

Forward School

Program Code: J620-002-4:2020

Program Name: FRONT-END SOFTWARE DEVELOPMENT

Title : Introduction of BeautifulSoup - Part 2

Name: Chong Mun Chen

IC Number: 960327-07-5097

Date : 3/7/2023

Introduction : Learning to Web Scrape with BeautifulSoup

Conclusion : Getting adept at using the BeautifulSoup package and its methods.

P13 - Introduction of BeautifulSoup - Part 2

BeautifulSoup is a library for parsing HTML and XML documents.

Requests with BeautifulSoup (a parsing library) are the best package tools for **small and quick web scraping**.

<https://www.crummy.com/software/BeautifulSoup/>
[\(https://www.crummy.com/software/BeautifulSoup/\)](https://www.crummy.com/software/BeautifulSoup/)

<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
[\(https://www.crummy.com/software/BeautifulSoup/bs4/doc/\)](https://www.crummy.com/software/BeautifulSoup/bs4/doc/)

Installation

pip3 install bs4

Methods

Find_All

Find

```
In [1]: ┏ ┌ from bs4 import BeautifulSoup
      ┌ import requests
```

```
In [2]: ┏ ┌ #page = requests.get("https://www.ebay.com.my/sch/Car-Truck-Parts/6030/i.htm
      page = requests.get("https://www.ebay.com.my/sch/Car-Truck-Parts/6030/i.htm
      page.status_code
```

Out[2]: 200

```
In [3]: ┏ ┌ page.content
```

```
Out[3]: b'<!DOCTYPE html><!--[if IE 9]><html class="ie9" lang="en"><![endif]-->
<!--[if gt IE 9]><!--<html lang="en"><!--<![endif]--><!--M#s0-2--><noscript
class=x-page-config></noscript><!--M/--><head><meta http-equiv=X-UA-Compatible content=IE=edge><script>"use strict";if(window.PerformanceObserver&&performance&&performance.mark&&performance.getEntriesByName){window.SRP=window.SRP||{};var paintObserver=new window.PerformanceObserver((function(e){var r=e.getEntries();if(r.sort((function(e,r){return e.startTime-r.startTime}),r&&!(r.length<2)){var n=r[1].startTime;window.SRP.TTI_TIMER={lastInteractiveWindow:n};var t=new window.PerformanceObserver((function(e){for(var r=e.getEntries(),i=0,a=r.length;i<a;i++)r[i].startTime-n>=5e3&&(window.SRP.TTI_TIMER.timeToInteract=n,t.disconnect()),n=r[i].startTime+r[i].duration,window.SRP.TTI_TIMER.lastInteractiveWindow=n});t.observe({entryTypes:["longtask"]}),paintObserver.disconnect()}));paintObserver.observe({entryTypes:["paint"]})}};</script><meta name="description" content="Find Proton from a vast selection of Car & Truck Parts & Accessories. Get great deals on eBay!" /><title>Proton in Car & Truck Parts & Accessories | eBay</title><meta name="referrer" content="unsafe-url" /><meta name="y_key" content="cdcb aefb6bbb375f" /><meta name="msvalidate.01" content="34E98E6F27109BE1A9D
CE10C50E5E522" /><link rel="stylesheet" href="https://www.ebay.com.my/sch/Car-Truck-Parts/6030/i.htm
page.css" />'
```

In [5]:

```
soup = BeautifulSoup(page.content, 'html.parser')
soup
```

WINDOW.SRP.ISSAMPLEFORMONITORING = true;
}
</script><!--M#s0-10-1--><noscript></noscript><!--M/--><!--M#s0-11--><d
iv class="x-header"><script type="text/javascript">var GHpre = {si:fals
e}</script><style>.gh-hide-if-nocss{display:none;}.gh-ar-hdn{color:#ff
f}</style> <div class="gh-acc-exp-div gh-hide-if-nocss"><a class="gh-ac
c-a" href="#mainContent" id="gh-hdn-stm">Skip to main content</div>
<!--[if lt IE 9]><div id="gh" role="banner" class="gh-IE8 gh-flex gh-pr
e-js gh-w "><![endif]--><header class="gh-ui-6-5 gh-flex gh-pre-js gh-w
gh-sch-prom" data-treatment="" id="gh" role="banner"><table class="gh-t
bl" role="presentation"><tbody><tr><td class="gh-td"><a _sp="m570.1258
6" href="https://www.ebay.com.my" id="gh-la"></td><td class="gh-td"><div
class="gh-hide-if-nocss" id="gh-shop"><button aria-controls="gh-sbc-o"
aria-expanded="false" class="gh-control" id="gh-shop-a"> Shop by catego
ry<i class="gh-sprRetina" id="gh-shop-ei"></i></button><div id="gh-sbc
-o"><h2 class="gh-ar-hdn">Shop by category</h2></div></div></td><td clas
s="gh-td-s"><form action="/sch/i.html" id="gh-f" method="get"><input na

Observe that the downloaded page

In [6]:

```
print(soup.prettify())
```

<!DOCTYPE html>
<!--[if IE 9]><html class="ie9" lang="en"><![endif]-->
<!--[if gt IE 9]><!-->
<html lang="en">
<!--<![endif]-->
<!--M#s0-2-->
<noscript class="x-page-config">
</noscript>
<!--M/-->
<head>
<meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<script>
 "use strict";if(window.PerformanceObserver&&performance&&performanc
e.mark&&performance.getEntriesByName){window.SRP=window.SRP||{};var pai
ntObserver=new window.PerformanceObserver((function(e){var r=e.getEntry
es();if(r.sort((function(e,r){return e.startTime-r.startTime})),r&&(r.
length<2)){var n=r[1].startTime;window.SRP.TTI_TIMER={lastInteractiveWi
ndow:n};var t=new window.PerformanceObserver((function(e){for(var r=e.g
etEntries(),i=0,a=r.length;i<a;i++)r[i].startTime-n>=5e3&&(window.SRP.T
TI_TIMER.lastInteractiveWindow=n)});t.observe({type:"idle",timeDelta:500});}});}};

Observe that the HTML doc is now a human readable document

In [7]: `list(soup.children)`

```
inwindow){var g=window._tti={e:[]};g.o=new PerformanceObserver(function(l){g.e=g.e.concat(l.getEntries())});g.o.observe({entryTypes:["longtask"]})}();</script><link href="https://ir.ebaystatic.com/rs/v/1bhvbn1jr i3ldkqaxcguwmjwuw.css?proc=DU:N" rel="stylesheet" type="text/css"/><link href="https://ir.ebaystatic.com/rs/c/inception-Robg8qPC.css" rel="stylesheet"/>
<link href="https://ir.ebaystatic.com/rs/c/search-page-large-rhm3pPJt.css" rel="stylesheet"/><!--M#s0-5-9--><noscript data-marko-key="@_wbind s0-5-9" id="s0-5-9"></noscript><!--M/--><style type="text/css">
:root {
    --sd-s-y14d697: inline;
    --sd-s-jst2988: none;
    --p4-s-y14d697: 100%;
    --p4-s-jst2988: 98%;
}
span.s-y14d697 {
    display: inline;
}
span.s-jst2988 {
    display: none.
}
```

Now go back to Ebay Page and observe the listing for repeatable items

Observe the listed items and corresponding HTML tags

We want to grab all the product listed on the page

In [8]: `soup_list = soup.findAll("li", attrs={"class": "s-item s-item__pl-on-bottom"})`
soup_list

```
<li class="s-item s-item__pl-on-bottom" data-view="mi:1686|iid:1" data-viewport='{"trackableId":"01GDF2P4GS7EBZ8M2AWN707B"}' id=""><div class="s-item__wrapper clearfix"><!--F#f_0--><div class="s-item__image-section"><div class="s-item__image"><a _sp="" aria-hidden="true" data-carousel-tabindex="-1" data-interactions='[{"actionKind": "NAVSRC", "interactionName": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAAgn3NDAwDE5BV1NSQwA="}]' data-s-y14d697='{"eventFamily": "LST", "eventAction": "ACTN", "actionKind": "NAVSRC", "actionKinds": ["NAVSRC"], "operationId": "", "flushImmediately": false, "eventProperty": {"trackableId": "01GDF2P4GS7EBZ8M2AWN707B", "interactionName": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAAgn3NDAwDE5BV1NSQwA="}}' href="https://ebay.com/itm/123456?hash=item28cae0a3a:g:E3kAAOSwlGJiMikD&amdata=enc%3AAQAHAAAAAsJoWXGf0hxNZspTmh8%2FTJCCurAWCHuXJ2Xi3S9cwXL6BX04zSEiVaDMCvsUbApftgXEAHGJU1ZGugZ0%2FnW1U7Gb6vgol%2BmXlqCbLkw0Zff3AUAK8YvJ5B4%2BnhFA7ID4dxpYs4jjExEnN5SR2g1mQe7QtLkmGt%2FZ%2FbH2W62cXPuKbf550ExbnBP02QJyZTXYCuw5KVkMdFMDuoB4p3FwJKcSPzez5kyQyVjyiIq6PB2q%7Ctkp%3AB1BMULq7kqyXYA" tabindex="-1" target="_blank"><div class="s-item__image-wrapper image-treatment"><img alt="Shop on eBay" fetchpriority="high" loading="eager" onerror="window.SRP.metrics.incrementCounter('imageLoadError');" onload="if(this.width === 80 &amp; this.height
```

```
In [9]: ┆ for items in range(len(soup_list)):  
    print(items)
```

```
0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19
```

```
In [10]: soup_list = soup.find("li", attrs={"class": "s-item s-item__pl-on-bottom"})
soup_list
```

```
Out[10]: <li class="s-item s-item__pl-on-bottom" data-view="mi:1686|iid:1" data-viewport='{"trackableId": "01GDF2P4GS7EBZ8M2AWN707B"}' id=""><div class="s-item__wrapper clearfix"><!--F#f_0--><div class="s-item__image-section"><div class="s-item__image"><a _sp="" aria-hidden="true" data-carousel-tabindex="-1" data-interactions='[{"actionKind": "NAVSRC", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAg3NDAwDE5BV1NSQwA="}]' data-s-y14d697='{"eventFamily": "LST", "eventAction": "ACTN", "actionKind": "NAVSRC", "actionKinds": ["NAVSRC"], "operationId": "", "flushImmediately": false, "eventProperty": {"trackableId": "01GDF2P4GS7EBZ8M2AWN707B", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAg3NDAwDE5BV1NSQwA="}}' href="https://ebay.com/itm/123456?hash=item28cae0a3a:g:E3kAAOSwlGJiMikD&amdata=enc%3AAQAHAAAAsJowXGf0hxNZspTmhb8%2FTJCCurAWCHuXJ2Xi3S9cwXL6BX04zSEiVaDMCvsUbApftgXEAHGJU1ZGugZ0%2FnW1U7Gb6vgoL%2BmXlqCbLkwoZff3AUAK8YvJ5B4%2BnhFA7ID4dxpYs4jjExEnN5SR2g1mQe7QtLkmGt%2FZ%2FbH2W62cXPuKbf550ExbnBP02QJyZTXYCuw5KVkMdFMDuoB4p3FwJKcSPzez5kyQyVjyiIq6PB2q%7Ctkp%3AB1BMULq7kqyXYA" tabindex="-1" target="_blank"><div class="s-item__image-wrapper image-treatment"></div></a></div></div><div class="s-item__info clearfix"><a _sp="" class="s-item__link" data-interactions='[{"actionKind": "NAVSRC", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAg3NDAwDE5BV1NSQwA="}]' data-s-y14d697='{"eventFamily": "LST", "eventAction": "ACTN", "actionKind": "NAVSRC", "actionKinds": ["NAVSRC"], "operationId": "", "flushImmediately": false, "eventProperty": {"trackableId": "01GDF2P4GS7EBZ8M2AWN707B", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAg3NDAwDE5BV1NSQwA="}}' href="https://ebay.com/itm/123456?hash=item28cae0a3a:g:E3kAAOSwlGJiMikD&amdata=enc%3AAQAHAAAAsJowXGf0hxNZspTmhb8%2FTJCCurAWCHuXJ2Xi3S9cwXL6BX04zSEiVaDMCvsUbApftgXEAHGJU1ZGugZ0%2FnW1U7Gb6vgoL%2BmXlqCbLkwoZff3AUAK8YvJ5B4%2BnhFA7ID4dxpYs4jjExEnN5SR2g1mQe7QtLkmGt%2FZ%2FbH2W62cXPuKbf550ExbnBP02QJyZTXYCuw5KVkMdFMDuoB4p3FwJKcSPzez5kyQyVjyiIq6PB2q%7Ctkp%3AB1BMULq7kqyXYA" tabindex="-1" target="_blank"><div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0-->Shop on eBay<!--F--></span></div><span class="clipped">Opens in a new window or tab</span></a><div class="s-item__subtitle"><!--F#f_0--><span class="SECONDARY_INFO">Brand New</span><!--F--></div><div class="s-item__details clearfix"><div class="s-item__detail s-item__detail--primary"><span class="s-item__price"><!--F#f_0--><!--F#f_0-->$20.00<!--F--><!--F--></span></div><div class="s-item__detail s-item__detail--primary"><span class="s-item__purchase-options"><!--F#f_0-->or Best Offer<!--F--></span></div><div class="s-item__detail s-item__detail--primary"><span><div aria-labelledby="s-y14d697" class="s-y14d697_s-jst2988" role="text"><span aria-hidden="true"><wbr/>Sponsored</span></div></span><span class="s-item__space_bar"></span></div></div><!--F--></div></li>
```

```
In [11]: #show only 1
soup_list = soup.find("li", attrs={"class":"s-item s-item__pl-on-bottom"})
soup_list
```

```
Out[11]: <li class="s-item s-item__pl-on-bottom" data-view="mi:1686|iid:1" data-viewport='{"trackableId":"01GDF2P4GS7EBZ8M2AWN707B"}' id=""><div class="s-item__wrapper clearfix"><!--F#f_0--><div class="s-item__image-section"><div class="s-item__image"><a _sp="" aria-hidden="true" data-carousel-tabindex="-1" data-interactions='[{"actionKind": "NAVSRC", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAAgn3NDAwDE5BV1NSQwA="}]' data-s-y14d697='{"eventFamily": "LST", "eventAction": "ACTN", "actionKind": "NAVSRC", "actionKinds": ["NAVSRC"], "operationId": "", "flushImmediately": false, "eventProperty": {"trackableId": "01GDF2P4GS7EBZ8M2AWN707B", "interaction": "wwFVrK2vRE01hQY0MREUzME5KNVY2VE1XNkJEMTUAAAgn3NDAwDE5BV1NSQwA="}}' href="https://ebay.com/itm/123456?hash=item28cae0a3a:g:E3kAAOSwlGJiMikD&amdata=enc%3AAQAHAAAAAsJoWxGf0hxNZspTmhB%2FTJCCurAWChuXJ2Xi3S9cwXL6BX04zSEiVaDMCvsUbApftgXEAHGU1ZGugZ0%2FnW1U7Gb6vg0L%2BmXlqCbLkwOzFF3AUAK8YvJ5B4%2BnhFA7ID4dxpYs4jjExEnN5SR2g1mQe7QtLkmGt%2FZ%2FbH2W62cXPuKbf550ExbnBP02QjyZTXYCuw5KVkMdFMDuoB4p3FwJKcSPzez5kyQyVjyiIq6PB2q%7Ctkp%3AB1BMULq7kqyXYA" tabindex="-1" target="_blank"><div class="s-item__image-wrapper image-treatment"><img alt="Shop on eBay" fetchpriority="high" loading="eager" onerror="window.SRP.metrics.incrementCounter('imageLoadError');" onload=";if(this.width === 80 &amp; this.height<80) { window.SRP.metrics.incrementCounter('imageLoadError');}">
```

```
In [12]: # FindAll but get the Nth item from the Match
soup_list = soup.findAll("li", attrs={"class": "s-item s-item__pl-on-bottom"})
soup_list
```

```
Out[12]: <li class="s-item s-item__pl-on-bottom" data-view="mi:1686|iid:1" data-viewport='{"trackableId":"01H4RBQ710AB0AQCHTJNRVNGA9"}' id="item54d3c37da6"><div class="s-item__wrapper clearfix"><!--F#f_0--><div class="s-item__image-section"><div class="s-item__image"><a _sp="p2351460.m1686.17400" aria-hidden="true" data-carousel-tabindex="-1" data-interactions='[{"actionKind": "NAVSRC", "interaction": "wwFVrK2vRE01hQ0MDFINFJCUTc0OUtxODZLWENGMEREDJQN1o0MDFINFJCUTcxMEFCMEFRQ0hUSk5SVk5HQTkAAAg3NDAwDE5BV1NSQwA="}]' data-s-y14d697='{"eventFamily": "LST", "eventAction": "ACTN", "actionKind": "NAVSRC", "actionKinds": ["NAVSRC"], "operationId": "2351460", "flushImmediately": false, "eventProperty": {"$1": "241248621493171"} }' href="https://www.ebay.com.my/itm/364330057126?hash=item54d3c37da6:g:XMKAAOSwGD5koEyF&amdata=enc%3AAQAIAAAwPER7bPUorUVIp8aqU0zE%2BwY%2BCQyLTxWtm9HV%2BujcJ1cDf%2FCRc5Qw6m7hHY5YshQ%2FJX1bvoDAn3hPBeXdzKEUMjs4hQ%2BSnB%2Fmj57%2Bmc0gFDYyE5SzCFi0wDbrqyPaQ364J9YeAm%2Fi%2ByUXZCsR2M9zgjN1%2B1RpIG0SXLB0Ja76ZURypbPJqGufa%2FL01jtG6pzNG5SD50c2Qt8AwsRIsigSYmlBF02NMokvtDfWT7pg4qlk3pp6CSfb%2FvLGF5aVrSoWA%3D%3D%7Ctkp%3Abk9SR8jw3IumYgtabindex="-1" target="_blank"><div class="s-item__image-wrapper image-treatment"><img alt="Proton Saga Mitsuboshi CPK865 Original Fan Belt Brand New" fetchpriority="high" loading="eager" onerror="window.SRP.metric
```

```
In [13]: soup_list = soup.findAll("li", attrs={"class":"s-item s-item__pl-on-bottom"})
for items in range(len(soup_list)):
    print(soup_list[items].find("div", attrs={"class":"s-item__title"}))

<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0-->Shop on eBay<!--F/-></span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0-->Proton Saga Mitsubishi CPK865 Original Fan Belt Brand New<!--F/->
</span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0-->4D68 For MITSUBISHI GALANT OR PROTON PERSONA 400 Engine Rebuild Kit
s Auto Engine<!--F/-></span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0-->Tail Lamp Proton Perdana V6 RH<!--F/-></span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0--><span class="LIGHT_HIGHLIGHT">New Listing</span><!--F/-><!--F#f_0-->Timing Belt fits PROTON SAVVY 1.2 05 to 12 Gates Genuine Top Quality
Guaranteed<!--F/-></span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0--><span class="LIGHT_HIGHLIGHT">New Listing</span><!--F/-><!--F#f_0-->ENGINE TIMING BELT CAM BELT OYODO 30R5024-0YO P FOR MITSUBISHI LANCER
V,GALANT V<!--F/-></span></div>
<div class="s-item__title"><span aria-level="3" role="heading"><!--F#f_0--><span class="LIGHT_HIGHLIGHT">New Listing</span><!--F/-><!--F#f_0-->Timing Belt fits PROTON SAVVY 1.2 05 to 12 Gates Genuine Top Quality
Guaranteed<!--F/-></span></div>
```

```
In [14]: ┆ soup_list = soup.findAll("li", attrs={"class":"s-item s-item__pl-on-bottom"})
      for items in range(len(soup_list)):
          print(soup_list[items].find("div", attrs={"class":"s-item__title"}).te
```

Shop on eBay

Proton Saga Mitsubishi CPK865 Original Fan Belt Brand New
4D68 For MITSUBISHI GALANT OR PROTON PERSONA 400 Engine Rebuild Kits Auto Engine
Tail Lamp Proton Perdana V6 RH
New Listing Timing Belt fits PROTON SAVVY 1.2 05 to 12 Gates Genuine Top Quality Guaranteed
New Listing ENGINE TIMING BELT CAM BELT OYODO 30R5024-OYO P FOR MITSUBISHI LANCER V, GALANT V
New Listing V-Ribbed Belts for CITROËN FIAT LADA INA FB 4PK698
New Listing ZAHNRÄMEN STEUERRÄMEN OYODO 30R5024-OYO P NEU OE QUALITÄT
New Listing TIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT842
5PK865 GATES V-Ribbed Belt for BMW, BYD, FERRARI, MITSUBISHI, PROTON, TOYOTA, Z HONGHUA
New Listing TIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT812
New Listing TIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT821
KEILRIPPERÄMEN FÜR MITSUBISHI NISSAN PROTON INA FB 4PK760
New Listing OYODO ZAHNRÄMEN STEUERRÄMEN 30R5024-OYO P FÜR MITSUBISHI LAN CER V, GALANT V 2L
1x Meyle 0500040763 V-Ribbed Belt for Mitsubishi Nissan Rover Suzuki Proton GATES Micro-V Drive Belt for Proton Satria 4G93 1.8 August 1998 to March 2000
Timing Belt for Proton Persona 4cyl 4G15 1.5L - 4G13 1.3L 1996 to 2004
Genuine SKF Timing Belt Kit for Proton Savvy 1149cc D4F 1.2 (01/2006-12/2014)
SKF VKMA 36067 V-Ribbed Belt Set for DACIA, NISSAN, PROTON, RENAULT
INA 529 0280 10 V-Ribbed Belt Set for DACIA, NISSAN, PROTON, RENAULT
DAYCO TIMING BELT KIT for PROTON M21 10/97-11/00 1.8L 4CYL 16V DOHC MPFI 4G93
New Listing MITSUBOSHI ZAHNRÄMEN STEUERRÄMEN 158XR32 L FÜR MITSUBISHI LAN CER V, GALANT V
Balance Shaft Belt fits HYUNDAI LANTRA J-1 1.6 1.8 90 to 95 Gates 2335633 110 New
New Listing MITSUBOSHI ZAHNRÄMEN STEUERRÄMEN 92ZBS19 L FÜR HYUNDAI PONY EXCEL, PONY
New Listing ENGINE TIMING BELT CAM BELT BOSCH 1 987 949 518 P NEW OE REPLACEMENT
DAYCO TIMING BELT KIT for PROTON Satria 01/07-08/12 1.6L 4cyl DOHC MPFI B S S4PH
GATES Timing Belt Kit To Suit Proton M21 1.8 EXi Petrol Coupe
Genuine OEM Crank Sensor For Proton (SC576GEN)
Dayco Timing Belt T889 fits MITSUBISHI/ PROTON 94510
New Listing Genuine Proton Saga Iswara Fuel Filler Hose MB400890
ENGINE TIMING BELT CAM BELT GATES 5577XS P FOR PROTON SAVVY 1.2 1.2L 55KW
Dayco Timing Belt T935 fits MITSUBISHI/ PROTON 94494
530 0518 10 INA Timing Belt Set for MITSUBISHI, PROTON
1 987 948 364 BOSCH V-Ribbed Belt for DACIA, NISSAN, PEUGEOT, PROTON, RENAULT
ZAHNRÄMEN FÜR MITSUBISHI PROTON BOSCH 1 987 949 429
Stabilizer Link Set of 2pcs Fit For Suzuki Ertiga Proton Ertiga
Timing Belt fits PROTON WIRA 2.0D 96 to 00 QH MD192319 Top Quality Guaranteed
GATES Micro-V Drive Belt for Proton Satria 4G15M 1.5 March 1996 to March 2000
Dayco Timing Belt T191 fits HYUNDAI/ MITSUBISHI/ PROTON 94216
GATES Timing Belt To Suit Proton Satria 1.6 Petrol
Timing Belt fits PROTON PERSONA 2.0D 96 to 99 4D68 ADL MD192319 Quality New

KEILRIPPENRIEMEN FÜR MITSUBISHI NISSAN PROTON MEYLE 050 004 0760
050 004 0763 MEYLE V-Ribbed Belt for MITSUBISHI,NISSAN,PROTON,ROVER,SUZUKI
ZAHNRIEMEN STEUERRIEMEN GATES 5550XS G FÜR PROTON IMPIAN 1.8 1.8L 93KW
GATES Timing Belt To Suit Proton Persona 313i Petrol
Timing Belt fits PROTON COMPACT C9 1.3 1.5 00 to 06 QH MD145813 Quality New
Timing Belt fits PROTON PERSONA 1.6 94 to 00 4G92 QH MD176386 Quality Guaranteed
GATES Timing Belt To Suit Proton Persona 1.6 Petrol
KEILRIPPENRIEMEN FÜR MITSUBISHI NISSAN PROTON CONTINENTAL CTAM 4PK760
ZAHNRIEMEN STEUERRIEMEN DAYCO 94195 I FÜR HYUNDAI LANTRA I 1.6L,1.8L 84KW,93KW
GATES Timing Belt Kit To Suit Proton Persona 313i Petrol
GATES Timing Belt To Suit Proton M21 1.8 EXi Petrol Coupe
GATES Timing Belt Kit To Suit Proton Persona 316 Petrol
ZAHNRIEMEN STEUERRIEMEN SUN A444YU25MM L FÜR PROTON SALOON 1.5,1.3 1.5L,1.3L
ZAHNRIEMEN FÜR MITSUBISHI PROTON GATES 5350XS
GATES Timing Belt Kit To Suit Proton Persona 315 GLi/GLSi Petrol
V-RIBBED FAN BELT FOR PROTON SATRIA 5PK1148 OEM QUALITY
New ListingProton Gen2 Persona 2005-2010 Hatch Rear Wiper Arm
GATES 5257XS Timing Belt for ,CHERY,FORD AFRICA,GREAT WALL,HYUNDAI,KIA,LANDWIND
Timing Belt fits MITSUBISHI DELICA 2.0 96 to 00 QH MB182293 MD197141 MD338288
ZAHNRIEMEN STEUERRIEMEN BOSCH 1 987 949 518 P FÜR PROTON SAVVY 1.2 1.2L 55KW
Brand New DAYCO Timing Belt to fit Proton Persona 1995-1996
GATES Micro-V Drive Belt for Proton Compact 4G15 1.5 July 2000 to July 2006
3x2 High Quality Rubber Tie Rod End Ball Joint Dust Boots Dust Cover Boot Set

```
In [15]: ┆ soup_list = soup.findAll("li", attrs={"class":"s-item s-item__pl-on-bottom"})
      for items in range(len(soup_list)):
          print(soup_list[items].find("span", attrs={"class":"s-item__price"}).te
```

\$20.00
RM 18.00
RM 410.26
RM 250.00
RM 169.63
RM 247.70
RM 100.56
RM 202.40
RM 289.42
RM 203.36
RM 289.42
RM 206.56
RM 90.69
RM 207.17
RM 188.57
RM 124.23
RM 194.18
RM 616.25
RM 408.66
RM 402.02
RM 647.74
RM 207.17
RM 94.82
RM 121.04
RM 189.10
RM 616.89
RM 567.57
RM 686.94
RM 234.40
RM 118.31
RM 229.94
RM 268.33
RM 440.64
RM 188.80
RM 156.55
RM 455.68
RM 211.71
RM 110.68
RM 209.72
RM 339.31
RM 365.12
RM 75.49
RM 110.50
RM 252.56
RM 225.18
RM 128.79
RM 182.18
RM 302.29
RM 90.69
RM 141.10
RM 434.93
RM 302.29
RM 496.62
RM 176.57
RM 166.69
RM 404.08
RM 107.13

```
RM 50.90
RM 67.13
RM 164.01
RM 176.57
RM 209.75
RM 128.26
RM 74.08
```

In [16]:

```
# Exercise
# Find the Cheapest Item
import re

soup_list = soup.findAll("li", attrs={"class": "s-item s-item__pl-on-bottom"})
cheapest = []
for items in range(1, len(soup_list)):
    cheapest.append(float(re.sub(r'^0-9.', "", soup_list[items].find("span").text)))

# for items in range(len(soup_list)):
#     cheapest.append(double(soup_list[items].find("span", attrs={"class": "s-item__price"}).text))

#     x = soup_list[items].find("span", attrs={"class": "bold"}).text
#     cheapest.append(re.sub(r'^0-9.', '', x))

# print("Cheapest Item is: RM", min(cheapest))

# Should not work, why?
```

Cheapest Item is: RM 18.0

In [17]:

```
import re

input_value="RM 1,000.00"

result = re.sub(r"RM ", "", input_value)

print(result)
```

1,000.00

In [18]:

```
import re

input_value="RM 1,000.00"

result = re.sub(r'^0-9', '', input_value)

print(result)
```

100000

```
In [19]: ➤ import re

input_value="RM 1,000.00"

result = re.sub(r'^[^\d.]', '', input_value)

print(result)
```

1000.00

```
In [40]: ➤ # Exercise
# Find the Cheapest Item
# Use regex to remove the RM Suffix
import re

soup_list = soup.findAll("li", attrs={"class": "sresult lvresult clearfix li"})

cheapest = []

for items in range(len(soup_list)):
    cheapest.append(float(re.sub(r'^[^\d.]', '', soup_list[items].find("span" "s-item__price").text)))

print(cheapest)

print("The Cheapest is: ", min(cheapest))
```

[]

```
--> ValueError                                     Traceback (most recent call last)
t)
Cell In[40], line 15
      11     cheapest.append(float(re.sub(r'^[^\d.]', '', soup_list[items].f
ind("span", attrs={"class": "s-item__price"}).text)))
      12 print(cheapest)
--> 13 print("The Cheapest is: ", min(cheapest))
```

ValueError: min() arg is an empty sequence

```
In [21]: ➤ cheapest = [1,2,3,4,5]
min(cheapest)
```

Out[21]: 1

```
In [22]: ┏ soup_list = soup.findAll("li", attrs={"class":"s-item s-item__pl-on-bottom"}  
  for items in range(1, len(soup_list)):  
    print(soup_list[items].find("span", attrs={"role":"heading"}).text)  
    print()  
  ━
```

New ListingTiming Belt fits PROTON SAVVY 1.2 05 to 12 Gates Genuine Top Quality Guaranteed

New ListingENGINE TIMING BELT CAM BELT OYODO 30R5024-OYO P FOR MITSUBIS HI LANCER V,GALANT V

New ListingV-Ribbed Belts for CITROËN FIAT LADA INA FB 4PK698

New ListingZAHNRIEMEN STEUERRIEMEN OYODO 30R5024-OYO P NEU OE QUALITÄT

New ListingTIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT842

5PK865 GATES V-Ribbed Belt for BMW,BYD,FERRARI,MITSUBISHI,PROTON,TOYOT A,ZHONGHUA

New ListingTIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT812

New ListingTIMING BELT FOR MITSUBISHI PROTON CONTINENTAL CTAM CT821

KETLRIPPENRIEMEN FÜR MITSUBISHI NISSAN PROTON INA FB 4PK760

```
In [39]: ┏ soup_list = soup.findAll("li", attrs={"class":"s-item s-item__pl-on-bottom"}  
  for items in range(1, len(soup_list)):  
    print(soup_list[items].find("div", attrs={"class":"s-item__title"}).text)  
    print(soup_list[items].find("span", attrs={"class":"s-item__price"}).text)  
    print()  
  ━
```

GATES Timing Belt Kit To Suit Proton Persona 313i Petrol RM 434.93

GATES Timing Belt To Suit Proton M21 1.8 EXi Petrol Coupe RM 302.29

GATES Timing Belt Kit To Suit Proton Persona 316 Petrol RM 496.62

ZAHNRIEMEN STEUERRIEMEN SUN A444YU25MM L FÜR PROTON SALOON 1.5,1.3 1.5 L,1.3L RM 176.57

ZAHNRIEMEN FÜR MITSUBISHI PROTON GATES 5350XS RM 166.69

GATES Timing Belt Kit To Suit Proton Persona 315 GLi/GLSi Petrol RM 404.08

Exercise

Store the results obtained into a Dataframe and save it as CSV file

1. Column 1 is "Product Name"
2. Column 2 is "Product Price"

```
In [47]: ┆ import csv
import requests
from bs4 import BeautifulSoup

page = requests.get("https://www.ebay.com.my/sch/Car-Truck-Parts/6030/i.htm")
soup = BeautifulSoup(page.content, 'html.parser')
soup_list = soup.findAll("li", attrs={"class": "s-item s-item__pl-on-bottom"})

csvFile = open("ebay_products.csv", 'w', encoding='utf8', newline=' ')
writer = csv.writer(csvFile)
writer.writerow(["Product Name", "Product Price"])

try:
    for item in soup_list[1:]:
        product_name = item.find("div", attrs={"class": "s-item__title"}).text
        product_price = item.find("span", attrs={"class": "s-item__price"}).text
        writer.writerow([product_name, product_price])
finally:
    csvFile.close()
```