	40% off	4.3
8	SAMSUNG Galaxy M14 5G (Smoky Teal, 128 GB)	₹13,865	26% off	4.2
9	SAMSUNG Galaxy F14 5G (OMG Black, 128 GB)	₹12,490	32% off	4.2
10	SAMSUNG Galaxy F13 (Sunrise Copper, 128 GB)	₹10,199	40% off	4.3
11	SAMSUNG Galaxy M14 5G (Icy Silver, 128 GB)	₹14,195	25% off	4.2
12	SAMSUNG Galaxy F23 5G (Aqua Blue, 128 GB)	₹15,499	35% off	4.3
13	SAMSUNG Galaxy M14 5G (Berry Blue, 128 GB)	₹14,028	26% off	4.2
14	SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)	₹11,490	34% off	4.2
15	SAMSUNG Galaxy F14 5G (GOAT Green, 128 GB)	₹12,490	32% off	4.2
16	SAMSUNG Galaxy M14 5G (Smoky Teal, 128 GB)	₹13,373	25% off	4.2
17	SAMSUNG Galaxy M14 5G (ICY Silver, 128 GB)	₹13,722	23% off	4.2

(7) Then store above all data in CSV file and again open this file in application.

```python
samsung_phone.to_csv("samsung phone names and prices.csv",index= False)
```

```python
mobile_info=pd.read_csv("samsung phone names and prices.csv")
print(mobile_info)
```

Output:

	phone_name	phone_price	phone_discount	phone_rating
0	SAMSUNG Galaxy F04 (Opal Green, 64 GB)	₹6,499	43% off	4.2
1	SAMSUNG Galaxy F04 (Jade Purple, 64 GB)	₹6,499	43% off	4.2
2	SAMSUNG Galaxy F13 (Waterfall Blue, 64 GB)	₹9,199	38% off	4.3
3	SAMSUNG Galaxy F13 (Waterfall Blue, 128 GB)	₹10,199	40% off	4.3
4	SAMSUNG Galaxy F13 (Nightsky Green, 64 GB)	₹9,199	38% off	4.3
5	SAMSUNG Galaxy F13 (Sunrise Copper, 64 GB)	₹9,199	38% off	4.3
6	SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)	₹12,490	32% off	4.2
7	SAMSUNG Galaxy F13 (Nightsky Green, 128 GB)	₹10,199	40% off	4.3
8	SAMSUNG Galaxy M14 5G (Smoky Teal, 128 GB)	₹13,865	26% off	4.2
9	SAMSUNG Galaxy F14 5G (OMG Black, 128 GB)	₹12,490	32% off	4.2
10	SAMSUNG Galaxy F13 (Sunrise Copper, 128 GB)	₹10,199	40% off	4.3
11	SAMSUNG Galaxy M14 5G (Icy Silver, 128 GB)	₹14,195	25% off	4.2
12	SAMSUNG Galaxy F23 5G (Aqua Blue, 128 GB)	₹15,499	35% off	4.3
13	SAMSUNG Galaxy M14 5G (Berry Blue, 128 GB)	₹14,028	26% off	4.2
14	SAMSUNG Galaxy F14 5G (B.A.E. Purple, 128 GB)	₹11,490	34% off	4.2
15	SAMSUNG Galaxy F14 5G (GOAT Green, 128 GB)	₹12,490	32% off	4.2
16	SAMSUNG Galaxy M14 5G (Smoky Teal, 128 GB)	₹13,373	25% off	4.2
17	SAMSUNG Galaxy M14 5G (ICY Silver, 128 GB)	₹13,722	23% off	4.2

Conclusion :

In this web scraping project, we set out to collect valuable information from websites automatically. We successfully achieved this goal using tools like Python, BeautifulSoup, and Selenium. We collected data from websites which belonging to our project.

Along the way, we encountered some challenges like handling website changes and respecting website terms of use. However, we overcame these challenges effectively.

Once gathering the data, we cleaned and organized it to make it useful. We stored the data in [mention where you stored it in required file extension. Although we didn't perform detailed data analysis in this demo, the possibilities are endless.

This data can be used for various purposes, such as research, decision-making, or creating informative visualizations. In the future, we could automate data updates and expand our sources to gather even more valuable information.

In conclusion, web scraping is a powerful tool for extracting data from the web, and with responsible use, it can provide valuable insights and information.

THANK YOU 😊