

Project For Python Module

Web Scrapping iPhone 15 Insights from Flipkart

Institute Name:- ITvedant Education Pvt.Ltd

Name:- Arun Shankar Ovhal

Email Address:-arun77ovhal@gmail.com

Date Of Submission:-26th Nov 2023

AIM OF THE PROJECT

The aim of this project is to create a Python web scraping script to extract specific details of the iPhone 15 from the Flipkart website. Using BeautifulSoup and Requests libraries, the objective is to gather essential information such as product names, prices, customer reviews, ratings, and comments associated with the iPhone 15 listings



Description:-

This Python-based web scraping project focuses on collecting structured data from Flipkart's website, specifically targeting the iPhone 15 product details. The script employs Requests to fetch the HTML content of the product page and BeautifulSoup to parse and extract relevant information.

The data extracted includes:

- **Product Details:** Name and price of the iPhone 15.
- **Customer Review:** Capturing customer names, ratings, reviews, and comments.
- **Product Popularity:** Analyzing the number of likes and dislikes for the product.

The extracted data is then structured into a Pandas DataFrame and exported to a CSV file for further analysis and insights.

This database contains 4 tables:

- 1.CName
- 2.Reveiw
- 3.Star
- 4.Comment

Project Steps:-

- **URL Retrieval:** You've defined the URL for the iPhone 15 product page on Flipkart.
- **HTTP Request:** Using the requests library, you've fetched the HTML content from the URL.
- **Parsing HTML:** Utilizing BeautifulSoup, you've parsed the HTML content to navigate and extract specific elements.
- **Data Extraction:** Extracted data includes product prices, customer names, comments, reviews, ratings, likes, and dislikes.
- **Data Manipulation:** You've stored extracted data into lists and a pandas DataFrame.
- **Data Export:** Exported the DataFrame containing the extracted data to a CSV file.

Project Benefits:-

- **Real-time Insights:** Gather up-to-date iPhone 15 prices, reviews, and trends from Flipkart for informed decision-making.
- **Competitive Advantage:** Analyze pricing strategies and customer sentiments, aiding in competitive analysis.
- **Efficiency and Automation:** Streamline data collection processes, enabling efficient market monitoring and operational enhancements.

Project Risk:-

- The project faces legal risks concerning website terms, potential data inaccuracies from changing website structures, and ethical concerns regarding data usage.
- Technical challenges like IP blocking, rate limiting, and data interpretation issues could affect the reliability of scraped information and disrupt the scraping process.

iPhone15 Web Scraping from Flipkart coding command:

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

```
import requests
from bs4 import BeautifulSoup
URL="https://www.flipkart.com/apple-iphone-15-blue-128-gb/p/itmfb14ef54f645d?pid=MOBGTAGPAQNVFZZY&lid=LSTMOBGTAGPAQNVFZZYSCIIOB&m"

r=requests.get(URL)
htmlcontent=r.content
print(htmlcontent)

b'<!doctype html><html lang="en"><head><link href="https://rukminim2.flixcart.com" rel="preconnect"><link rel="stylesheet" h
ref="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app_modules.chunk.905c37.css"/><link rel="stylesheet" h
ref="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app.chunk.615ed9.css"/><meta http-equiv="Content-typ
e" content="text/html; charset=utf-8"/><meta http-equiv="X-UA-Compatible" content="IE=Edge"/><meta property="fb:page_id" cont
ent="102988293558"/><meta property="fb:admins" content="658873552,624500995,100000233612389"/><meta name="robots" content="no
odp"/><link rel="shortcut icon" href="https://www.promos/new/20150528-140547-favicon-retina.ico"/><link type="application/op
ensearchdescription+xml" rel="search" href="/osdd.xml?v=2"/><meta property="og:type" content="website"/><meta name="og_site_n
ame" property="og:site_name" content="Flipkart.com"/><link rel="apple-touch-icon" sizes="57x57" href="/apple-touch-icon-57x5
7.png"/><link rel="apple-touch-icon" sizes="72x72" href="/apple-touch-icon-72x72.png"/><link rel="apple-touch-icon" sizes="11
4x114" href="/apple-touch-icon-114x114.png"/><link rel="apple-touch-icon" sizes="144x144" href="/apple-touch-icon-144x144.pn
g"/><link rel="apple-touch-icon" href="/apple-touch-icon-57x57.png"/><meta name="twitter:card" content="app"><meta name="twit
ter:site" content="@flipkart"><meta name="twitter:creator" content="@flipkart"><meta name="twitter:title" content="APPLE iPho
ne 15 (Blue, 128 GB) | Flipkart"><meta name="twitter:description" content="Shop for electronics, apparels & more using our Fl
ipkart app Free shipping & COD."><meta property="twitter:image" content="http://rukmini1.flixcart.com/image/300/300/xif0q/mob
ile/k/1/l/-original-imagtc5fz9spysyk.jpeg"/><meta name="twitter:app:country" content="in"><meta name="al:ios:app_name" conten
t="Flipkart"><meta name="al:ios:app_store_id" content="742044692"><meta name="twitter:app:name:iphone" content="Flipkart"><me
ta name="twitter:app:id:iphone" content="742044692"><meta name="twitter:app:url:iphone" content="http://dl.flipkart.com/dl/ho
me?><meta name="twitter:app:name:ipad" content="Flipkart"><meta name="twitter:app:id:ipad" content="742044692"><meta name="t
witter:app:url:ipad" content="http://dl.flipkart.com/dl/home?><meta name="twitter:app:name:googleplay" content="Flipkart"><m
...

soup=BeautifulSoup(htmlcontent,"html.parser")
print(soup)
print(soup.prettify)
```

Output:

```
<!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.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app.chunk.615ed9p.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_id><meta content=658873552,624500995,100000233612389 property=fb:admins><meta content=nooodp name=robots><link href=https://www.promos/new/20150528-140547-favicon-retina.ico rel=shortcut icon><link href=/osdd.xml?v=2 rel=search_tpe=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 href=/apple-touch-icon-72x72.png rel=apple-touch-icon sizes=72x72><link href=/apple-touch-icon-114x114.png rel=apple-touch-icon sizes=114x114><link href=/apple-touch-icon-144x144.png rel=apple-touch-icon sizes=144x144><link href=/apple-touch-icon-57x57.png rel=apple-touch-icon><meta content=app name=twitter:card><meta content=@flipkart name=twitter:site><meta content=@flipkart name=twitter:creator><meta content=APPLE iPhone 15 (Blue, 128 GB) | Flipkart name=twitter:title><meta content=Shop for electronics, apparels & more using our Flipkart app Free shipping & COD. name=twitter:description><meta content=http://rukmini1.flixcart.com/image/300/300/xif0q/mobile/k/l/l/-original-image15cf29spysyk.jpeg property=twitter:image><meta content=in name=twitter:app:country><meta 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.com/dl/home? name=twitter:app:url:iphon e><meta content=Flipkart name=twitter:app:name:ipad><meta content=742044692 name=twitter:app:id:ipad><meta content=
```

(2) To access iPhone 15 price by using this class.

```
price=soup.find_all(class_="_30jeq3 _16Jk6d")
price
```

[<div class=" 30]eq3 16]k6d">₹79,900</div>]

```
price[0].get_text()
```

Output:

['₹79,900']

(3) To access iPhone 15 Customer Name by using this class.

```
: name=soup.find_all(class_="_2sc7ZR _2V5EHH")
name=list(name)
print(type(name))
print(name)
len(name)
```

```
<class 'list'>
[<p class="_2sc7ZR _2V5EHH">Mousam Guha Roy</p>, <p class="_2sc7ZR _2V5EHH">Ajin V</p>, <p class="_2sc7ZR _2V5EHH">Prithivi Bo
ruah</p>, <p class="_2sc7ZR _2V5EHH">Ankit Verma </p>, <p class="_2sc7ZR _2V5EHH">Suman Jain</p>, <p class="_2sc7ZR _2V5EHH">R
ahul Shedge</p>, <p class="_2sc7ZR _2V5EHH">Jayanta Roy</p>, <p class="_2sc7ZR _2V5EHH">Sheetla Prasad Maurya</p>, <p class="_
2sc7ZR _2V5EHH">Saurabh Gothwad</p>, <p class="_2sc7ZR _2V5EHH">Praveen Kumar</p>]
```

10

```

cust_name=[]
for i in range(0,len(name)):
    cust_name.append(name[i].get_text())
cust_name

```

Output:

```
['Mousam Guha Roy',  
'Ajin V',  
'Prithivi Boruah',  
'Ankit Verma ',  
'Suman Jain',  
'Rahul Shedge',  
'Jayanta Roy',  
'Sheetla Prasad Maurya',  
'Saurabh Gothwad',  
'Praveen Kumar']
```

(4) To access iPhone 15 Customer Review by using this class.

```
comment=soup.find_all(class_='t-ZTKy')  
comment  
  
[<div class="t-ZTKy"><div><div class="">Very nice</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">High quality camera 📷</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Camera Quality Is Improved Loving It</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Awesome camra quality superb camra 😊</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Simply Awesome 🥰</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Totally happy!<br>Camera 5<br>Battery 5<br>Display 5<br>Design 5<br></div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Nice 😊😊😊😊😊</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Best mobile phone<br>Camera quality is very nice<br>Battery backup is very good<br>Sound quality is amazing.</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">Simply premium.</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>,<br><div class="t-ZTKy"><div><div class="">This is my first iphone ever used loving it already this is worth it 🥰</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>]  
  
comment_review=[]  
for i in range(0,len(comment)):  
    comment_review.append(comment[i].get_text())  
comment_review
```

Output:

```
['Very niceREAD MORE',  
'High quality camera 📷 READ MORE',  
'Camera Quality Is Improved Loving ItREAD MORE',  
'Awesome camra quality superb camra 😊 READ MORE',  
'Simply Awesome 🥰 READ MORE',  
'Totally happy!Camera 5Battery 5 Display 5Design 5READ MORE',  
'Nice 😊😊😊😊😊 READ MORE',  
'Best mobile phoneCamera quality is very nice Battery backup is very good Sound quality is amazing.READ MORE',  
'Simply premium.READ MORE',  
'This is my first iphone ever used loving it already this is worth it 🥰 READ MORE']
```

(5) To access iPhone 15 Customer Review star by using this class.

[illegible]

Output:

```
['4', '5', '5', '4', '5', '5', '5', '4', '5', '5']
```

(5) To access iPhone 15 Customer Likes & Dislikes by using this class with the segregation.

```
product_likedislike=soup.find_all("span",class="_3c3Px5")
product_likedislike
```

```
product_likedislike1=[]
for i in range(0,len(product_likedislike)):
    product_likedislike1.append(product_likedislike[i].get_text())
product_likedislike1
```

```
like=[]
dislike=[]
for i in range(0,len(product_likedislike1)):
    if i%2==0:
        like.append(product_likedislike1[i])
    elif i%2!=0:
        dislike.append(product_likedislike1[i])
print(like)
print(dislike)
```

Output:

['88', '925', '550', '283', '560', '75', '559', '64', '104', '301', '44', '84', '41', '1', '1', '2', '3', '0', '0', '0']
['4', '145', '103', '74', '161', '15', '171', '13', '26', '91', '17', '49', '38', '0', '0', '0', '2', '0', '0', '0']

(6) To arrange properly in column all above data by using import Pandas library.

```
import pandas as pd

df=pd.DataFrame({"CName":cust_name,"Review":comment_review,"star":review_star,"comment":content_review},index=range(101,111))
print(df)
```

Output:

	CName	Review \
101	Mousam Guha Roy	Very niceREAD MORE
102	Ajin V	High quality camera🥰READ MORE
103	Prithivi Boruah	Camera Quality Is Improved Loving ItREAD MORE
104	Ankit Verma	Awesome camra quality superb camra 😊READ MORE
105	Suman Jain	Simply Awesome 🥰READ MORE
106	Rahul Shedge	Totally happy!Camera 5Battery 5 Display 5Desig...
107	Jayanta Roy	Nice 😊😊😊😊😊READ MORE
108	Sheetla Prasad Maurya	Best mobile phoneCamera quality is very nice B...
109	Saurabh Gothwad	Simply premium.READ MORE
110	Praveen Kumar	This is my first iphone ever used loving it al...

	star	comment
101	4	Very niceREAD MORE
102	5	High quality camera🥰READ MORE
103	5	Camera Quality Is Improved Loving ItREAD MORE
104	4	Awesome camra quality superb camra 😊READ MORE
105	5	Simply Awesome 🥰READ MORE
106	5	Totally happy!Camera 5Battery 5 Display 5Desig...
107	5	Nice 😊😊😊😊😊READ MORE
108	4	Best mobile phoneCamera quality is very nice B...
109	5	Simply premium.READ MORE
110	5	This is my first iphone ever used loving it al...

(7)To see quick overview of the central tendency, spread, and distribution of the data in the DataFrame.

```
df.describe()
```

Output:

	Unnamed: 0	star
count	10.00000	10.000000
mean	105.50000	4.700000
std	3.02765	0.483046
min	101.00000	4.000000
25%	103.25000	4.250000
50%	105.50000	5.000000
75%	107.75000	5.000000
max	110.00000	5.000000

This summary can help in identifying outliers, understanding the range of values, and getting a general sense of the data distribution for further analysis.

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

```
df.to_csv("flipkart_iphone15_review.csv")
```

Output:

	CName	Review	star	comment
101	Mousam Guha Roy	Very nice	4	Very nice
102	Ajin V	High quality camera	5	High quality camera
103	Prithivi Boruah	Camera Quality Is Improved Loving It	5	Camera Quality Is Improved Loving It
104	Ankit Verma	Awesome camra quality superb camra	4	Awesome camra quality superb camra
105	Suman Jain	Simply Awesome	5	Simply Awesome
106	Rahul Shedge	Totally happy!	5	Totally happy!
107	Jayanta Roy	Nice	5	Nice
108	Sheetla Prasad Maurya	Best mobile phone	4	Best mobile phone
109	Saurabh Gothwad	Simply premium.	5	Simply premium.
110	Praveen Kumar	This is my first iphone ever used loving it already this is worth it	5	This is my first iphone ever used loving it already this is worth it

CONCLUSION

- **Successful Data Extraction:** The project effectively gathered iPhone 15 details and customer feedback from Flipkart using web scraping techniques.
- **Insights Unveiled:** Analysis revealed pricing trends, customer sentiments, and product popularity metrics.
- **Challenges Acknowledged:** Adapting to website changes and maintaining data accuracy were key hurdles.
- **Recommendations for the Future:** Emphasize regular script updates, enhanced analysis techniques, and data enrichment for comprehensive insights.

