

PROJECT FOR WEB SCRAPING

Web Scrapping of From Wings earbuds flipkart

**Institute Name:- ITvedant Education
Pvt.Ltd**

Name:- Sejal Santosh Bhekare

Email Address:-

bhekaresejal2002@gmail.com

Date Of Submission:-03 Dec 2023

AIM OF THE PROJECT

The aim of this project is to create a Python Web scraping to extract specific details of the wings earbuds 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 wings earbuds listings



DESCRIPTION:-

Web scraping of Flipkart products involves programmatically extracting information about various products listed on the Flipkart e-commerce platform. This process allows users to gather data such as product names, prices, ratings and reviews available on Flipkart's product pages.

To perform web scraping on Flipkart, you'd typically utilize tools, programming languages, or libraries like Python's BeautifulSoup. These tools can simulate human browsing behavior, sending requests to Flipkart's servers, retrieving the HTML content of product pages, and then parsing the HTML to extract the desired product information.

PROJECT STEPS :-

- **URL Retrieval:** You've defined the URL for the Apple Macbook 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:-

Market Analysis and Pricing Strategy: Extracting product details, pricing information, discounts, and stock availability helps in analyzing market trends, understanding competitor pricing strategies, and adjusting your own pricing strategy accordingly.

Competitor Monitoring: Regularly scraping Flipkart's product listings allows for tracking competitors' product ranges, availability, and pricing changes. This information helps in identifying gaps in the market or potential niches to explore.

Product Research and Development: Scraping Flipkart can provide insights into customer reviews, ratings, and feedback on products. This data aids in understanding customer preferences, identifying popular features, and guiding product development decisions.

Sales and Marketing Insights: Extracting data on product descriptions, specifications, and customer reviews can guide marketing campaigns, content creation, and SEO strategies by understanding what features or aspects customers value the most.

Price Comparison and Adjustment: Scraping Flipkart allows for comparing prices of similar products across different vendors. This enables businesses to adjust their pricing strategies to stay competitive in the market.

Consumer Behavior Analysis: Analyzing data from Flipkart's product reviews and ratings can provide valuable insights into consumer behavior, preferences, and sentiments toward certain products or brands.

Personalized Customer Experience: Insights gained from scraping Flipkart can be used to tailor product offerings, customer service, and shopping experiences to meet specific customer needs and preferences.

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

Wings earbuds Web Scrapping from Flipkart Code command:

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

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

```
In [4]: url = "https://www.flipkart.com/wings-phantom-pro-earbuds-low-latency-game-mode-40h-playtime-50ms-bluetooth-headset/p/itm9ceff01a"
<img alt="Progress bar" data-bbox="228 271 398 278"/>
```

```
In [5]: r = requests.get(url)
data = r.content
print(data)
```

```
b'<!doctype html><html lang="en"><head><link href="https://rukminim2.flixcart.com" rel="preconnect"><link rel="stylesheet" href="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app_modules.chunk.905c37.css"/><link rel="stylesheet" href="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/css/app.chunk.615ed9.css"/><meta http-equiv="Content-type" content="text/html; charset=utf-8"/><meta http-equiv="X-UA-Compatible" content="IE=Edge"/><meta property="fb:page_id" content="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/opensearchdescription+xml" rel="search" href="/osdd.xml?v=2"/><meta property="og:type" content="website"/><meta name="og:site_name" property="og:site_name" content="Flipkart.com"/><link rel="apple-touch-icon" sizes="57x57" href="/apple-touch-icon-57x57.png"/><link rel="apple-touch-icon" sizes="72x72" href="/apple-touch-icon-72x72.png"/><link rel="apple-touch-icon" sizes="114x114" href="/apple-touch-icon-114x114.png"/><link rel="apple-touch-icon" sizes="144x144" href="/apple-touch-icon-144x144.png"/><link rel="apple-touch-icon" href="/apple-touch-icon-57x57.png"/><meta name="twitter:card" content="app"><meta name="twitter:site" content="@flipkart"><meta name="twitter:creator" content="@flipkart"><meta name="twitter:title" content="Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Price in India - Buy Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Online - Wings : Flipkart.com"><meta name="twitter:description" content="Shop for electronics, apparels & more using our Flipkart app Free shipping & COD."><meta property="twitter:image" content="http://rukmini1.flixcart.com/image/300/300/xif0q/headphone/d/w/k/-original-imageznw4hxna7tjg.jpeg"/><meta name="twitter:app:country" content="in"><meta name="al:ios:app_name" content="Flipkart"><meta name="al:ios:app_store_id" content="742044692"><meta name="twitter:app:name:iphone" content="Flipkart"><meta name="twitter:app:id:iphone" content="742044692"><meta name="twitter:app:url:iphone" content="http://dl.flipkart.com/dl/home?><meta name="twitter:app:nam
```

```
In [6]: #prettyfy the codes of html page
soup = BeautifulSoup(data, "html.parser")
soup.prettify
```

```
Out[6]: <bound method Tag.prettify of <!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.615ed9.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="no
odp" name="robots"/><link href="https://www.promos/new/20150528-140547-favicon-retina.ico" rel="shortcut icon"/><link href="/osdd.xml?v=2" rel="search" type="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="Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Price in India - Buy Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Online - Wings : Flipkart.com" 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/headphone/d/w/k/-original-imageznw4hxna7tjg.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?><meta name="twitter:app:nam
```

```
In [10]: title = soup.title
print(type(soup))           #beautifulsoup
print(type(title))          #tag
print(type(title.string))    #navigablestring
```

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

(2) To access wings earbuds price by using this class.

```
In [11]: title = soup.title
         title

Out[11]: <title>Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Price in India - Buy Wings Phantom Pro earbuds with Low latency Game mode, 40H Playtime, 50ms low latency Bluetooth Headset Online - Wings : Flipkart.com</title>

In [13]: price = soup.find_all(class_="_30jeq3 _16jk6d")
         price

Out[13]: [<div class="_30jeq3 _16jk6d">₹1,099</div>]

In [14]: price[0].get_text()

Out[14]: '₹1,099'
```

(3) To access wings earbuds Customer Name by using the respective class from its HTML content.

```
In [17]: names = soup.find_all(class_="_2sc7ZR _2V5EH#")
         names = list(names)
         print(type(names))
         print(names)
         len(names)

<class 'list'>
[<p class="_2sc7ZR _2V5EH#">saif ali khan</p>, <p class="_2sc7ZR _2V5EH#">Ankit Sahoo</p>, <p class="_2sc7ZR _2V5EH#">Sohail Khan</p>, <p class="_2sc7ZR _2V5EH#">Flipkart Customer</p>, <p class="_2sc7ZR _2V5EH#">Julen Nath</p>, <p class="_2sc7ZR _2V5EH#">Ashwani Kumar</p>, <p class="_2sc7ZR _2V5EH#">Sahil</p>, <p class="_2sc7ZR _2V5EH#">Flipkart Customer</p>, <p class="_2sc7ZR _2V5EH#">Akash Patke </p>, <p class="_2sc7ZR _2V5EH#">Suvam Paul</p>]

Out[17]: 10

In [18]: names[0].get_text()

Out[18]: 'saif ali khan'

In [19]: c_names = []
         for i in names:
             c_names.append(i.get_text())
         c_names

Out[19]: ['saif ali khan',
          'Ankit Sahoo',
          'Sohail Khan',
          'Flipkart Customer',
          'Julen Nath',
          'Ashwani Kumar',
          'Sahil',
          'Flipkart Customer',
          'Akash Patke ',
          'Suvam Paul']
```

(4) To access wings earbuds Customer Comment by using the respective class.

```
In [21]: comment = soup.find_all(class_="_2-N8zT")
         print(len(comment))

10

In [22]: review_comment = []
         for i in comment:
             review_comment.append(i.get_text())
         review_comment

Out[22]: ['Nice product',
          'Wonderful',
          'Awesome',
          'Highly recommended',
          'Value-for-money',
          'Good choice',
          'Highly recommended',
          'Highly recommended',
          'Highly recommended',
          'Classy product',
          'Worth every penny']
```


(5) To access wings earbuds Customer Review star by using the respective class.

[illegible]

```
In [25]: s_starts = []
         for i in stars:
             s_starts.append(i.get_text())
         s_starts
```

```
Out[25]: ['4', '5', '5', '5', '4', '4', '5', '5', '5', '5']
```

(6)To access wings earbuds Customer Comment by using the respective class.

```
In [17]: review = soup.find_all('div', class_ = "t-ZTKy")
print(review)
```

[<div class="t-ZTKy"><div><div>ss="">Pros -
1. Highly recommended Budget Gaming TWS.
2. Sound quality is good in th
e. price range.
3. Gaming lights look awesome
4. Latency in gaming mode is the lowest.
5. Battery backup is awesom
e.

Cons -
1. Difficulty in opening the case can be seen.
2. While playing PUBG difficulty in understanding
the footsteps direction.
3. Starts paining on the left ear, don't know if it's only for me.

Overall a good budge
t gaming TWS in this range.</div>READ MORE</div><div><div class="t-ZTKy"><div>
class="">Good design good product</div>READ MORE</div><div><div class="t-ZTKy"><di
v><div class="">Good for gaming 🎮</div>READ MORE</div><div><div class="t-ZTKy"><d
iv><div class="">It's all awesome 😍
More than i expected ❤️</div>READ MORE</div>
</div><div><div class="t-ZTKy"><div><div class="">My personal opinion sound quality is very good.
Bass is awesome 🙌
Cal
ling is excellent work but gaming mode is not good. Not working in the same firing mode.
Battery backup very good 🙌
M
usic lovers by this product.
Gaming lover don't by this product.</div>READ MORE
</div><div><div class="t-ZTKy"><div><div class="">Value for money</div>READ MORE</div>
</div><div><div class="t-ZTKy"><div><div class="">Unique design solid build quality</div>READ MORE
</div><div><div class="t-ZTKy"><div><div class="">Good product light is good but case opening is not that good
overall good at this price</div>READ MORE</div></div><div><div class="t-ZTKy"><div><div
class="">Honestly, one of the best product available in the market for gaming with 60 mili seconds of lag which is not availab
le in the market right now and the sound quality is superb.
Best buy for gaming 🎮</div>READ M
ORE</div></div></div>


```
In [29]: cust_review =[]
for i in review:
    cust_review.append(i.get_text())
cust_review

Out[29]: ["Pros - 1. Highly recommended Budget Gaming TWS.2. Sound quality is good in this price range.3. Gaming lights look awesome 4. Latency in gaming mode is the lowest.5. Battery backup is awesome. Cons - 1. Difficulty in opening the case can be seen. 2. While playing PUBG difficulty in understanding the footsteps direction. 3. Starts paining on the left ear, don't know if it's only for me. Overall a good budget gaming TWS in this range.READ MORE",
'Good design good productREAD MORE',
'Good for gaming 🎮 READ MORE',
'It's all awesome 🥰 More than i expected ❤️ READ MORE',
'My personal opinion sound quality is very good.Bass is awesome 🎵.Calling is excellent work but gaming mode is not good .Not working in the same firing mode.Battery backup very good 🎧.Music lovers by this product. Gaming lover don't by this product.READ MORE",
'Value for moneyREAD MORE',
'Unique design solid build qualityREAD MORE',
'Good product light is good but case opening is not that good overall good at this priceREAD MORE',
'Honestly, one of the best product available in the market for gaming with 60 mili seconds of lag which is not available in the market right now and the sound quality is superb. Best buy for gaming 🎮 READ MORE',
'Osm look 🥰🥰 READ MORE']
```

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

```
In [30]: import pandas as pd
```

```
In [31]: df= pd.DataFrame({"customers_Name":c_names,"Review_Comment":review_comment,"review_stars":s_stars,"review":cust_review})
df
```

```
Out[31]:
```

	customers_Name	Review_Comment	review_stars	review
0	saif ali khan	Nice product	4	Pros - 1. Highly recommended Budget Gaming TWS...
1	Ankit Sahoo	Wonderful	5	Good design good productREAD MORE
2	Sohail Khan	Awesome	5	Good for gaming 🎮 READ MORE
3	Flipkart Customer	Highly recommended	5	It's all awesome 🥰 More than i expected ❤️ READ MORE
4	Julen Nath	Value-for-money	4	My personal opinion sound quality is very good...
5	Ashwani Kumar	Good choice	4	Value for moneyREAD MORE
6	Sahil	Highly recommended	5	Unique design solid build qualityREAD MORE
7	Flipkart Customer	Highly recommended	5	Good product light is good but case opening is...
8	Akash Patke	Classy product	5	Honestly, one of the best product available in...
9	Suvam Paul	Worth every penny	5	Osm look 🥰🥰 READ MORE

```
In [32]: import os
os.getcwd()
```

```
Out[32]: 'C:\\Users\\HP'
```

```
In [33]: df.to_csv("wings_earbuds.csv",index= False )
df1= pd.read_csv("wings_earbuds.csv")
df1
```

```
Out[33]:
```

	customers_Name	Review_Comment	review_stars	review
0	saif ali khan	Nice product	4	Pros - 1. Highly recommended Budget Gaming TWS...
1	Ankit Sahoo	Wonderful	5	Good design good productREAD MORE
2	Sohail Khan	Awesome	5	Good for gaming 🎮 READ MORE
3	Flipkart Customer	Highly recommended	5	It's all awesome 🥰 More than i expected ❤️ READ MORE
4	Julen Nath	Value-for-money	4	My personal opinion sound quality is very good...
5	Ashwani Kumar	Good choice	4	Value for moneyREAD MORE
6	Sahil	Highly recommended	5	Unique design solid build qualityREAD MORE
7	Flipkart Customer	Highly recommended	5	Good product light is good but case opening is...
8	Akash Patke	Classy product	5	Honestly, one of the best product available in...
9	Suvam Paul	Worth every penny	5	Osm look 🥰🥰 READ MORE

```
In [34]: df.to_csv("wings_earbuds.csv")
```

```
In [36]: import os
os.getcwd()
```

```
Out[36]: <function nt.getcwd()>
```

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

```
In [37]: df1.describe()
```

```
Out[37]:
```

	review_stars
count	10.000000
mean	4.700000
std	0.483046
min	4.000000
25%	4.250000
50%	5.000000
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.

```
In [34]: df.to_csv("wings_earbuds.csv")
```

```
In [2]: import os  
os.getcwd()
```

```
Out[2]: 'C:\\Users\\HP'
```

(10) Find max, mean, min etc

```
In [37]: df1.describe()
```

```
Out[37]:
```

	review_stars
count	10.000000
mean	4.700000
std	0.483046
min	4.000000
25%	4.250000
50%	5.000000
75%	5.000000
max	5.000000

(11) Find INFO of data

```
In [44]: 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

```
In [39]: df1.customers_Name.sort_values()

Out[39]: 8      Akash Patke
1      Ankit Sahoo
5      Ashwani Kumar
3      Flipkart Customer
7      Flipkart Customer
4      Julen Nath
6      Sahil
2      Sohail Khan
9      Suvam Paul
0      saif ali khan
Name: customers_Name, dtype: object
```

(13) Sort the Names of Customer in Descending order

```
In [40]: df1.customers_Name.sort_values(ascending=False)

Out[40]: 0      saif ali khan
9      Suvam Paul
2      Sohail Khan
6      Sahil
4      Julen Nath
3      Flipkart Customer
7      Flipkart Customer
5      Ashwani Kumar
1      Ankit Sahoo
8      Akash Patke
Name: customers_Name, dtype: object
```

15) Find who had given full rating

```
In [41]: df1[df1.review_stars==5]

Out[41]:
```

	customers_Name	Review_Comment	review_stars	review
1	Ankit Sahoo	Wonderful	5	Good design good productREAD MORE
2	Sohail Khan	Awesome	5	Good for gaming 🍷 READ MORE
3	Flipkart Customer	Highly recommended	5	It's all awesome 🍷 More than i expected 🍷 READ MORE
6	Sahil	Highly recommended	5	Unique design solid build qualityREAD MORE
7	Flipkart Customer	Highly recommended	5	Good product light is good but case opening is...
8	Akash Patke	Classy product	5	Honestly, one of the best product available in...
9	Suvam Paul	Worth every penny	5	Osm look 🍷 🍷 READ MORE

16) Find who gave less than 5 rating

```
In [43]: df1[df1.review_stars<5]
```

```
Out[43]:
```

	customers_Name	Review_Comment	review_stars	review
0	saif ali khan	Nice product	4	Pros - 1. Highly recommended Budget Gaming TWS...
4	Julen Nath	Value-for-money	4	My personal opinion sound quality is very good...
5	Ashwani Kumar	Good choice	4	Value for moneyREAD MORE

CONCLUSION

The web scraping process for the wings earbuds 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.

**THANK
YOU**