

# **“RYNOXGEAR WEBSITE WEB-SCRAPING”**

**A project Submitted to the**  
**I.T. VEDANT Institute, Thane.**



**Python-Web-Scraping Project**

**BY**

**KARAN SHRIRAM GUPTA**

**Under the Guidance Of**

**Mr. Sameer Warsolkar**

## DESCRIPTIONS :

Riding Jacket web-scraping on Rynoxgear project involves extracting data from websites. This data can be used for various purposes, such as data analysis, research, reporting, or populating databases. We are going to use Riding Jacket to scrape site. To do so, we are going to use tools like Python, Requests, BeautifulSoup.



## STEALTH EVO 3 JACKET

## **Important Libraries for Project :-**

Python has several web scrapping libraries. We will use the following libraries:

1. Requests :- for making HTTP requests to website.
2. BeautifulSoup :- for parsing the HTML code .
3. Pandas :- for storing the scraped data in data frame.

## **Project Outline :-**

From this site, we are going to grab the following information:-

- Name of the product.
- Price of the product.
- Feature of the product.
- Reviews.

## **BENEFITS:-**

- Access to valuable data for analysis or research.
- Automation of data collection, saving time and effort.
- Stay up to date with change changes on the target websites.

## **Risk:-**

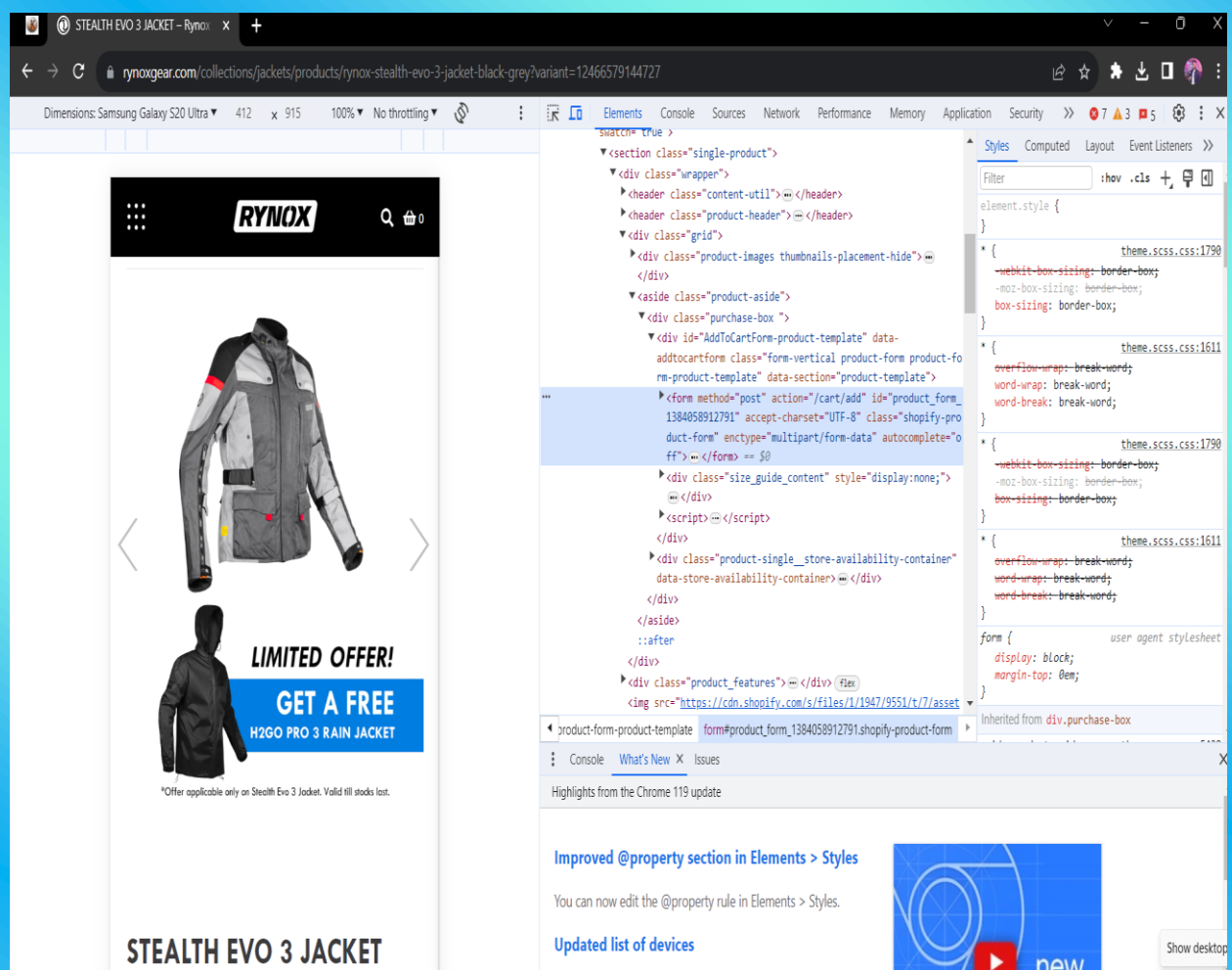
- Legal issues related to web scraping.
- Technical challenges due to website changes.

## Exporting the extracted data:-

We'll export the data as a CSV file. We will use the pandas library. We'll use the pandas library.

## Inspect the website:-

Now the next step is to understand the website structure. Understand what the attributes of the elements that are of your interest are. Right click on the website to select "Inspect". This will open HTML code. Use the inspector tool to see the name of all the elements to use in the code.





## RYNOXGEAR Web-Scraping Code :-

### 1. Access Rynoxgear website by using import requests URL and BeautifulSoup library.

```
In [1]: 1 import requests
        2 from bs4 import BeautifulSoup
        3 import pandas as pd
```

```
In [2]: 1 url = "https://rynoxgear.com/collections/jackets/products/rynox-stealth-evo-3-jacket-black-grey?variant=12466579144727"
```

```
In [3]: 1 r = requests.get(url)
        2 htmlcontent = r.content
        3 htmlcontent
```

s of the site are most popular.", "targetingCookiesDescriptionText": {"en": "These cookies may be set through our site by our advertising partners. They may be used by those companies to build a profile of your interests and show you relevant adverts on other sites. They do not store directly personal information, but are based on uniquely identifying your browser and internet device. If you do not allow these cookies, you will experience less targeted advertising."}, "unclassifiedCookiesDescriptionText": {"en": "Unclassified cookies are cookies that we are in the process of classifying, together with the providers of individual cookies."}, "auto": {"declName": {"en": "Name"}, "declPath": {"en": "Path"}, "declType": {"en": "Type"}, "declDomain": {"en": "Domain"}, "declPurpose": {"en": "Purpose"}, "declProvider": {"en": "Provider"}, "declRetention": {"en": "Retention"}, "declFirstParty": {"en": "First-party"}, "declThirdParty": {"en": "Third-party"}, "cookiesDetailsText": {"en": "Cookies details"}, "preferencesPopupAlwaysAllowedText": {"en": "Always allowed"}, "submitButton": {"en": "Submit"}, "submittingButton": {"en": "Submitting..."}, "cancelButton": {"en": "Cancel"}, "guestsSupportInfoText": {"en": "Please login with your customer account to further proceed."}, "guestsSupportEmailPlaceholder": {"en": "E-mail address"}, "guestsSupportEmailValidationError": {"en": "Email is not valid"}, "guestsSupportEmailSuccessTitle": {"en": "Thank you for your request"}, "guestsSupportEmailFailureTitle": {"en": "A problem occurred"}, "guestsSupportEmailSuccessMessage": {"en": "If you are registered as a customer of this store, you will soon receive an email with instructions on how to proceed."}, "guestsSupportEmailFailureMessage": {"en": "Your request was not submitted. Please try again and if problem persists, contact store owner for assistance."}, "confirmationSuccessTitle": {"en": "Your request is verified"}, "confirmationFailureTitle": {"en": "A problem occurred"}, "confirmationSuccessMessage": {"en": "We will soon get back to you as to your request."}, "confirmationFailureMessage": {"en": "Your request was not verified. Please try again and if problem persists, contact store owner for assistance."}, "consentSectionTitleText": {"en": "Your cookie consent"}, "consentSectionNoConsentText": {"en": "You have not consented to the cookies policy of this website."}, "consentSectionConsentedText": {"en": "You consented to the cookies policy of this website on"}, "consentStatus": {"en": "Consent preference"}, "consentDate": {"en": "Consent date"}, "consentId":

### 2. Parse the HTML Link.

```
In [4]: 1 soup = BeautifulSoup(htmlcontent, "html.parser")
        2 #print(soup)
        3 soup.prettify #Prettify the codes of the HTML page
```

```
Out[4]: <bound method Tag.prettify of <!DOCTYPE html>
```

```
<!--[if IE 8]><html class="no-js lt-ie9" lang="en"> <![endif]-->
<!--[if IE 9]><html class="ie9 no-js"> <![endif]-->
<!--[if (gt IE 9)]!(IE)><!--> <html class="no-js" lang="en"> <!--<![endif]-->
<head>
<!-- PANDECTES-GDPR: DO NOT MODIFY AUTO GENERATED CODE BELOW --> <script id="pandectes-rules">
```

```
    window.PandectesSettings = {"store":{"plan":"basic","theme":"Copy of District (Simpl Checkout)","primaryLocale":"en","adminMode":false,"tsPublished":1682405406,"declaration":{"showPurpose":false,"showProvider":false,"showDateGenerated":false},"language":{"languageMode":"Single","fallbackLanguage":"en","languageDetection":"browser","languagesSupported":[]},"texts":{"managed":{"headerText":{"en":"We respect your privacy"},"consentText":{"en":"This website uses cookies to ensure you get the best experience. By continuing to browse this site you agree to the use of cookies"},"dismissButtonText":{"en":"Ok"},"linkText":{"en":"Learn more"},"imprintText":{"en":"Imprint"},"preferencesButtonText":{"en":"Preferences"},"allowButtonText":{"en":"Accept"},"denyButtonText":{"en":"Decline"},"leaveSiteButtonText":{"en":"Leave this site"},"cookiePolicyText":{"en":"Cookie policy"},"preferencesPopupTitleText":{"en":"Manage consent preferences"},"preferencesPopupIntroText":{"en":"We use cookies to optimize website functionality, analyze the performance, and provide personalized experience to you. Some cookies are essential to make the website operate and function correctly. Those cookies cannot be disabled. In this window you can manage your preference of cookies."},"preferencesPopupCloseButtonText":{"en":"Close"},"preferencesPopupAcceptAllButtonText":{"en":"Accept all"},
```

### 3. Access Rynox Jacket Name using the Source Code.

#### ACCESS NAME OF PRODUCT

```
In [5]: 1 title = soup.title
2 print(type(soup)) # BeautifulSoup
3 print(type(title)) # Tag
4 print(type(title.string)) # NavigableString

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

```
In [6]: 1 title=soup.title
2 title

Out[6]: <title>
STEALTH EVO 3 JACKET – Rynox Gear
</title>
```

### 4. Access Rynox Jacket Price using this source code.

#### ACCESS PRICE

```
In [7]: 1 Price = soup.find_all("div",class_="selection-wrapper price product-single__price-product-template")
2 Price

Out[7]: [<div class="selection-wrapper price product-single__price-product-template" data-price="">
<span class="money" data-productprice="" id="ProductPrice-product-template"><span class="money">₹ 12,750</span></span>
<span style="font-size:11px; font-weight:normal; color:#999; ">MRP <i style="font-size:11px; font-weight:normal; font-style:
normal; text-transform:lowercase !important;">inclusive of all taxes</i></span>
<p data-compareprice="" id="ComparePrice-product-template" style="display:none;">
<span class="compareprice__label">Compare at </span>
<span class="money"></span>
</p>
<div class="price_unit">
<dt>
<span class="visually-hidden">Unit price</span>
</dt>
<dd class="price-unit-price"><span data-unit-price=""></span><span aria-hidden="true"></span><span class="visually-hidden">pe
r </span><span data-unit-price-base-unit=""></span></dd>
</div><div class="product_policies rte" data-product-policies=""></div></div>]
```

```
In [8]: 1 Product_Price = []
2 for i in range(0,len(Price)):
3     Product_Price.append(Price[i].get_text())
4 Product_Price

Out[8]: ['\n₹ 12,750\nMRP inclusive of all taxes\n\nCompare at \n\n\nUnit price\n\n/per\xa0\n']
```

### 5. Access Rynox Jacket Features Using Code.

#### Product Features

```
In [23]: 1 feature = soup.find("div",class_="product_features").text
2 feature.strip().replace('\n\n\n\n\n',' ')

Out[23]: 'Product Features Tap on features to know more IMPACT PROTECTION \nABRASION RESISTANCE \nVENTILATION \nALL-SEASON COM
PATIBILITY \nVISIBILITY \nFIT AND COMFORT'
```

## 6. Access Customers Names.

### ACCESS CUSTOMERS NAMES

```
In [10]: 1 customers = soup.find_all("span",class_="jdgm-rev__author-wrapper")
2         customers

Out[10]: [<span class="jdgm-rev__author-wrapper"> <span class="jdgm-rev__author">Niladri P.</span> <span class="jdgm-rev__location"> </span> </span>,
<span class="jdgm-rev__author-wrapper"> <span class="jdgm-rev__author">Pavithra Krishna</span> <span class="jdgm-rev__location"> </span> </span>,
<span class="jdgm-rev__author-wrapper"> <span class="jdgm-rev__author">Nikhil Manoharan</span> <span class="jdgm-rev__location"> </span> </span>,
<span class="jdgm-rev__author-wrapper"> <span class="jdgm-rev__author">Yogesh P.</span> <span class="jdgm-rev__location"> </span> </span>,
<span class="jdgm-rev__author-wrapper"> <span class="jdgm-rev__author">Akhil M.</span> <span class="jdgm-rev__location"> </span> </span>]

In [11]: 1 cust_name = []
2         for i in range(0,len(customers)):
3             cust_name.append(customers[i].get_text())
4         cust_name

Out[11]: [' Niladri P. ',
' Pavithra Krishna ',
' Nikhil Manoharan ',
' Yogesh P. ',
' Akhil M. ']
```

## 7. Arrange Customers Reviews.

### ACCESS REVIEW

```
In [24]: 1 Comment = soup.find_all("div",class_="jdgm-rev__body")
2         Comment

Out[24]: [<div class="jdgm-rev__body"><p>This is a jacket which is been designed for Touring purpose and which can be used in summer as well as in winters. One who ride and tour this is the must go jacket.</p></div>,
<div class="jdgm-rev__body"><p>I'm extremely satisfied with the Stealth Evo 3 riding jacket. Not only does it provide high level of protection that I can trust, but it also offers exceptional comfort during long rides. The attention to detail in its design is evident, and it's impressive how well it performs. Overall, the Stealth Evo 3 has exceeded my expectations and is a fantastic choice for any rider looking for a comfortable and protective jacket for long rides.</p></div>,
<div class="jdgm-rev__body"><p>One of the best one I used ever. It's very comfortable with all uses. I suggest this one for all riders.</p></div>,
<div class="jdgm-rev__body"><p>I used it winter's it protected me from cold , I used it summer it has enuff ventilation. The guards give a perfect protection . The looks are amazing . It fits well. Easy to clean and it becomes new after wash . Rynox team rocks</p></div>,
<div class="jdgm-rev__body"><p>Awesome riding gear, Rynox a brand which can be trusted without any doubt</p></div>]

In [13]: 1 Review_Comment = []
2         for i in range(0,len(Comment)):
3             Review_Comment.append(Comment[i].get_text())
4         Review_Comment

Out[13]: ['This is a jacket which is been designed for Touring purpose and which can be used in summer as well as in winters. One who ride and tour this is the must go jacket.',
'I'm extremely satisfied with the Stealth Evo 3 riding jacket. Not only does it provide high level of protection that I can trust, but it also offers exceptional comfort during long rides. The attention to detail in its design is evident, and it's impressive how well it performs. Overall, the Stealth Evo 3 has exceeded my expectations and is a fantastic choice for any rider looking for a comfortable and protective jacket for long rides.',
'One of the best one I used ever. It's very comfortable with all uses. I suggest this one for all riders',
'I used it winter's it protected me from cold , I used it summer it has enuff ventilation. The guards give a perfect protection . The looks are amazing . It fits well. Easy to clean and it becomes new after wash . Rynox team rocks',
'Awesome riding gear, Rynox a brand which can be trusted without any doubt']
```

## 8. Access Ratings.

### ACCESS RATINGS

```
In [14]: 1 rating_one = soup.find("span",class_="jdgm-rev__rating")['data-score']
2         rating_one

Out[14]: '4'

In [15]: 1 ratings = [span['aria-label'] for span in soup.find_all("span",class_="jdgm-rev__rating")]
2         ratings

Out[15]: ['4 star review',
'5 star review',
'5 star review',
'5 star review',
'5 star review']
```

## 9. Arrange All Data Using Pandas Library.

### ARRANGE ALL DATA USING PANDAS LIBRARY

```
In [16]: 1 import pandas as pd
2 df = pd.DataFrame({"Customers_Name": cust_name,
3                   "Review_Comments": Review_Comment,
4                   "Review_Stars": ratings})
5 print(df)
```

	Customers_Name	Review_Comments	Review_Stars
0	Niladri P.	This is a jacket which is been designed for To...	4 star review
1	Pavithra Krishna	I'm extremely satisfied with the Stealth Evo 3...	5 star review
2	Nikhil Manoharan	One of the best one I used ever. It's very com...	5 star review
3	Yogesh P.	I used it winter's it protected me from cold ,...	5 star review
4	Akhil M.	Awesome riding gear, Rynox a brand which can b...	5 star review

## 10. DATASET Information :-

```
In [26]: 1 df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   Customers_Name  5 non-null     object
1   Review_Comments 5 non-null     object
2   Review_Stars    5 non-null     object
dtypes: object(3)
memory usage: 248.0+ bytes
```

## 11. Arrange All Data Into CSV File.

### Export data to CSV file

```
In [17]: 1 df.to_csv("Rynox_Gear_reviews.csv",index = False)
```

```
In [18]: 1 import os
2 os.getcwd()
```

```
Out[18]: 'C:\\Users\\KARAN'
```

```
In [27]: 1 df1 = pd.read_csv("Rynox_Gear_reviews.csv")
2 df1
```

```
Out[27]:
```

	Customers_Name	Review_Comments	Review_Stars
0	Niladri P.	This is a jacket which is been designed for To...	4 star review
1	Pavithra Krishna	I'm extremely satisfied with the Stealth Evo 3...	5 star review
2	Nikhil Manoharan	One of the best one I used ever. It's very com...	5 star review
3	Yogesh P.	I used it winter's it protected me from cold ,...	5 star review
4	Akhil M.	Awesome riding gear, Rynox a brand which can b...	5 star review



## 12. Store above all data in CSV file and again open this file in application.

A	B	C	D
Customers_Name	Review_Comments	Review_Stars	
Niladri P.	This is a jacket which is been designed for Touring purpose and which can be used in summer as well as in winters. One who ride and tour t	4 star review	
Pavithra Krishna	I'm extremely satisfied with the Stealth Evo 3 riding jacket. Not only does it provide high level of protection that I can trust, but it also offe	5 star review	
Nikhil Manoharan	One of the best one I used ever. It's very comfortable with all uses. I suggest this one for all riders	5 star review	
Yogesh P.	I used it winter's it protected me from cold , I used it summer it has enuff ventilation. The guards give a perfect protection . The looks are z	5 star review	
Akhil M.	Awesome riding gear, Rynox a brand which can be trusted without any doubt	5 star review	

## **CONCLUSION:-**

In conclusion, the web scraping project focused on extracting customer Product Name, Price, Features and Reviews from the RynoxGear website has been executed successfully using Python. Data scraping is a powerful technique to automatically extract information from websites. It is important to use web scraping responsibly and be mindful of the website's performance to avoid any legality issues.

This is all for my RynoxGear web scraping project.

