

Project for Web Scrapping

Web Scrapping of OnePlus-nord From Flipkart

Institute Name:- ITvedant Education Pvt.Ltd

Name:- Kapil Vijay Ghag

Email Address:- kapilghag439@gmail.com

Date Of Submission:-06th Dec 2023

AIMOFTHEPROJECT

The aim of this project is to create a Python Web scraping to extract specific details of the oneplus-nord from flipkart . 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 oneplus-nord listings



Description:

The primary goal of this web scraping project is to extract relevant information about the oneplus-nord from the specified Flipkart product page for analysis purpose. The key data points include product title, price, comment, and customer reviews.

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

Project Steps :-

- **URL Retrieval**: You've defined the URL for the oneplus-nord 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, filtering.
- **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 oneplus-nord 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.

Challenges Faced:

- **Dynamic Content:** Some parts of the webpage, such as customer reviews, are loaded dynamically using JavaScript. This required additional handling to ensure complete data retrieval.
- **Anti-Scraping Measures:** Flipkart employs anti-scraping measures like rate limiting and IP blocking. The scraping process incorporated delays to mitigate these issues.

Output Format:

The scraped data is stored in a structured format, such as CSV or JSON, for easy integration with downstream processes.

Analysis of project:

Based on our analysis we conclude that this project has most sales in the smartwatch in

Flipkart as this product has no negative review and has average of 4.8

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

```
<class 'bs4.BeautifulSoup'>
<class 'bs4.element.Tag'>
<class 'bs4.element.NavigableString'>
```

(2) To access oneplus-nord price by using this class.

```
title = soup.title
title
```

```
<title>Oneplus Nord Ce3 5g Aqua Surge 256 Gb Reviews: Latest Review of Oneplus Nord Ce3 5g Aqua Surge 256 Gb | Price in India | Flipkart.com</title>
```

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

```
[<div class="_30jeq3">₹28,968</div>]
```

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

```
'₹28,968'
```

(3) To access oneplus-nord Customer Name by using the respective class from its HTML content.

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

```
<class 'list'>
```

```
[<p class="_2sc7ZR _2V5EHH">Prem Ashok</p>, <p class="_2sc7ZR _2V5EHH">T.keshava Keshava</p>, <p class="_2sc7ZR _2V5EHH">Kanhaiya Dixit</p>, <p class="_2sc7ZR _2V5EHH">Lalit Ray</p>, <p class="_2sc7ZR _2V5EHH">Indu Kumari </p>, <p class="_2sc7ZR _2V5EHH">saddam hussain shah</p>, <p class="_2sc7ZR _2V5EHH">VIPIN VISHWAKARMA</p>, <p class="_2sc7ZR _2V5EHH">ATIF KHAN</p>, <p class="_2sc7ZR _2V5EHH">Flipkart Customer</p>, <p class="_2sc7ZR _2V5EHH">Amit Kumar</p>]
```

```
10
```

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

```
'Prem Ashok'
```

```
c_names=[]
for i in names:
    c_names.append(i.get_text())
c_names
```

```
['Prem Ashok',
 'T.keshava Keshava',
 'Kanhaiya Dixit',
 'Lalit Ray',
 'Indu Kumari ',
 'saddam hussain shah',
 'VIPIN VISHWAKARMA',
 'ATIF KHAN',
 'Flipkart Customer',
```

(4) To access oneplus-nord Customer Comment by using the respective class.

```
comment = soup.find_all(class_ = "_2-N8zT")
print(len(comment))
```

10

```
review_comment = []
for i in comment:
    review_comment.append(i.get_text())
review comment
```

```
['Fabulous!',  
'Wonderful',  
'Good choice',  
'Good choice',  
'Worth the money',  
'Best in the market!',  
'Wonderful',  
'Fabulous!',  
'Mind-blowing purchase',  
'Worth every penny']
```

(5) To access oneplus-nord Customer Review star by using the respective class.

```
stars=soup.find_all('div', class_="_3LWZlK _1BLPMq")
print(stars)
```

[illegible]

```
: s_stars = []
  for i in stars:
      s_stars.append(i.get_text())
  s_stars

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


(6) To access oneplus-nord Customer Comment by using the respective class.

```
review = soup.find_all('div', class_ = "t-ZTKy")
print(review)
```

```
[<div class="t-ZTKy"><div><div class="">When I checked review there was full negative feedback, so I was worried about that,, a
fter mobile delivery, after using 2 day's I am leaving my review, mobile is super, and overall performance is good, mobile disp
lay charging, and camera is too good, when u use the mobile you feel like it's 1 1lakh rupees mobile, mobile display is very ve
ry smooth, it has a premium look, finger print is very fast, charging like within 25mins full of charge 0-100%, you buy this mo
bile, in thiss price it's ve...</div><span class="_1BWGvX"><span>READ MORE</span></span></div></div>, <div class="t-ZTKy"><div>
<div class="">Product Good</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <div class="t-ZTKy"><div><div
class="">Display is satisfied and other everything good</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <
div class="t-ZTKy"><div><div class="">Allrounder phone</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <d
iv class="t-ZTKy"><div><div class="">Good</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <div class="t-Z
TKy"><div><div class="">I like it</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <div class="t-ZTKy"><di
v><div class="">Best product 🥰🥰🥰</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <div class="t-ZTKy">
<div><div class="">Good camera quality 🙌</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>, <div class="t-Z
TKy"><div><div class="">Very nice and awesome alrounder phone..loved it</div><span class="_1H-bmy"><span>READ MORE</span></span>
</div></div>, <div class="t-ZTKy"><div><div class="">Thanks for the fastest delivery flipkart team awesome looking this phone
I love one+ mobile</div><span class="_1H-bmy"><span>READ MORE</span></span></div></div>]
```

```
cust_review =[]
for i in review:
    cust_review.append(i.get_text())
cust_review
```

```
["When I checked review there was full negative feedback, so I was worried about that,, after mobile delivery, after using 2 da
y's I am leaving my review, mobile is super, and overall performance is good, mobile display charging, and camera is too good,
when u use the mobile you feel like it's 1 1lakh rupees mobile, mobile display is very very smooth, it has a premium look, fing
er print is very fast, charging like within 25mins full of charge 0-100%, you buy this mobile, in thiss price it's ve...READ M
ORE",
'Product GoodREAD MORE',
'Display is satisfied and other everything goodREAD MORE',
'Allrounder phoneREAD MORE',
'GoodREAD MORE',
'I like itREAD MORE',
'Best product 🥰🥰🥰 READ MORE',
'Good camera quality 🙌 READ MORE',
'Very nice and awesome alrounder phone..loved itREAD MORE',
'Thanks for the fastest delivery flipkart team awesome looking this phone I love one+ mobilereAD MORE']
```

(7) To arrange properly in column all above data by using -Pandas library.

```
import pandas as pd
```

```
df= pd.DataFrame({"customers_Name":c_names,"Review_Comment":review_comment,"review_stars":s_stars,"review":cust_review})  
df
```

	customers_Name	Review_Comment	review_stars	review
0	Prem Ashok	Fabulous!	5	When I checked review there was full negative ...
1	T.keshava Keshava	Wonderful	4	Product GoodREAD MORE
2	Kanhaiya Dixit	Good choice	4	Display is satisfied and other everything good...
3	Lalit Ray	Good choice	4	Allrounder phoneREAD MORE
4	Indu Kumari	Worth the money	4	GoodREAD MORE
5	saddam hussain shah	Best in the market!	5	I like itREAD MORE
6	VIPIN VISHWAKARMA	Wonderful	4	Best product 🥰🥰🥰 READ MORE
7	ATIF KHAN	Fabulous!	5	Good camera quality 🙌 READ MORE
8	Flipkart Customer	Mind-blowing purchase	5	Very nice and awesome alrounder phone..loved i...
9	Amit Kumar	Worth every penny	5	Thanks for the fastest delivery flipkart team ...

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

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

```
import os  
os.getcwd
```

```
<function nt.getcwd()>
```

```
df1.describe()
```

review_stars	
count	10.000000
mean	4.500000
std	0.527046
min	4.000000
25%	4.000000
50%	4.500000
75%	5.000000
max	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

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

```
import os
os.getcwd()
```

```
'C:\\Users\\HP'
```

```
df.to_csv("oneplus_mobile.csv",index= False )
df1= pd.read_csv("oneplus_mobile.csv")
df1
```

	customers_Name	Review_Comment	review_stars	review
0	Prem Ashok	Fabulous!	5	When I checked review there was full negative ...
1	T.keshava Keshava	Wonderful	4	Product GoodREAD MORE
2	Kanhaiya Dixit	Good choice	4	Display is satisfied and other everything good...
3	Lalit Ray	Good choice	4	Allrounder phoneREAD MORE
4	Indu Kumari	Worth the money	4	GoodREAD MORE
5	saddam hussain shah	Best in the market!	5	I like itREAD MORE
6	VIPIN VISHWAKARMA	Wonderful	4	Best product 🥰🥰🥰 READ MORE
7	ATIF KHAN	Fabulous!	5	Good camera quality 🙌 READ MORE
8	Flipkart Customer	Mind-blowing purchase	5	Very nice and awesome alrounder phone..loved i...
9	Amit Kumar	Worth every penny	5	Thanks for the fastest delivery flipkart team ...

(10)Find max,mean,min etc

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

```
import os
os.getcwd
```

```
<function nt.getcwd()>
```

```
df1.describe()
```

	review_stars
count	10.000000
mean	4.500000
std	0.527046
min	4.000000
25%	4.000000
50%	4.500000
75%	5.000000
max	5.000000

(11) Find INFO of data

```
: df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype
---  -
0   customers_Name  10 non-null    object
1   Review_Comment  10 non-null    object
2   review_stars    10 non-null    int64
3   review          10 non-null    object
dtypes: int64(1), object(3)
memory usage: 452.0+ bytes
```

(12) sort the name of Customer in Ascending order

```
df1.customers_Name.sort_values()
```

```
7      ATIF KHAN
9      Amit Kumar
8  Flipkart Customer
4      Indu Kumari
2  Kanhaiya Dixit
3      Lalit Ray
0      Prem Ashok
1  T.keshava Keshava
6  VIPIN VISHWAKARMA
5  saddam hussain shah
Name: customers_Name, dtype: object
```

(13) Sort the Names of Customer in Descending order

```
df1.customers_Name.sort_values(ascending=False)
```

```
5  saddam hussain shah
6  VIPIN VISHWAKARMA
1  T.keshava Keshava
0  Prem Ashok
3  Lalit Ray
2  Kanhaiya Dixit
4  Indu Kumari
8  Flipkart Customer
9  Amit Kumar
7  ATIF KHAN
Name: customers_Name, dtype: object
```

(14) Find who had given full rating

```
df1[df1.review_stars==5]
```

	customers_Name	Review_Comment	review_stars	review
0	Prem Ashok	Fabulous!	5	When I checked review there was full negative ...
5	saddam hussain shah	Best in the market!	5	I like itREAD MORE
7	ATIF KHAN	Fabulous!	5	Good camera quality 🙌 READ MORE
8	Flipkart Customer	Mind-blowing purchase	5	Very nice and awesome alrounder phone..loved i...
9	Amit Kumar	Worth every penny	5	Thanks for the fastest delivery flipkart team ...

(15) Find who gave less than 5 rating

```
df1[df1.review_stars<5]
```

	customers_Name	Review_Comment	review_stars	review
1	T.keshava Keshava	Wonderful	4	Product GoodREAD MORE
2	Kanhaiya Dixit	Good choice	4	Display is satisfied and other everything good...
3	Lalit Ray	Good choice	4	Allrounder phoneREAD MORE
4	Indu Kumari	Worth the money	4	GoodREAD MORE
6	VIPIN VISHWAKARMA	Wonderful	4	Best product 🥰🥰🥰 READ MORE

CONCLUSION

The web scraping process for the **oneplus-nord** Flipkart product page was successful, The extracted data can be used for further analysis and revealed customer sentiments, and product popularity metrics. Regular monitoring and maintenance are recommended to adapt to any changes in the website structure.

