



# Advanced Web Scraping Using ParseHub

Table of Contents

Task 1: Scraping Images from a Website .....3

Task 2: Scrape a Password Protected Website .....7

Task 3: Scrape Through a List of Keywords.....10

Task 4: Scrape a list of URLs .....19

Task 5: Web Scraping using XPath .....26

## Task 1: Scraping Images from a Website

Now, let's proceed to another web scraping exercise. In this case, we will go over how to extract data from a website that goes beyond text on the page.

In this case, we will extract and download images from Amazon.

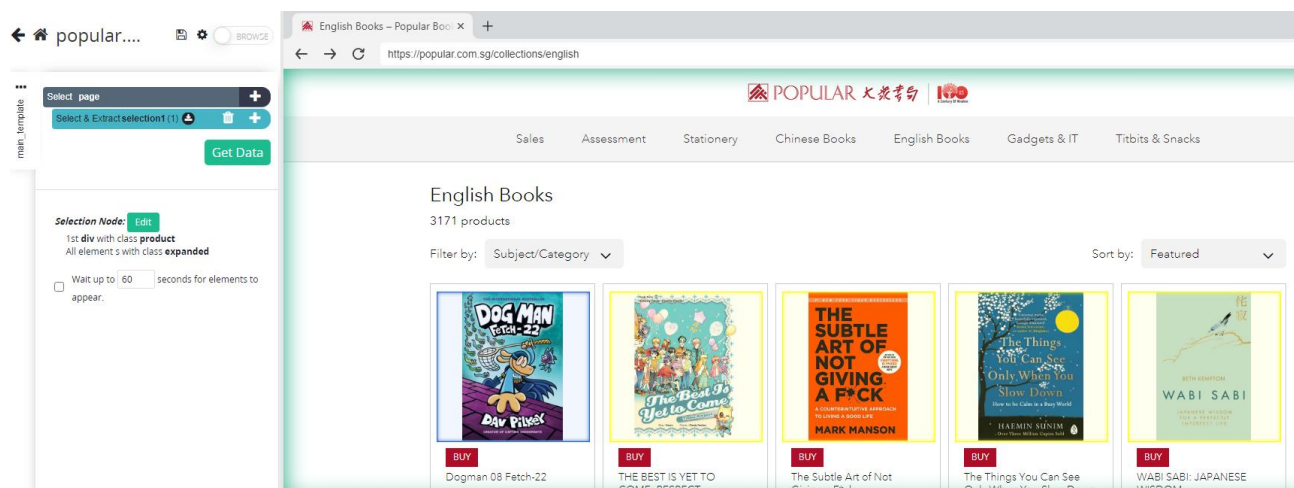
For this example, we will assume that we are interested in downloading every image for the first pages of results on Popular web site for "English book". This information could potentially be incredibly valuable for competitor analysis.

In ParseHub, click on "New Project" and enter the URL from the Amazon website that we will be scraping. In this case, the URL is: <https://popular.com.sg/collections/english>. Copy the link to the ParseHub.

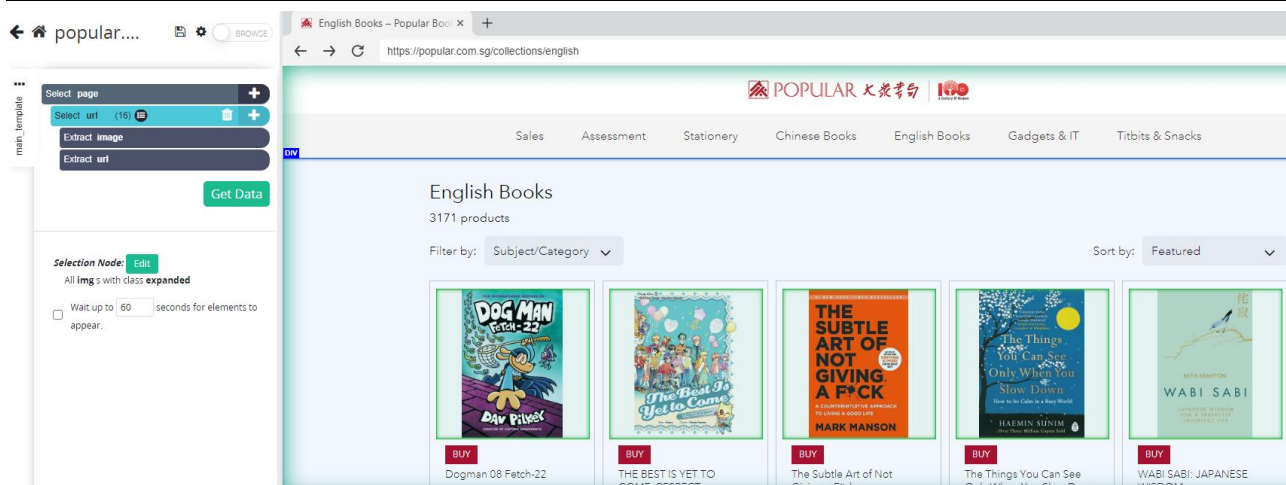
The webpage will now render in ParseHub and you will be able to choose the images you want to scrape.

### Select Images to Scrape

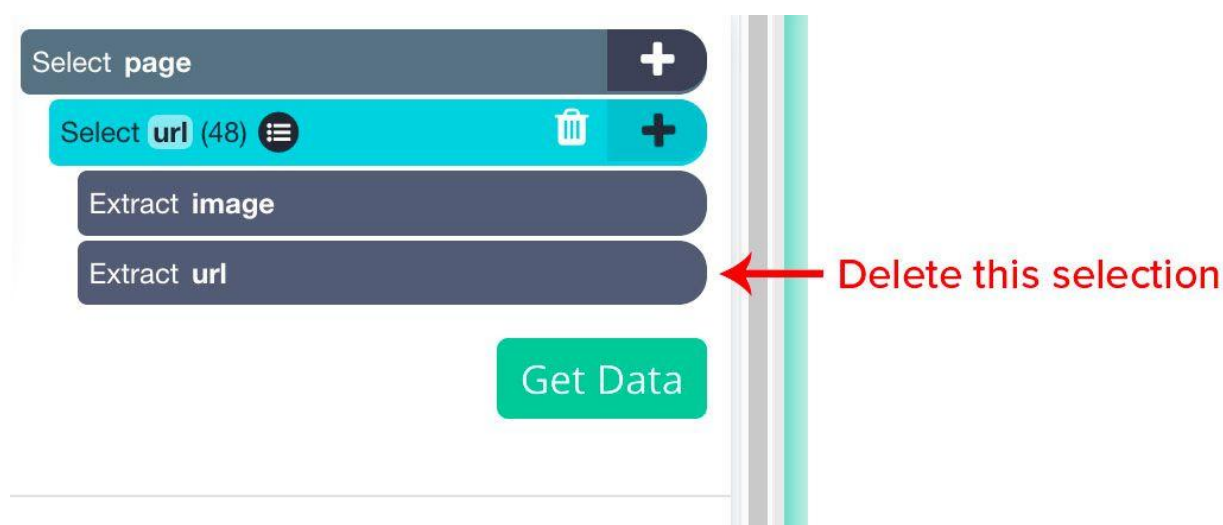
Begin by selecting the first image from the search results. It will then turn green, meaning it has been selected to be scrapped.



The rest of the images in the search results page will then turn yellow. Click on the second image to select all the images in the page. They will all turn green, which means they have been selected to be extracted.



Since these images also act as links to the book pages, ParseHub is extracting both the image URL and the link it is pointing to (book page). As a result, we will delete the URL selection from the left sidebar and only keep the image selection.



Now ParseHub will scrape every image URL for the first page of results.

## Scrape and Export Data

Now comes the fun part, we will let ParseHub run and extract the list of URLs for every image we have selected.

Click on the "Get Data" button on the left sidebar and run your project.

Once your data is ready to download, click on the CSV/Excel button. Now you can save and rename your file.

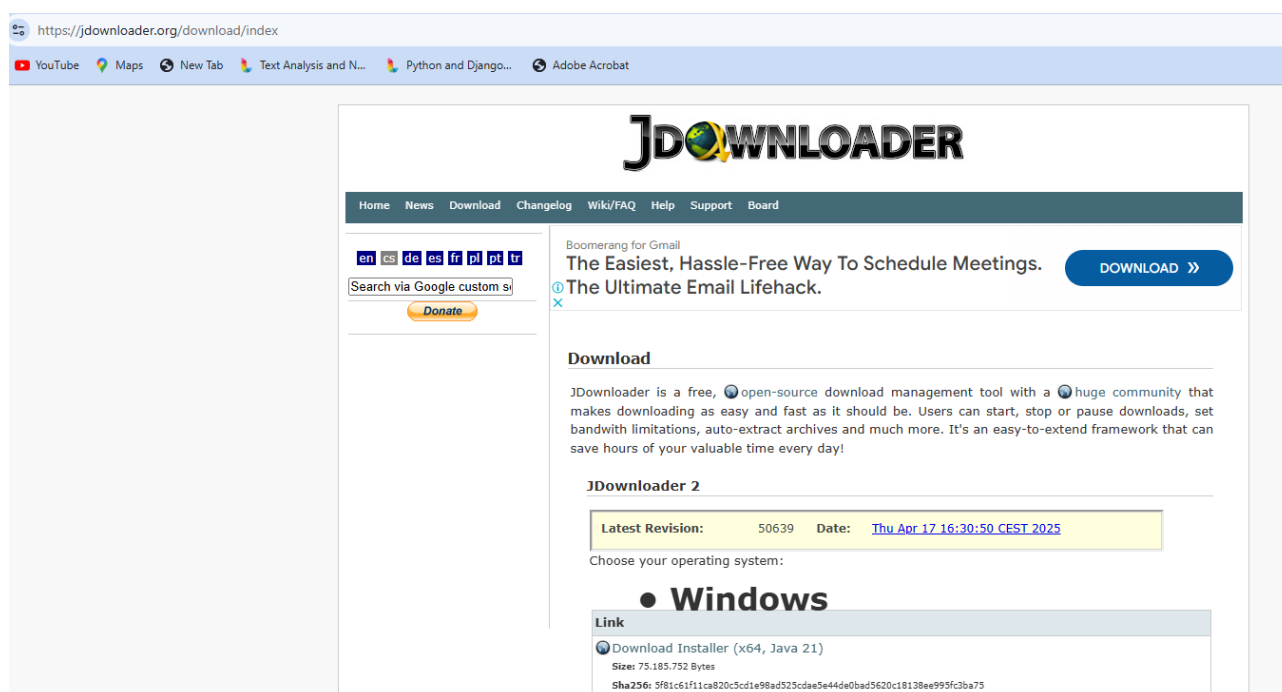
url\_image

[https://popular.com.sg/cdn/shop/files/download\\_0fc0490-ed65-465a-9393-c4e45c222d69.jpg?v=1744109479&width=338](https://popular.com.sg/cdn/shop/files/download_0fc0490-ed65-465a-9393-c4e45c222d69.jpg?v=1744109479&width=338)  
[https://popular.com.sg/cdn/shop/files/download\\_16af2d72-69a5-4499-ba40-4b5bef768c12.jpg?v=1743134272&width=825](https://popular.com.sg/cdn/shop/files/download_16af2d72-69a5-4499-ba40-4b5bef768c12.jpg?v=1743134272&width=825)  
[https://popular.com.sg/cdn/shop/files/download\\_ead4c06b-9efa-42ac-8995-0775d096d12f.jpg?v=1744111213&width=333](https://popular.com.sg/cdn/shop/files/download_ead4c06b-9efa-42ac-8995-0775d096d12f.jpg?v=1744111213&width=333)  
[https://popular.com.sg/cdn/shop/files/download\\_56abbcb85-f8c4-4d6e-bfb0-6dcb8c4874bd.jpg?v=1744111223&width=356](https://popular.com.sg/cdn/shop/files/download_56abbcb85-f8c4-4d6e-bfb0-6dcb8c4874bd.jpg?v=1744111223&width=356)  
[https://popular.com.sg/cdn/shop/files/download\\_faf95e91-2dee-4f90-a4ec-195439802d37.jpg?v=1742542383&width=316](https://popular.com.sg/cdn/shop/files/download_faf95e91-2dee-4f90-a4ec-195439802d37.jpg?v=1742542383&width=316)  
[https://popular.com.sg/cdn/shop/files/download\\_c82fd367-16f1-478a-a740-715a394bb024.jpg?v=1744110168&width=322](https://popular.com.sg/cdn/shop/files/download_c82fd367-16f1-478a-a740-715a394bb024.jpg?v=1744110168&width=322)  
[https://popular.com.sg/cdn/shop/files/download\\_3b901c0a-476d-4cdf-bbd0-8869535e3b3d.jpg?v=1742364618&width=326](https://popular.com.sg/cdn/shop/files/download_3b901c0a-476d-4cdf-bbd0-8869535e3b3d.jpg?v=1742364618&width=326)  
[https://popular.com.sg/cdn/shop/files/download\\_b1e7d127-9a47-4e31-9e76-b629779571be.jpg?v=1744111232&width=325](https://popular.com.sg/cdn/shop/files/download_b1e7d127-9a47-4e31-9e76-b629779571be.jpg?v=1744111232&width=325)  
[https://popular.com.sg/cdn/shop/files/download\\_dc0e8fb8-e289-49b1-8e5b-afafeac8e0db.jpg?v=1742359566&width=327](https://popular.com.sg/cdn/shop/files/download_dc0e8fb8-e289-49b1-8e5b-afafeac8e0db.jpg?v=1742359566&width=327)  
[https://popular.com.sg/cdn/shop/files/download\\_028e1768-9fd7-4db8-928a-c9e4361d456b.jpg?v=1744110392&width=302](https://popular.com.sg/cdn/shop/files/download_028e1768-9fd7-4db8-928a-c9e4361d456b.jpg?v=1744110392&width=302)  
[https://popular.com.sg/cdn/shop/files/download\\_71a979f1-edb9-48b1-a014-d8a4553de527.jpg?v=1743572887&width=326](https://popular.com.sg/cdn/shop/files/download_71a979f1-edb9-48b1-a014-d8a4553de527.jpg?v=1743572887&width=326)  
[https://popular.com.sg/cdn/shop/files/download\\_55a5770c-7910-47b6-855f-0fa53ea25846.jpg?v=1742364455&width=825](https://popular.com.sg/cdn/shop/files/download_55a5770c-7910-47b6-855f-0fa53ea25846.jpg?v=1742364455&width=825)  
[https://popular.com.sg/cdn/shop/files/download\\_3c75b68b-5305-4feb-b00d-e505efd77f0f.jpg?v=1742364599&width=330](https://popular.com.sg/cdn/shop/files/download_3c75b68b-5305-4feb-b00d-e505efd77f0f.jpg?v=1742364599&width=330)  
[https://popular.com.sg/cdn/shop/files/download\\_bad97c48-c7ac-4d61-8d2e-72648de14f19.jpg?v=1743754425&width=333](https://popular.com.sg/cdn/shop/files/download_bad97c48-c7ac-4d61-8d2e-72648de14f19.jpg?v=1743754425&width=333)  
[https://popular.com.sg/cdn/shop/files/download\\_afad4b42-0c16-46cb-be09-cffd1e323dc1.jpg?v=1744110158&width=326](https://popular.com.sg/cdn/shop/files/download_afad4b42-0c16-46cb-be09-cffd1e323dc1.jpg?v=1744110158&width=326)  
[https://popular.com.sg/cdn/shop/files/download\\_7369f27f-15fd-4f94-baf8-d05d253aa012.jpg?v=1744110188&width=327](https://popular.com.sg/cdn/shop/files/download_7369f27f-15fd-4f94-baf8-d05d253aa012.jpg?v=1744110188&width=327)

## Download Images to your Device

Now that we have a list of all the URLs for every image, we will go ahead and download them to our device with one simple tool.

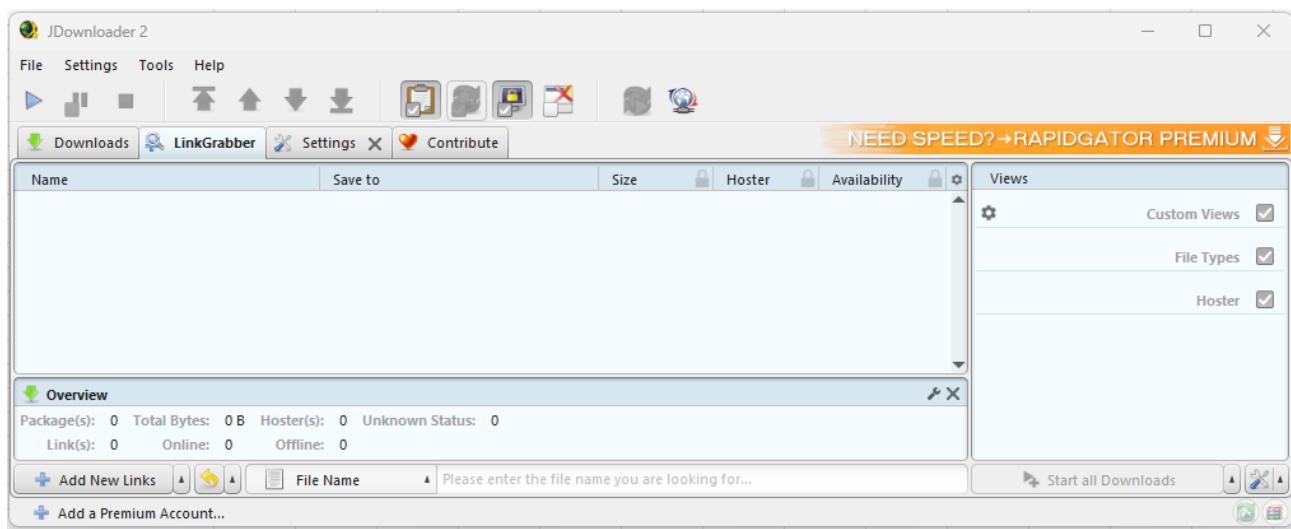
You can use a file downloader like [Jdownloader](https://jdownloader.org/). Choose the operating system from the webpage and download the software




The screenshot shows the JDownloader website interface. At the top, there's a navigation bar with links: Home, News, Download, Changelog, Wiki/FAQ, Help, Support, Board. Below this, there's a section for "Boomerang for Gmail" and "The Ultimate Email Lifhack" with a "DOWNLOAD" button. The main content area is titled "Download" and describes JDownloader as a free, open-source download management tool. It lists the latest revision as 50639, dated Thu Apr 17 16:30:50 CEST 2025. Under the heading "Choose your operating system:", the "Windows" option is selected, showing a link to the "Download Installer (x64, Java 21)" with a size of 75.185.752 Bytes and a SHA256 hash.

Once installed on your PC, open the JDownloader. This will open up software, then click on downloads on the top menu then click on the "Add new links" at the bottom to enter the URLs we just extracted.

## Choose "LinkGrabber"



Click on the "Add New Links" to paste the links from Excel.

When you click on the run  icon at the top menu, all images will automatically be downloaded to your device. This might take a couple of seconds depending on how many images you are downloading.

## Task 2: Scrape a Password Protected Website

Now, we will complete a series of exercises that use some of the advanced commands we've reviewed earlier.

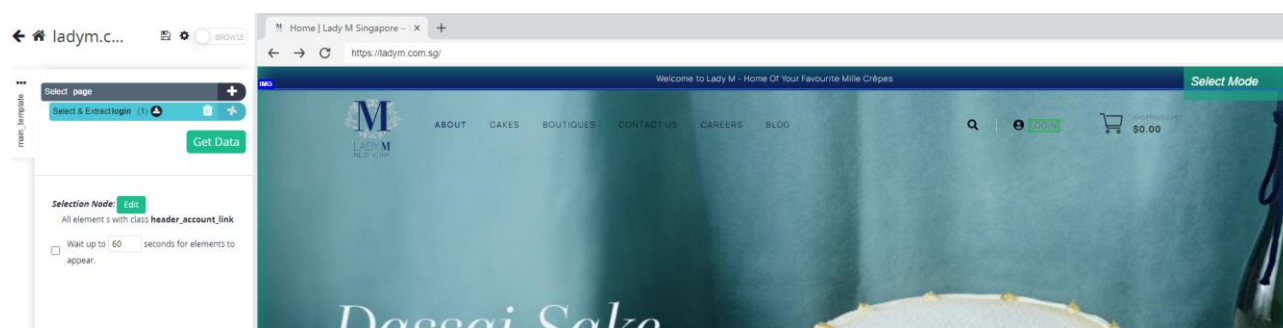
First, we will go over how to setup ParseHub to scrape data past a login screen or from a password-protected site.

### Scraping a Website with a Login Screen

Open ParseHub and enter the URL of the site you'd like to scrape, in this case, we will use the Lady M homepage.

<https://ladym.com.sg/> (Register a account with this website first)

ParseHub will now render the page inside the app. Start by clicking in the "Log In" button to select it. In the left sidebar, rename your selection to login.



Click on the PLUS(+) sign next to your login selection and choose the Click command.

A pop-up will appear asking you if this is a "Next Page" button. Click on "No", name your template to login\_page and click "Create New Template".

## Click setup



Is **LOG IN** a next page button?

Yes

No

This click takes you to a new page. Choose a template to use on that page ([click command help](#)).

☒ Create New Template login\_page

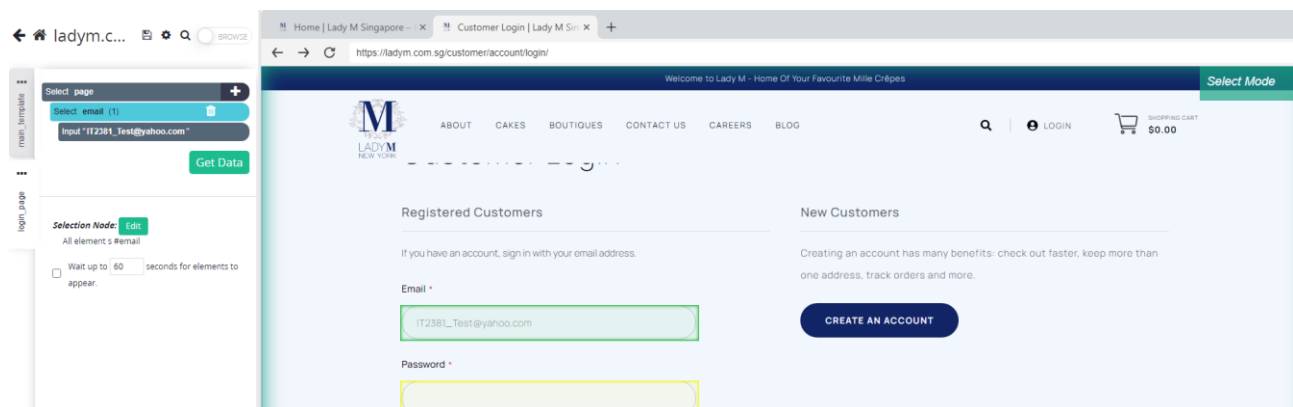
☐ Go to Template from Another Project

Advanced ▾

Create New Template

A new browser tab and new scraping template will open in ParseHub.

Start by clicking on the email field. ParseHub will automatically ask you for the text to enter in this field, enter your account email and rename the selection to email. This will automatically create an input command.

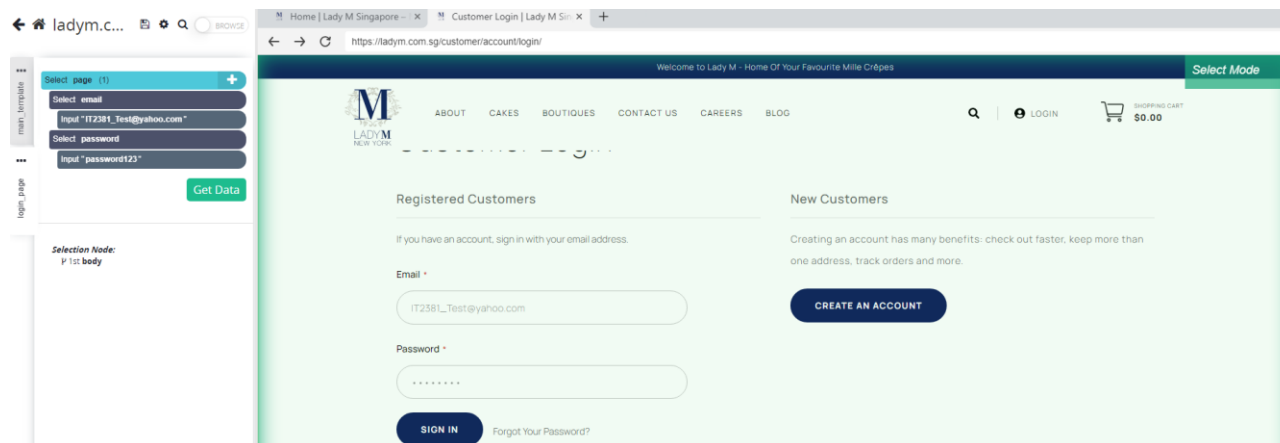


**Pro-Tip:** We recommend you use a dummy account for this step in order to keep your personal information safe.

Click on the PLUS(+) sign next to your page selection and use the Select command.

Now, select the password field and just like in the previous step, enter your password details and rename your selection to password.





Click on the PLUS(+) sign next to your page selection and choose the Select command.

With the select command, click on the blue "SIG IN" button and rename your selection to log\_in.

Click on the PLUS(+) sign next to your log\_in selection and choose the Click command.

A pop-up will appear asking you if this is a "Next Page" button. Click on "No" and create a new template. In this case, we will name it account\_page.

Click setup ✕

---

Is **SIGN IN** a next page button?

☐ Yes ☒ No

This click takes you to a new page. Choose a template to use on that page ([click command help](#)).

☒ Create New Template

☐ Go to Existing Template

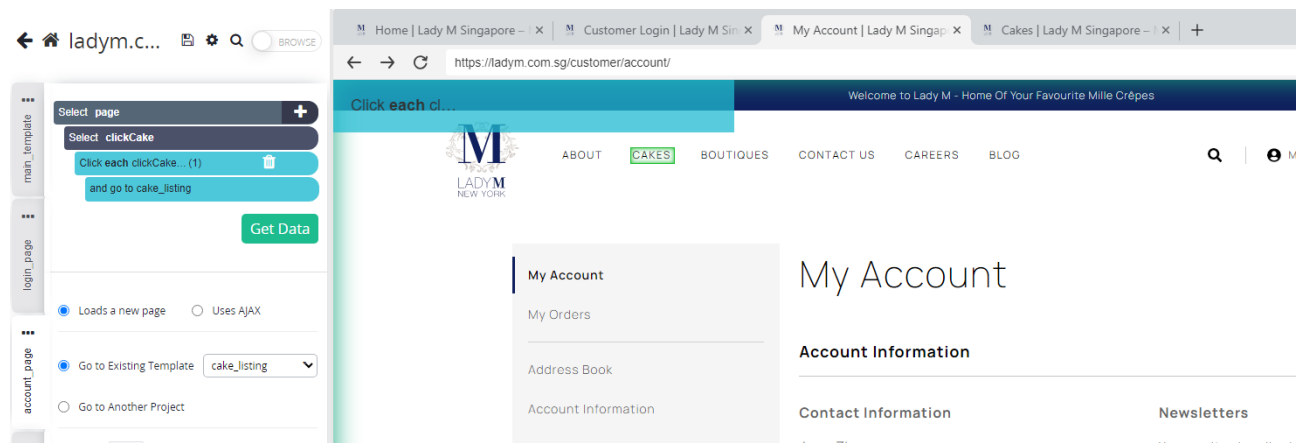
☐ Go to Template from Another Project

[Advanced](#) ▼

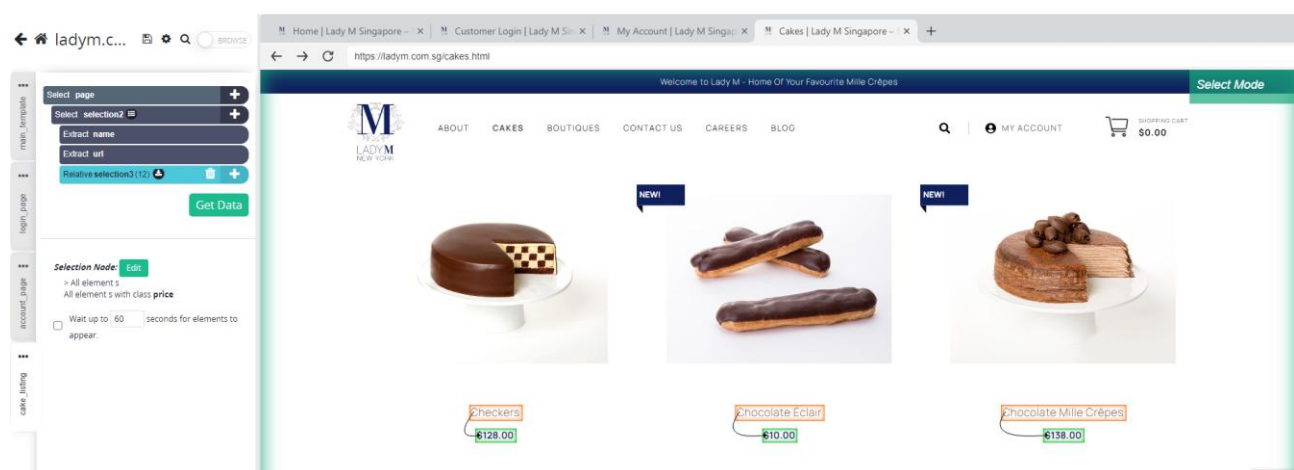
---

[Create New Template](#)

Click the "CAKES" in the menu and direct parsehub to cake listing page



Scrape the first page of cake listing



Before test the project, you need to logout Lady M website. Change the canvas to browse mode and click on "MY ACCOUNT" to sign out.

Now you can use ParseHub to login password protected website and scrape their data.

### Task 3: Scrape Through a List of Keywords

Next, we will create a new project and learn how to use more advanced ParseHub commands. In this exercise, we will go over how to search and scrape data from a list of keywords.

We will setup ParseHub to search through a list of keywords and extract product results.

First, let's create our list of keywords. Open your spreadsheet software of choice and create a CSV file with all keywords in one column. Make sure you have a header cell. In this case, it will be the word "keywords".

keywords

walnut

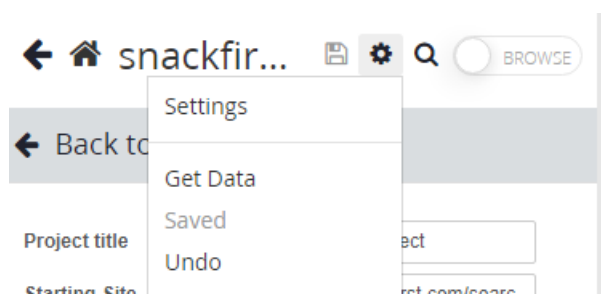
fruit

cereal

Our list contains the keywords: walnut, fruit and cereal.

Click on "New Project" and enter the URL of the website you will be scraping from. In this case, we will scrape data from [https://www.snackfirst.com/search?options%5Bunavailable\\_products%5D=last&options%5Bprefix%5D=last&options%5Bfields%5D=title%2Cvendor%2Cproduct\\_type%2Cvariants.title&q=cereal](https://www.snackfirst.com/search?options%5Bunavailable_products%5D=last&options%5Bprefix%5D=last&options%5Bfields%5D=title%2Cvendor%2Cproduct_type%2Cvariants.title&q=cereal). The page will then render inside the ParseHub and allow you to start extracting data.

Now, we need to give ParseHub our list of keywords we will be searching through to extract data. To do this, click on the settings icon at the top left and click on "settings".



Under the "Starting Value" section you can enter your list of keywords either as a CSV file or in JSON format right in the text box below it.

Project title

Starting Site

Starting Template

Web Hook

Project Token

API Key

Max Workers

Max Pages

Enable Email Notifications

Load JavaScript & Images ☒

Rotate IP Addresses ☒

Disable HTTPS verification ☐

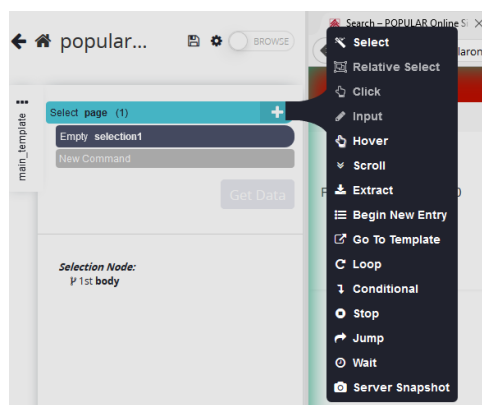
Starting Value

```

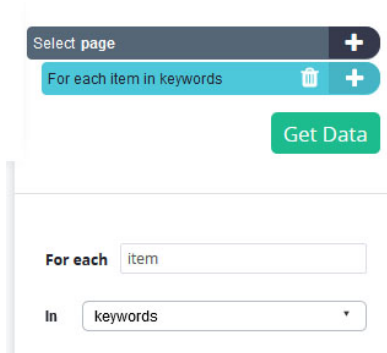
{"keywords":["walnut","fruit","cereal"]}

```

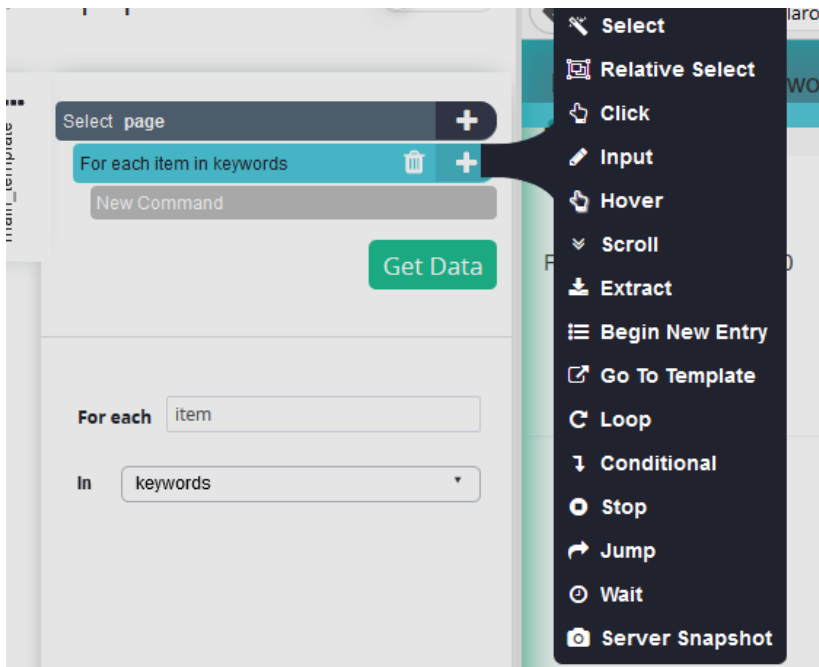
Once you've submitted your list of keywords, click on "Back to Commands" to go back to your project. Click on the PLUS (+) sign next to your "page" selection, click on Advanced and click on the "Loop" command.



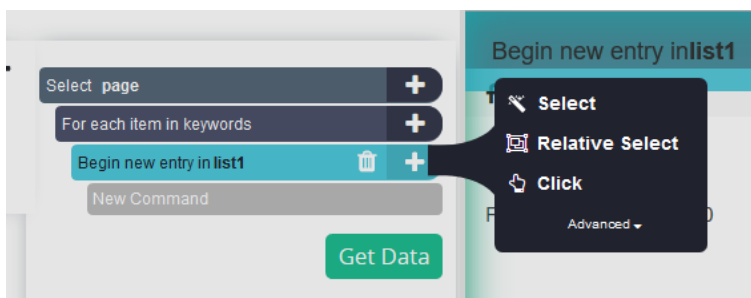
By default, your list of keywords will be selected as the list of items to loop through. If not, make sure to select "keywords" from the dropdown.



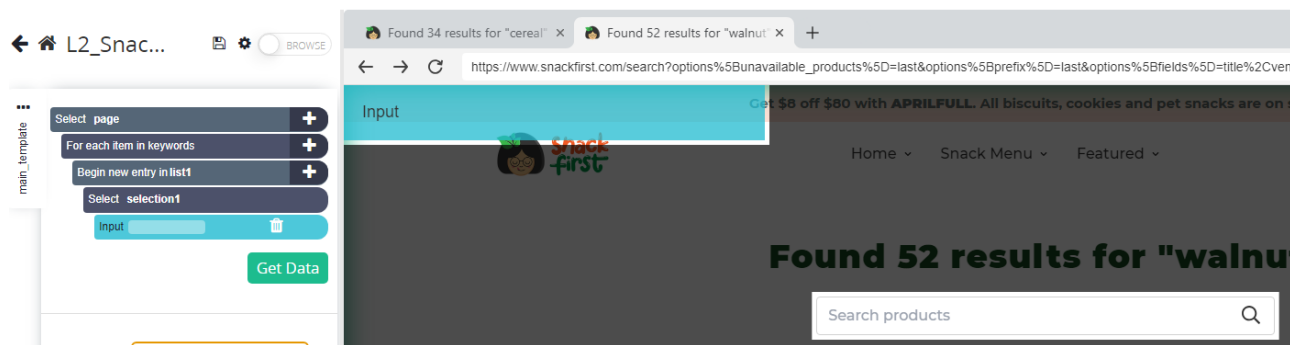
Click on the PLUS (+) sign next to your “For each item in keywords” selection and choose the “Begin New Entry” command. This command will be named “list1” by default.



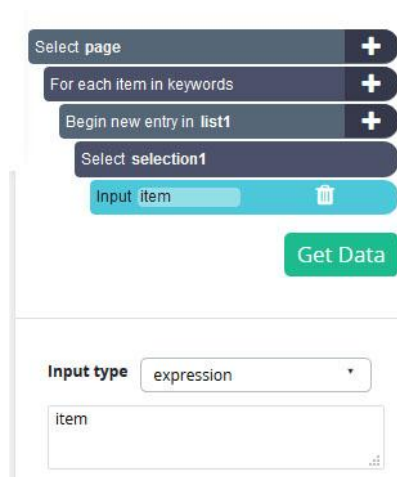
Click on the PLUS (+) sign next to your “list1” tool and choose the “Select” command.



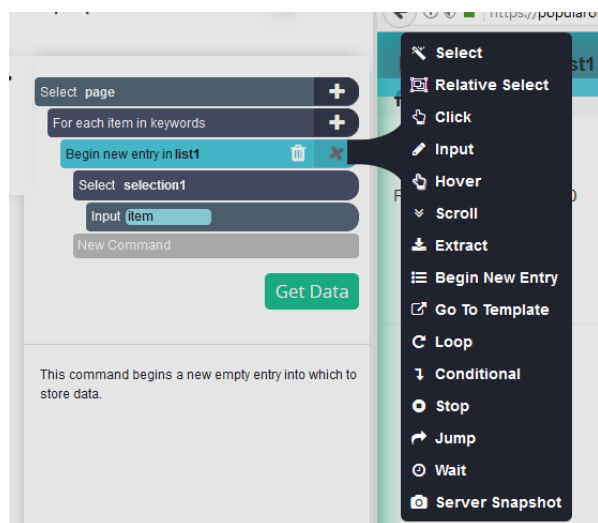
With the select command, click directly on the Snackfirst search bar to select it.



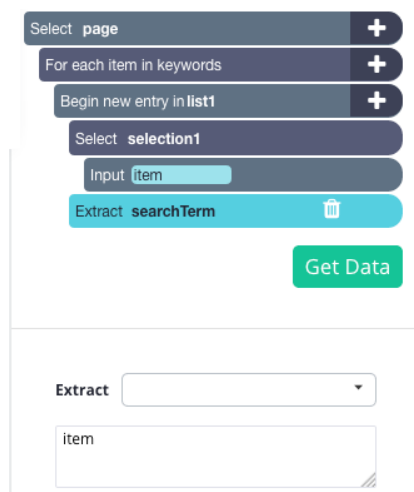
This will create an input command, under it, choose “expression” from the dropdown and enter the word “item” on the text box.



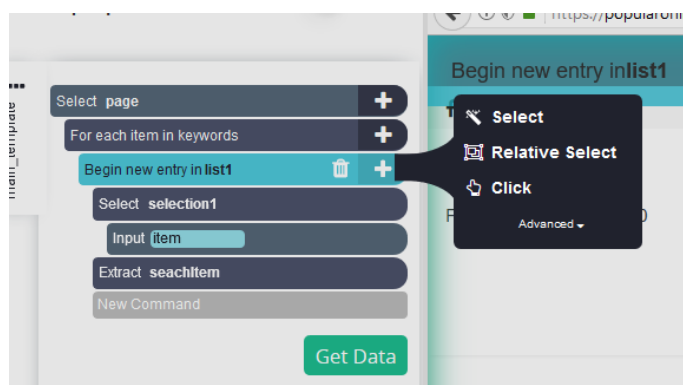
Now we will make it so ParseHub adds the keyword for each result next to it in our final export. To do this, click on the PLUS (+) sign next to the “list1” command and choose the “extract” command.



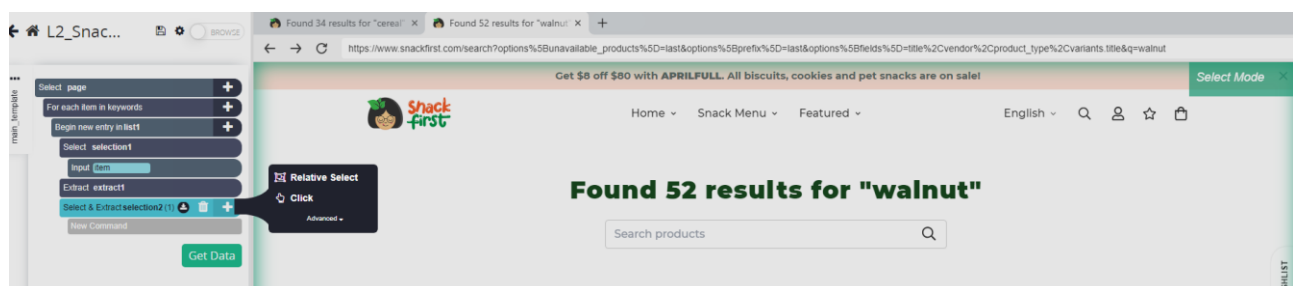
Under the extract command, enter the word “item” into the first text box.



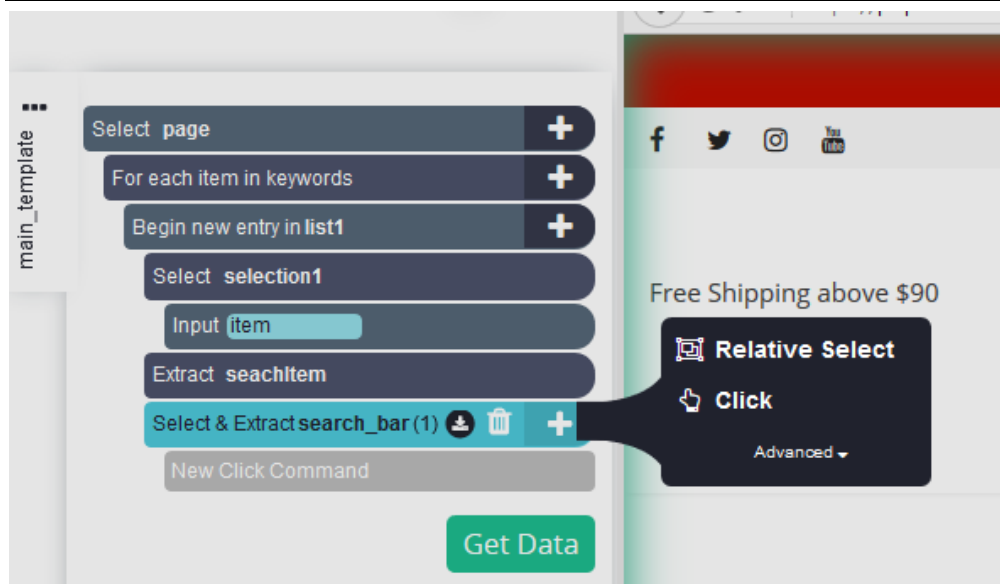
Now, let's tell ParseHub to perform the search for the keywords in the list. Click on the PLUS (+) sign next to your "list1" selection and choose the select command.



Click on the Search Button to select it and rename it to "search\_bar"




Click on the PLUS (+) sign next to your "search\_bar" selection and choose the "Click" command



A pop-up will appear asking you if this is a “next page” button. Click on “No” and rename your new template to “search\_results”

Click setup
×

Is  a next page button?

Yes
No

This click takes you to a new page. Choose a template to use on that page ([click command help](#)).

☒ Create New Template

☐ Go to Template from Another Project

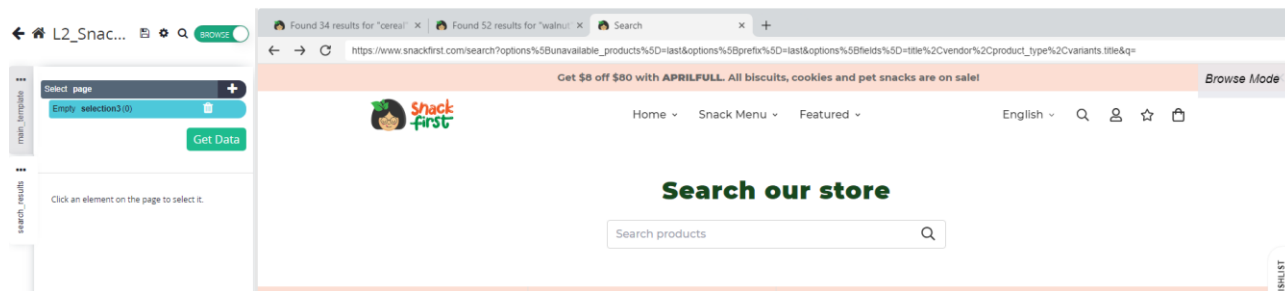
Advanced ▾

Create New Template

Now, let’s navigate to the search results page of the first keyword on the list and extract some data.

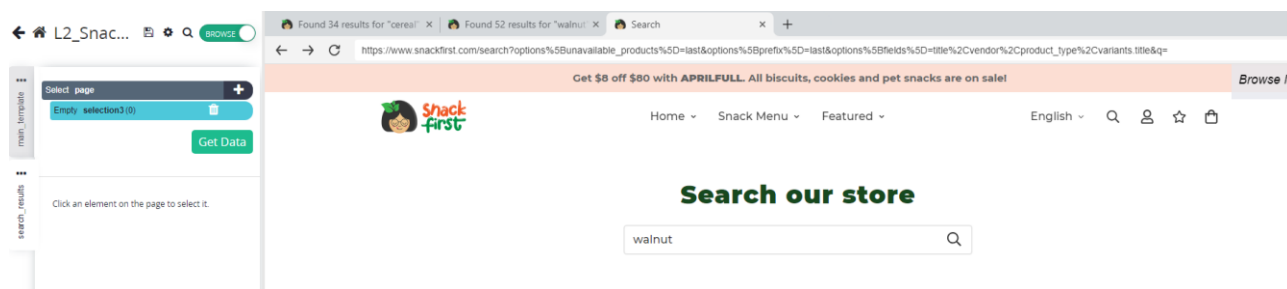
Start by switching over to browse mode on the top left and search for the first keyword on the list.



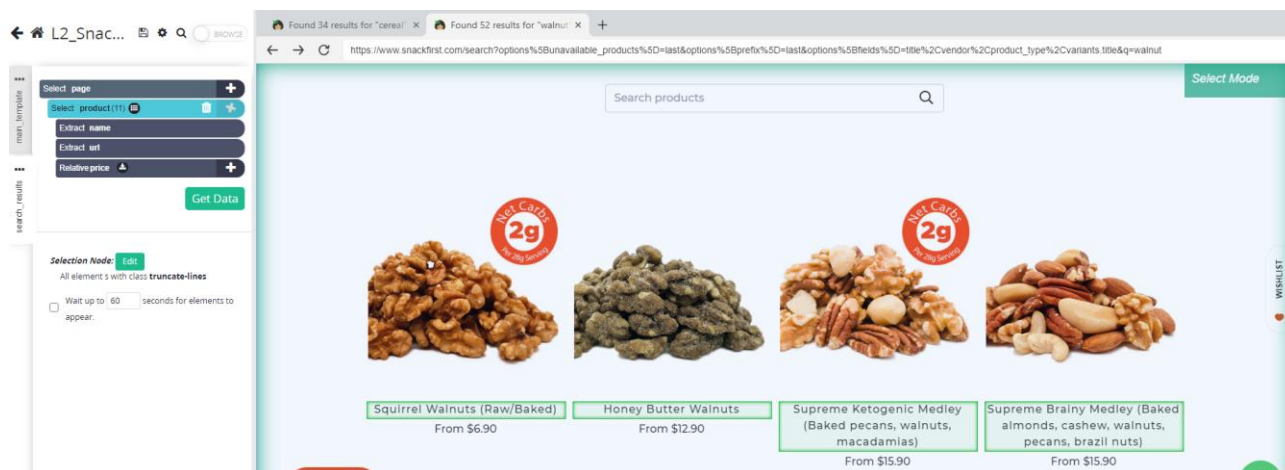


Once the page renders, make sure you are still working on your new “search\_results” template by selecting it with the tabs on the left.

Click on the search icon in the browse mode



Now, turn off Browse Mode and click on the name of the first result on the page to select it. It will be highlighted in green to indicate that it has been selected.



The rest of the products on the page will be highlighted in yellow. Click on the second one on the list to select them all.

ParseHub is now extracting the product name and price for each product on the first page of results for each keyword.

It is now time to run your project and export it as an Excel file.

To do this, click on the green “Get Data” button on the left sidebar. And run your scrape.

Your results file will show both the data you selected plus the keyword associated with it.

A	B	C
list1_searchItem	list1_product_name	list1_product_price
walnut	Squirrel Walnuts (Raw/Baked)	From \$6.90
walnut	Honey Butter Walnuts	From \$12.90
walnut	Supreme Ketogenic Medley (Baked pecans, walnuts, macadamias)	From \$15.90
walnut	Supreme Brainy Medley (Baked almonds, cashew, walnuts, pecans, braz	From \$15.90
walnut	Supreme Hearty Medley (Baked almonds, cashew, walnuts, cranberries, i	From \$15.90
walnut	Honey Sesame Walnut-Cashew Jumble	From \$14.90
walnut	Antioxidants Medley (Dates, baked walnuts, almonds, hazelnuts, raisins	From \$10.90
walnut	Brainy Pecans (Raw/Baked)	From \$14.90
walnut	Honey Red Dates (Deglet Nour)	From \$6.90

## Task 4: Scrape a list of URLs

For this exercise, we will use some of the advanced commands we used last time. But this time, we will use them to load a list of URLs to scrape.

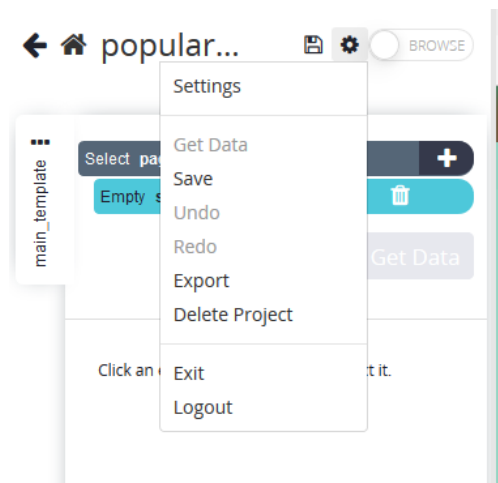
You might have a list of URLs rather than keywords to scrape data from. We will setup ParseHub to load each URL in order and scrape the data you need from each page.

For this example, we will extract data from 4 random Popular product listings. We will set up our scraper to open the URLs for each product page and extract some data we have selected.

Start by opening ParseHub. Click on “new project” and enter a basic URL. We will start with the Amazon.ca homepage.

ParseHub will now render the webpage inside the app.

Now, let’s give ParseHub the list of URLs we will be scraping data from. To do this, start by clicking on the Settings icon at the top right of the screen.



Here in the settings menu, we can submit our list of URLs under the “Starting Value” section.

We can do that either by clicking on the green “Import from CSV/JSON” button or by copy/pasting our list of URLs in JSON format right into the text box.

If you are submitting a CSV file with URLs, make sure the heading reads “urls” and the URLs are copy-pasted below it. Just like in the image below.

urls									
<a href="https://popular.com.sg/collections/gadgets-it-computer/products/elecom-m-ir07drs-wireless-ir-mouse-pink">https://popular.com.sg/collections/gadgets-it-computer/products/elecom-m-ir07drs-wireless-ir-mouse-pink</a>									
<a href="https://popular.com.sg/collections/gadgets-it-computer/products/hp-wireless-mouse-200">https://popular.com.sg/collections/gadgets-it-computer/products/hp-wireless-mouse-200</a>									
<a href="https://popular.com.sg/collections/gadgets-it-computer/products/logitech-desk-mat-rose">https://popular.com.sg/collections/gadgets-it-computer/products/logitech-desk-mat-rose</a>									

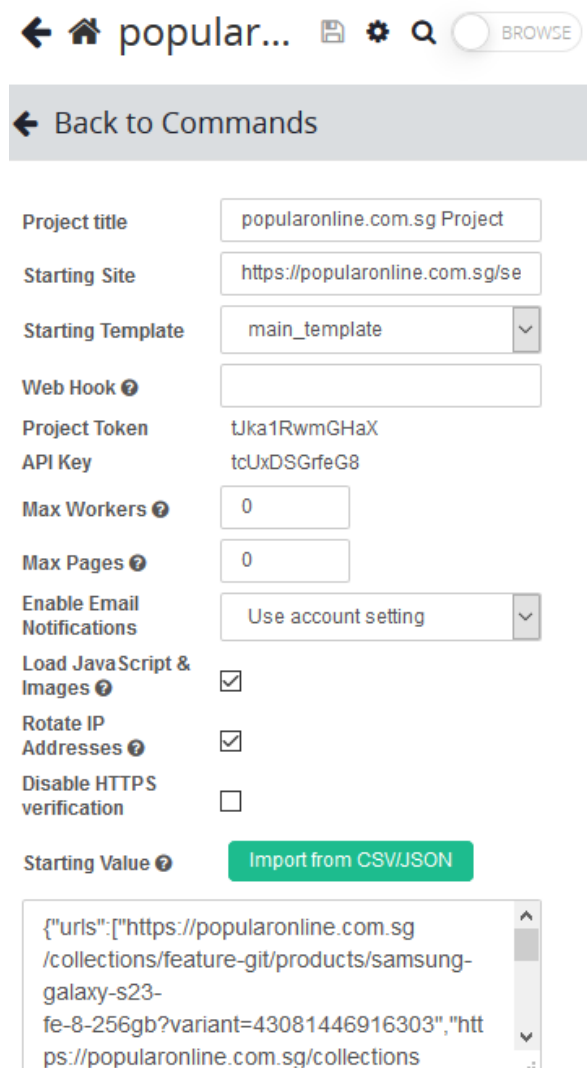
In this case, we will just copy-paste our list of URLs in a JSON format in the text box.

Here is the code we pasted in, in case you'd just want to use it and replace it with your own URLs.

```
{"urls":["https://popular.com.sg/collections/gadgets-it-computer/products/elecom-m-ir07drs-wireless-ir-mouse-pink","https://popular.com.sg/collections/gadgets-it-computer/products/hp-wireless-mouse-200","https://popular.com.sg/collections/gadgets-it-computer/products/logitech-desk-mat-rose"]}]}
```

**Notes:** check the links when coping from PDF file. Special characters needs to be removed and replaced with "-"

The starting site is : [Search – POPULAR Online Singapore](#)



The screenshot shows a web browser interface with a settings panel open. The browser address bar shows 'popular...' and a 'BROWSE' button. The settings panel has a 'Back to Commands' button at the top. The settings include:

- Project title:** popularonline.com.sg Project
- Starting Site:** https://popularonline.com.sg/se
- Starting Template:** main\_template (dropdown)
- Web Hook:** (empty field)
- Project Token:** tJka1RwmGHaX
- API Key:** tcUxDSGrfeG8
- Max Workers:** 0
- Max Pages:** 0
- Enable Email Notifications:** Use account setting (dropdown)
- Load JavaScript & Images:** ☒
- Rotate IP Addresses:** ☒
- Disable HTTPS verification:** ☐
- Starting Value:** Import from CSV/JSON (button)

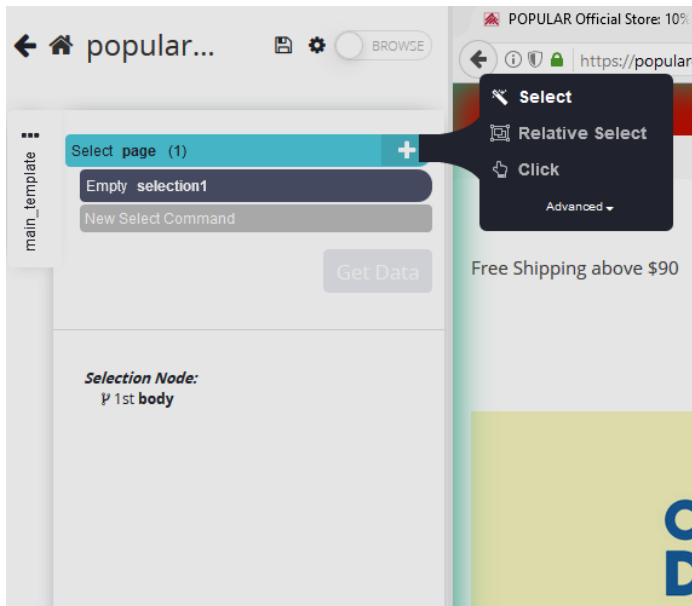
The 'Starting Value' field contains the following JSON text:

```
{"urls":["https://popularonline.com.sg/collections/feature-git/products/samsung-galaxy-s23-fe-8-256gb?variant=43081446916303","https://popularonline.com.sg/collections
```

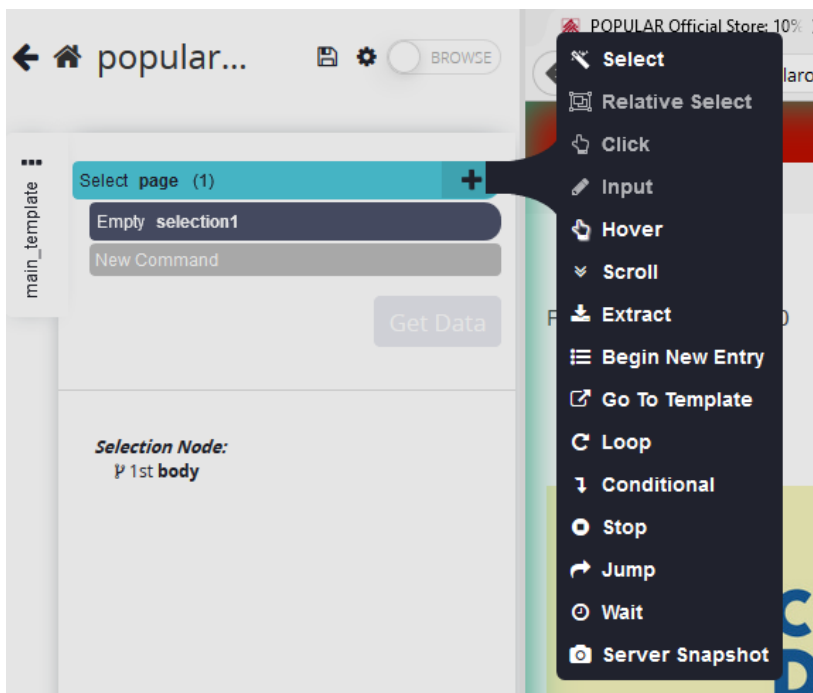
Close the settings panel by clicking on "Back to Commands" on the top right.

Now, that we have submitted our list of URLs, it is time to start setting up ParseHub to navigate through all the URLs and extract data.

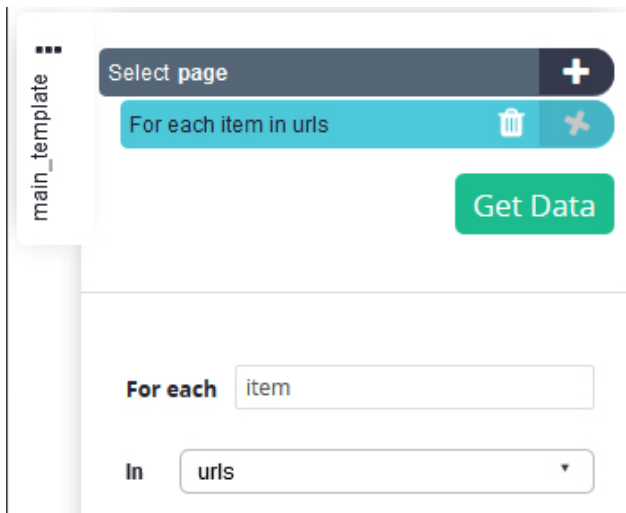
Back on the Commands screen, click on the PLUS (+) sign next to your “select page” command.



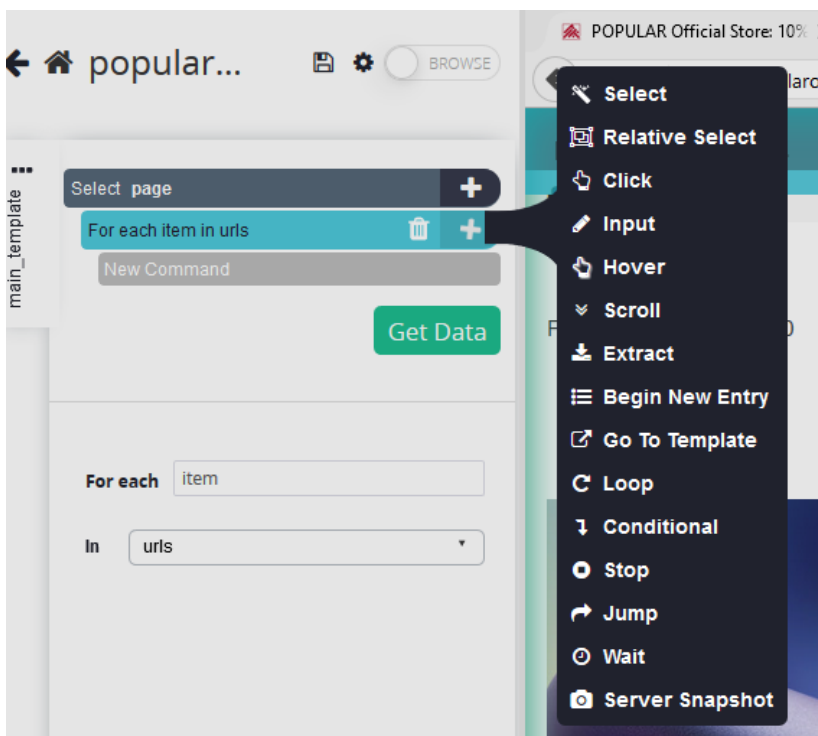
Click on “Advanced” and select the “Loop” command.



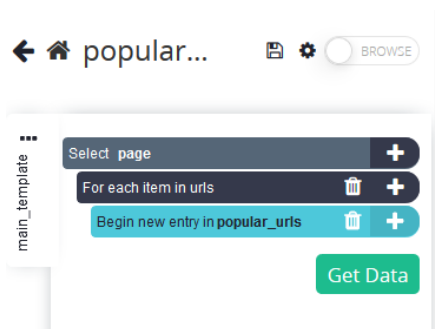
By default, the Loop command should be Looping though every item in “urls”. If it isn’t, use the dropdown to select “urls”



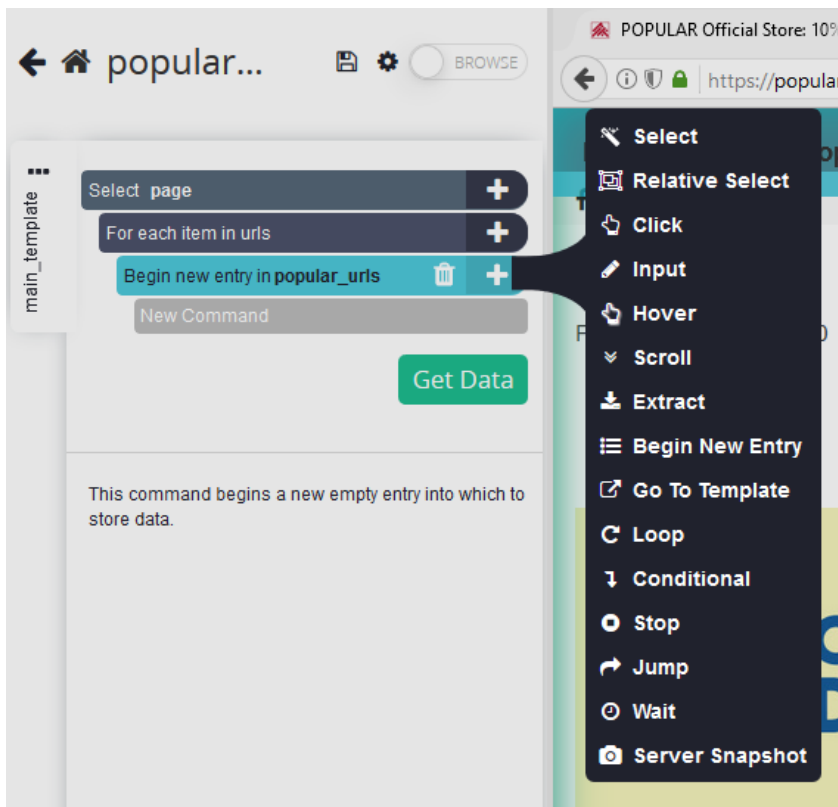
Click on the PLUS (+) sign next to your new Loop command and choose the "Begin New Entry" command under "Advanced".



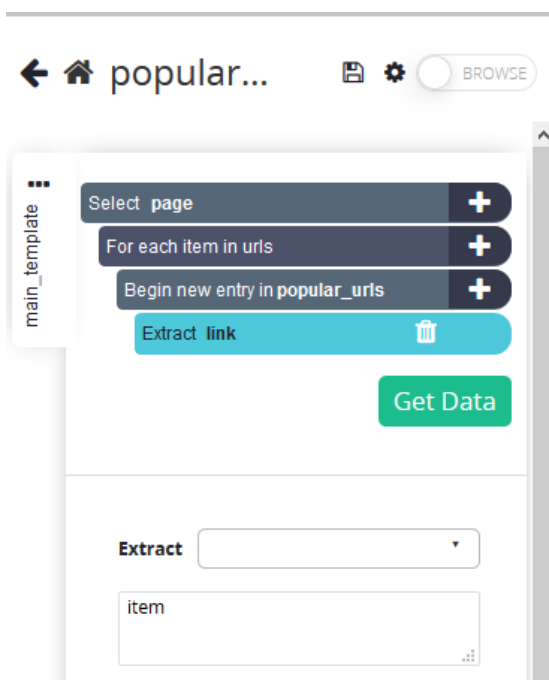
Make sure that this command has a different name than "urls". We will rename this one to "popular\_urls"



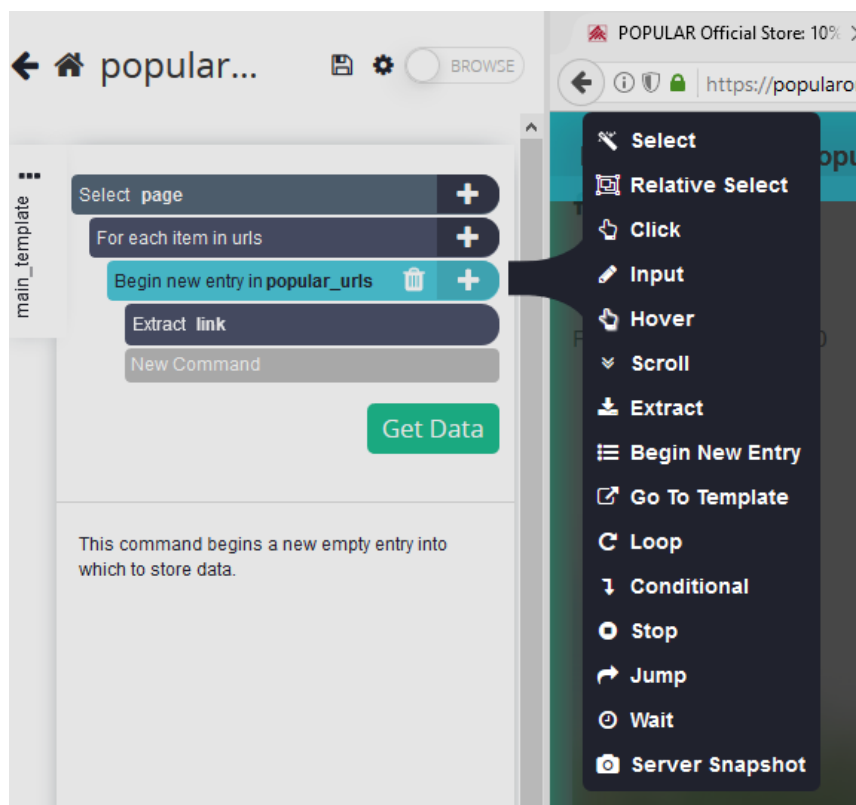
Next, we want our final scrape to keep the URLs for each product we are scraping data from. To do this, click on the PLUS (+) sign next to your “begin new entry command” and choose the “Extract” command under “Advanced”.



Rename this command to “link” and replace “\$location.href” with “item” in the text box below.



Now, we will tell ParseHub to start loading the URLs we've selected. Start by clicking on the PLUS (+) sign next to your "Begin new entry" command and choose the "Go To Template" command under "Advanced".



A pop up will appear with a few available selections. In the "Go to URL" field we will write in "link" without quotations. On the "Create New Template" field, we will name our new template "product\_page". Once done, click on the green "Create New Template" button.

Go To Template
×

Make ParseHub start executing a template. You can either execute the template on the same page or go to a URL. URLs are expressions, so you can combine them out of multiple parts.

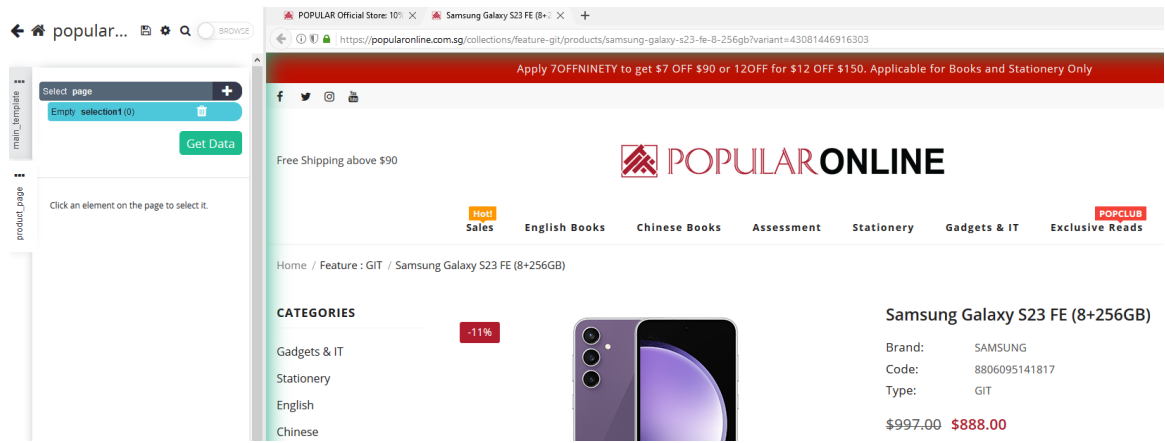
☐ Stay on the Same Page
   
☒ Go to URL: 
  
☐ Go to Existing Template 
  
☐ Go to Template from Another Project
   
☒ Create New Template

Advanced ▾

Create New Template

The page for the first URL on the list will now render inside the app. Make sure to select the new "product\_page" template tab on the left sidebar.





A “Select” command will be created by default. We will use this command to extract data from this page. Start by clicking on the product name on the page to extract it. The product name on the page will be highlighted in green to indicate that it has been selected. On the left sidebar, rename your selection to “product\_name”.

Now, click on the PLUS (+) sign next to your “select page” command to add a new “Select” command and extract more data. We will click on the product price to extract it. Rename this new command accordingly. Repeat these steps if you want to add more Select commands and extract more data. Our template ended up looking like this:

You are now ready to run your scrape job and download the data you’ve selected as a CSV file.

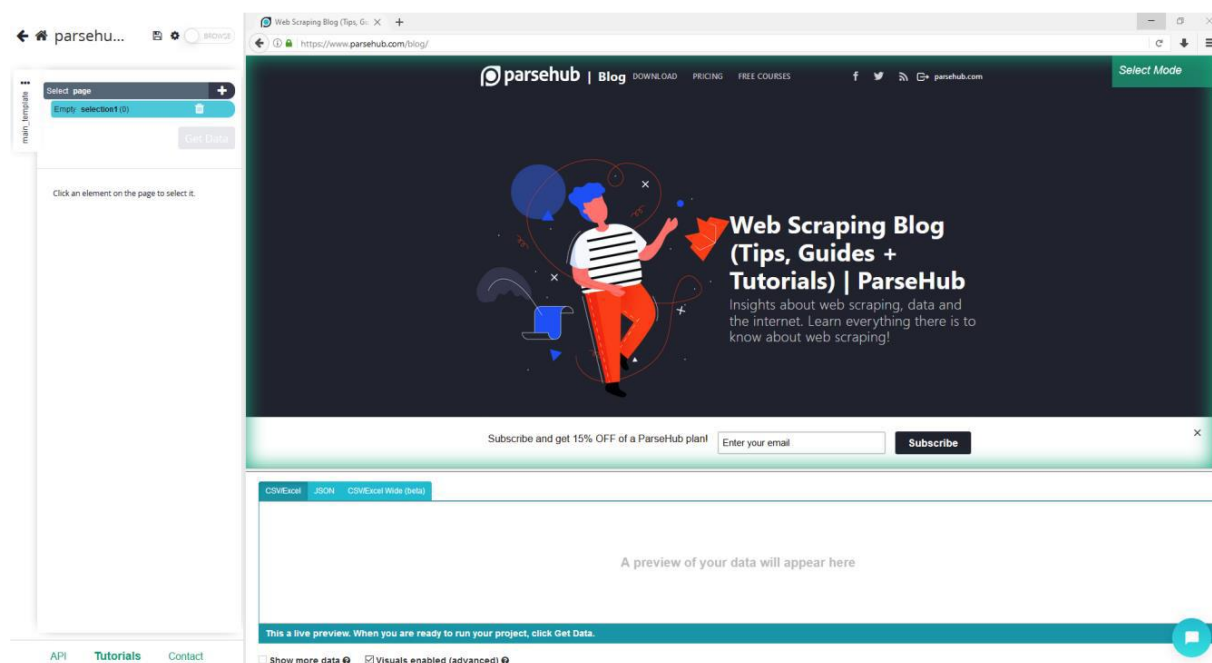
popular_urls_link	popular_urls_name	popular_urls_price		
https://popular.com.sg/collections/gadgets-it-computer/products/elecom-m-ir07drs-wireless-ir-mous	Elecom M-ir07drs Wireless I	\$17.90		
https://popular.com.sg/collections/gadgets-it-computer/products/hp-wireless-mouse-200	Hp Wireless Mouse 200	\$20.90		
https://popular.com.sg/collections/gadgets-it-computer/products/logitech-desk-mat-rose	Logitech Desk Mat - Rose	\$39.00		

## Task 5: Web Scraping using XPath

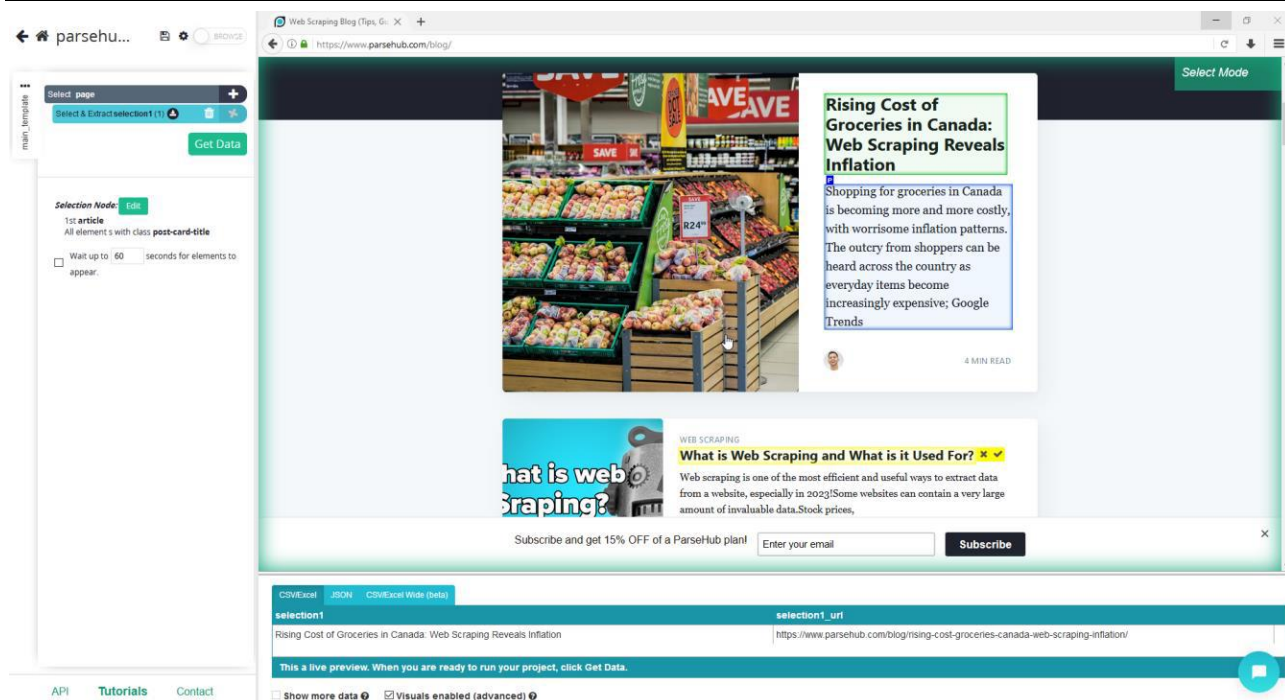
Next, we will use Xpath to scrape data from a website.

In this case, we will use Xpath to extract the meta Title and meta Description from numerous pages. Now it's time to get started.

Open ParseHub, click on "New Project" and enter the URL you will be scraping. In this case, we are scraping ParseHub Blog for the meta titles and descriptions of each blog post. Once submitted, the page will render inside the app.

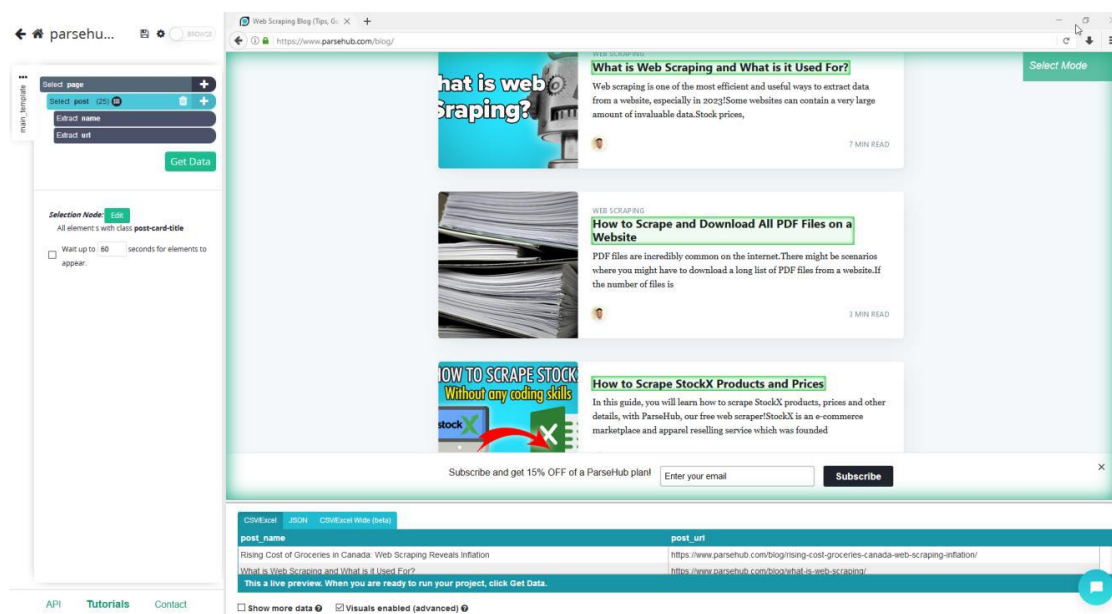


A "Select" command will be created by default, start by selecting the first blog post on the page by clicking on its headline. It will be highlighted in green to indicate that it's been selected. The rest of the posts on the page will be highlighted in yellow.



ParseHub interface showing the first selection step. The left sidebar shows "Select page" with "Select & Extract selection 1" chosen. The main area shows a preview of a blog post titled "Rising Cost of Groceries in Canada: Web Scraping Reveals Inflation". The bottom table shows the selected element with class "post-card-title" and the URL "https://www.parsehub.com/blog/rising-cost-groceries-canada-web-scraping-inflation/".

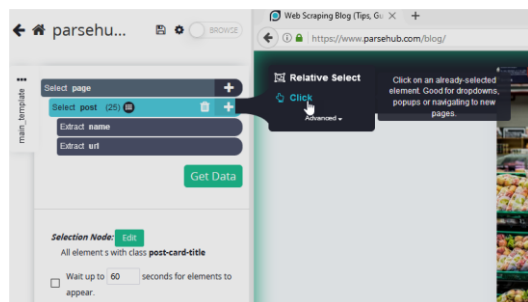
Now, click on the second post on the list to select all the posts. They will all now be highlighted in green. On the left sidebar, rename your selection to "post".



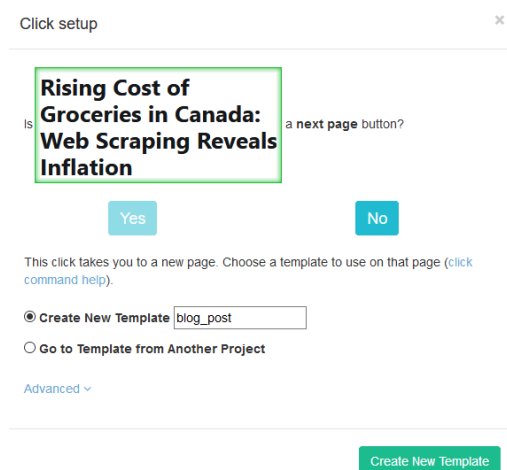
ParseHub interface showing the second selection step. The left sidebar shows "Select page" with "Select post (25)" chosen. The main area shows a preview of a blog post titled "What is Web Scraping and What is it Used For?". The bottom table shows the selected element with class "post-card-title" and the URL "https://www.parsehub.com/blog/what-is-web-scraping/".

ParseHub is now pulling the headline and URL for each post on the page. Let's now tell ParseHub to click on each post to also extract the Meta Title and Meta Description of each post. (Note: Remember, when it comes to SEO optimized content, the headline and meta title of each post are usually different).

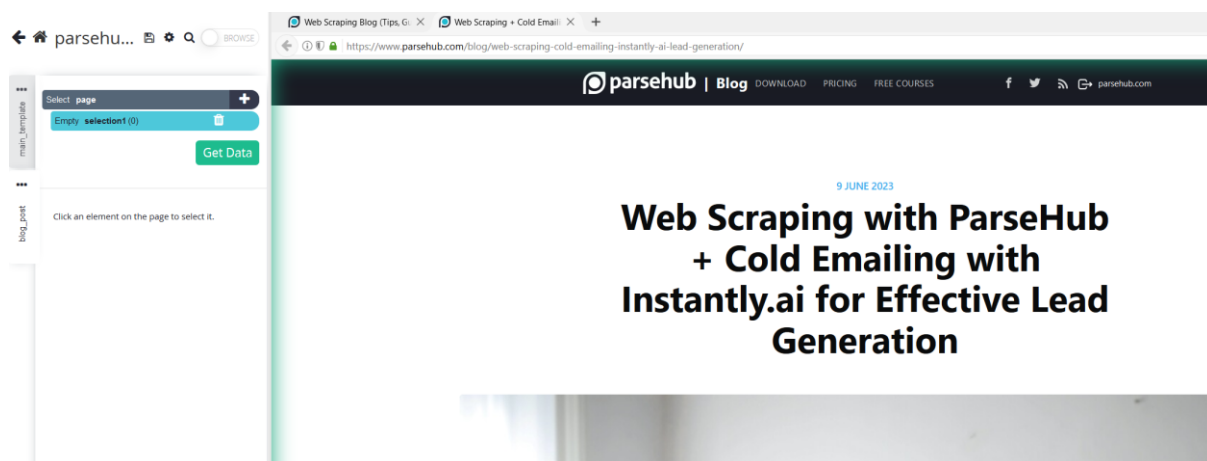
Click on the PLUS (+) sign next to your "post" selection and choose the "Click" command.



A pop-up will appear asking you if this is a "next page" link. Click on "No" and enter a name for your new template. In this case, we will call it "blog\_post". Lastly, click on the green "Create New Template" button.



The first blog post on the list will be rendered inside the app and a new select command will be created by default.



Extracting the Meta Title and Meta Description from each Page

We will now extract the meta title and meta description from each blog post we have selected.

First, we will inspect the code of the blog we are scraping to identify the two tags we want to extract.

In this case, we want to extract the <title> and <meta name="description"> tags of each page.

```
<!DOCTYPE html>
<html lang="en">
<head>

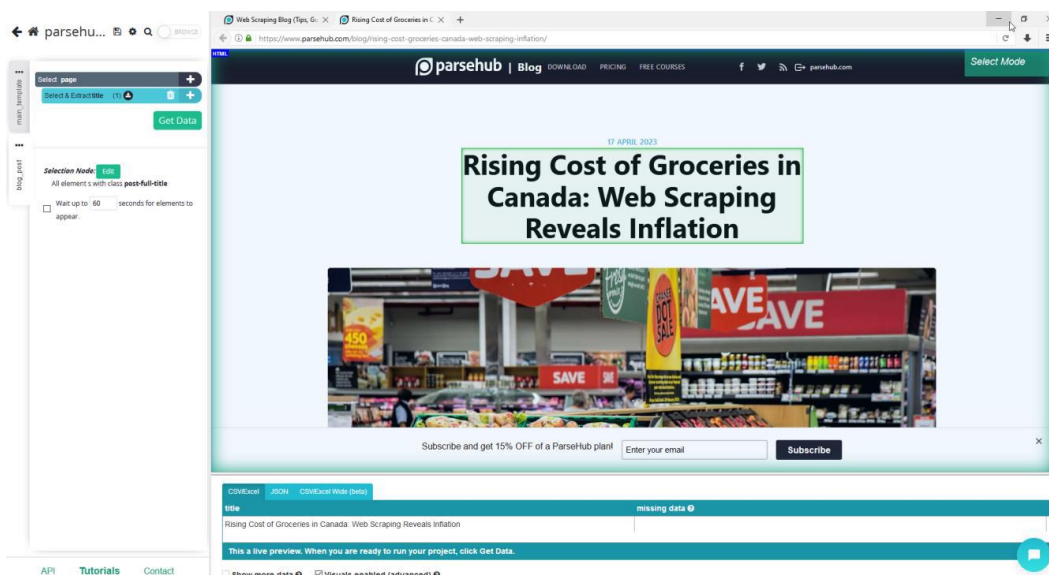
  <meta charset="utf-8" />
  <meta http-equiv="X-UA-Compatible" content="IE=edge" />

  <title>Rising Cost of Groceries in Canada: Web Scraping Reveals Inflation</title>
  <meta name="HandheldFriendly" content="true" />
  <meta name="viewport" content="width=device-width, initial-scale=1.0" />

  <link rel="stylesheet" type="text/css" href="/blog/assets/built/screen.css?v=103a430b2a" />

  <meta name="description" content="Learn how web scraping can be used to monitor grocery prices and inflation, as it has for Canadians over the last several years." />
  <link rel="shortcut icon" href="/blog/favicon.png" type="image/png" />
  <link rel="canonical" href="https://www.parsehub.com/blog/rising-cost-groceries-canada-web-scraping-inflation/" />
  <meta name="referrer" content="no-referrer-when-downgrade" />
  <link rel="amphtml" href="https://www.parsehub.com/blog/rising-cost-groceries-canada-web-scraping-inflation/amp/" />
```

Use the select command that was created by default on the previous step to select any element on the page. In this case, we will select the post's headline. It will be highlighted in green to indicate it has been selected. Rename your selection to "title".



The screenshot shows the ParseHub web scraper interface. The browser window displays the article "Rising Cost of Groceries in Canada: Web Scraping Reveals Inflation" from the ParseHub blog. The "Selection Node" panel on the left indicates that the selection is "All element's with class post-full-title". The "Data Table" at the bottom shows the extracted title: "Rising Cost of Groceries in Canada: Web Scraping Reveals Inflation".

Under your new selection, click on the "Edit" button next to "Selection Node" and choose the "use XPath selection" option.

Select page +

Select & Extract title (1) 🗑️ +

Get Data

---

**Selection Node:** Edit

All element s w

☐ Wait up to 6 to appear.

☐ Retry selection

- Use CSS Selection
- Use XPath Selection
- Advanced

A new text field for a selector will appear, here we will ask ParseHub to extract the

Select page +

Select title (1) 🗑️ +

Get Data

---

**XPath Selection Node:** Edit

**Selector**

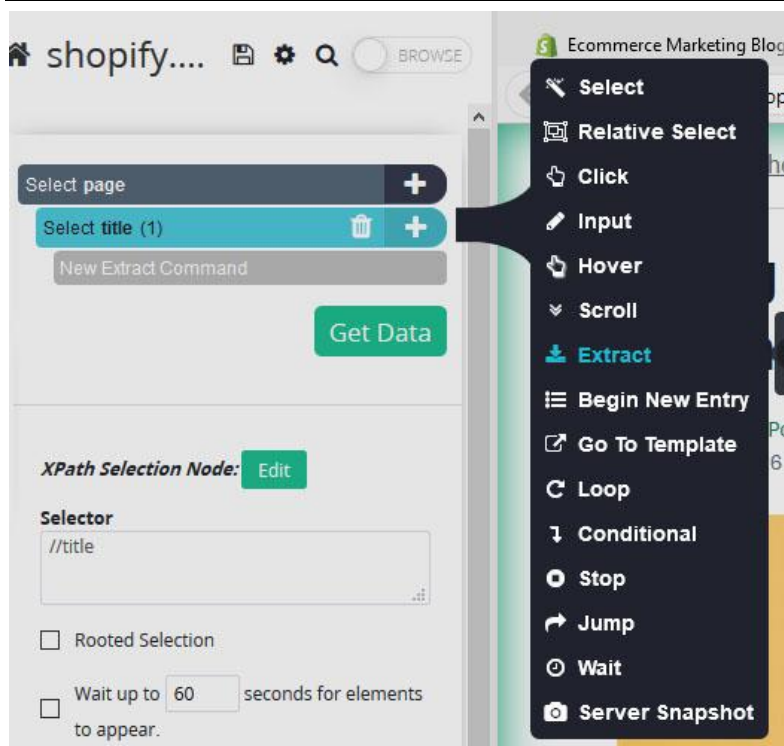
//title

☐ Rooted Selection

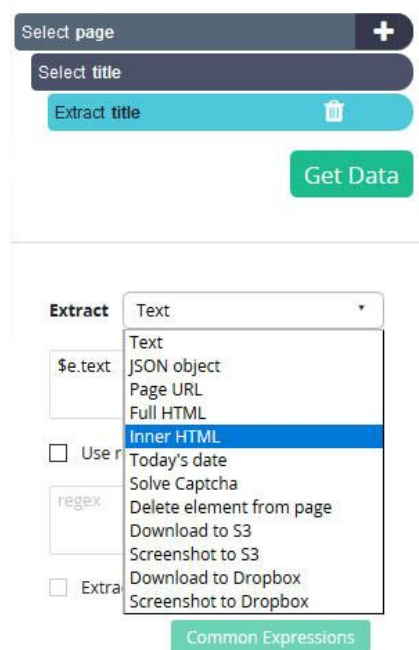
☐ Wait up to 60 seconds for elements to appear.

☐ Retry selection on missing element. ?

Now, click on the PLUS (+) sign next to your "title" selection, click on "Advanced" and choose the "Extract" command.

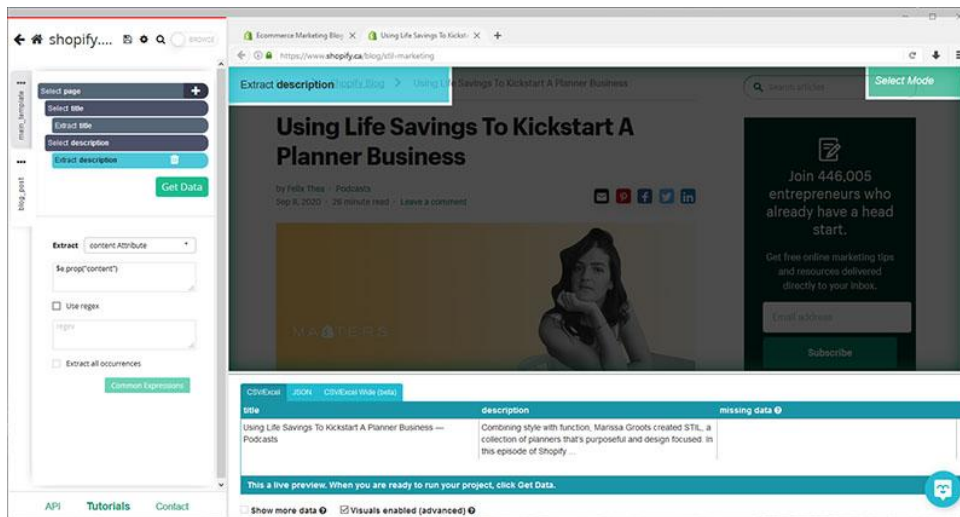


Under your new "Extract" command, use the dropdown and select the "Inner HTML" option. ParseHub will now pull the meta title of each page.



Now, click on the PLUS (+) sign next to your "page" command and add a new "select" command. With this select command repeat the steps 1-6 to extract the meta description for the page. For the Xpath command, use: **`//*[name='description']`** since there are more than one meta fields. On the extract command dropdown, choose "content Attribute".





Lastly, go back to your original template and add pagination to extract more pages worth of blog posts.

post_name	post_url	post_title	post_description
Rising Cost of Groceries in Canada: Web Scrapin	<a href="https://www.parsehub.com/blog/rising-cost-groceries-canada">https://www.parsehub.com/blog/rising-cost-groceries-canada</a>	Rising Cost of Groceries in Canada: Web Shopping for groceries in Canada is becoming more and more costly, with worrisome inflation patterns. The outcry from shoppers can be heard across the	
What is Web Scraping and What is it Used For?	<a href="https://www.parsehub.com/blog/what-is-web-scraping/">https://www.parsehub.com/blog/what-is-web-scraping/</a>	What is Web Scraping and What is it Used For? Web scraping is one of the most efficient and useful ways to extract data from a website, especially in 2023. Some websites can contain a very large amount of	
How to Scrape and Download All PDF Files on a Website	<a href="https://www.parsehub.com/blog/scrape-download-pdf-files/">https://www.parsehub.com/blog/scrape-download-pdf-files/</a>	How to Scrape and Download All PDF Files on a Website PDF files are incredibly common on the internet. There might be scenarios where you might have to download a long list of PDF files from a website. If that's	
How to Scrape StockX Products and Prices	<a href="https://www.parsehub.com/blog/scrape-stockx/">https://www.parsehub.com/blog/scrape-stockx/</a>	How to Scrape StockX Products and Prices In this guide, you will learn how to scrape StockX products, prices and other details, with ParseHub, our free web scraper ( <a href="https://www.parsehub.com/blog/scrape-stockx/">https://www.parsehub.com/blog/scrape-stockx/</a> )	
How to Scrape ThriftBooks	<a href="https://www.parsehub.com/blog/scrape-thriftbooks/">https://www.parsehub.com/blog/scrape-thriftbooks/</a>	How to Scrape ThriftBooks In this tutorial, we will show you how to scrape new or used books and other products from the ThriftBooks website. ThriftBooks was founded in 2003 in	
How to Scrape Calvin Klein Products	<a href="https://www.parsehub.com/blog/scrape-calvin-klein-products/">https://www.parsehub.com/blog/scrape-calvin-klein-products/</a>	How to Scrape Calvin Klein Products In this tutorial, we will show you how to scrape products from the Calvin Klein website, using ParseHub our free web scraper ( <a href="https://www.parsehub.com/blog/scrape-calvin-klein-products/">https://www.parsehub.com/blog/scrape-calvin-klein-products/</a> )	
How to Scrape Bath & Body Works	<a href="https://www.parsehub.com/blog/scrape-bath-body-works/">https://www.parsehub.com/blog/scrape-bath-body-works/</a>	How to Scrape Bath & Body Works In this guide, you will learn how to scrape products from the Bath & Body Works website, or any other e-commerce website that has a JSON API	
How to Scrape Company Details From Clutch.co	<a href="https://www.parsehub.com/blog/scrape-company-details-clutch.co">https://www.parsehub.com/blog/scrape-company-details-clutch.co</a>	How to Scrape Company Details From Clutch.co In this guide, we will show you how to scrape endless amounts of company and firm data from Clutch.co using our free web scraper ParseHub ( <a href="https://www.parsehub.com/blog/scrape-company-details-clutch.co">https://www.parsehub.com/blog/scrape-company-details-clutch.co</a> )	
How to Scrape Samsung Products	<a href="https://www.parsehub.com/blog/scrape-samsung-products/">https://www.parsehub.com/blog/scrape-samsung-products/</a>	How to Scrape Samsung Products In this tutorial, we will show you how to scrape any product on Samsung's website for free, including mobile phones, laptops, monitors and more with	
How to Scrape Flipkart Products	<a href="https://www.parsehub.com/blog/scrape-flipkart-products/">https://www.parsehub.com/blog/scrape-flipkart-products/</a>	How to Scrape Flipkart Products In this tutorial, you will learn how to scrape large amounts of Flipkart product data with our free web scraper ( <a href="https://www.parsehub.com/blog/scrape-flipkart-products/">https://www.parsehub.com/blog/scrape-flipkart-products/</a> )	
How to Scrape Trends: Free Google Trends Web	<a href="https://www.parsehub.com/blog/scrape-trends-google-trends/">https://www.parsehub.com/blog/scrape-trends-google-trends/</a>	How to Scrape Trends: Free Google Trends Web In this guide, you will learn how to scrape Google Trends' headlines and outlets for free, with ParseHub, in just a few easy steps. Google is the largest	
How to Scrape PubMed and Research Papers	<a href="https://www.parsehub.com/blog/scrape-pubmed-research-papers/">https://www.parsehub.com/blog/scrape-pubmed-research-papers/</a>	How to Scrape PubMed and Research Papers In this blog, we will show you how to scrape medical journals and research papers from PubMed with our free web scraper ( <a href="https://www.parsehub.com/blog/scrape-pubmed-research-papers/">https://www.parsehub.com/blog/scrape-pubmed-research-papers/</a> )	
How Much Are Your Pokemon Cards Worth in 2023?	<a href="https://www.parsehub.com/blog/pokemon-cards-ebay/">https://www.parsehub.com/blog/pokemon-cards-ebay/</a>	How Much Are Your Pokemon Cards Worth in 2023? We scraped over 1000 Pokemon cards that have recently been sold on eBay using our web scraping tool ( <a href="https://www.parsehub.com/blog/pokemon-cards-ebay/">https://www.parsehub.com/blog/pokemon-cards-ebay/</a> ) and we found out a few	
[2023 Tutorial] How to Scrape Amazon Product Data	<a href="https://www.parsehub.com/blog/scrape-amazon-product-data/">https://www.parsehub.com/blog/scrape-amazon-product-data/</a>	[2023 Tutorial] How to Scrape Amazon Product Data for In this 2023 updated scraping tutorial, we will show you how to web scrape Amazon products and their respective data, for free, without a line of code. A	
Job Board Web Scraping & How to Scrape Workopolis	<a href="https://www.parsehub.com/blog/how-to-scrape-workopolis/">https://www.parsehub.com/blog/how-to-scrape-workopolis/</a>	Job Board Web Scraping & How to Scrape Workopolis In this guide, we will show you how to scrape hiring companies, job listings, locations and salaries from the Workopolis website; all with our free web scraper	
How to Scrape Products From Any WooCommerce	<a href="https://www.parsehub.com/blog/scrape-products-woocommerce/">https://www.parsehub.com/blog/scrape-products-woocommerce/</a>	How to Scrape Products From Any WooCommerce In this tutorial, we will show you how to scrape products off a WordPress website, which is most likely built using the open-source WooCommerce plugin	
How to Scrape Capterra Reviews and Data	<a href="https://www.parsehub.com/blog/scrape-capterra-reviews-data/">https://www.parsehub.com/blog/scrape-capterra-reviews-data/</a>	How to Scrape Capterra Reviews and Data In this guide, we will teach you how to scrape Capterra reviews ( <a href="https://www.capterra.ca/">https://www.capterra.ca/</a> ), for free, in a few easy steps using ParseHub, our visual web scraper	
How to Scrape Canadian Tire Products	<a href="https://www.parsehub.com/blog/scrape-canadian-tire-products/">https://www.parsehub.com/blog/scrape-canadian-tire-products/</a>	How to Scrape Canadian Tire Products In this tutorial, we will show you how to easily scrape Canadian Tire ( <a href="https://www.canadiantire.ca/en.html">https://www.canadiantire.ca/en.html</a> ) products (in a single step) for free, using ParseHub	
2023 Guide: How to Scrape Airbnb Listing Data	<a href="https://www.parsehub.com/blog/scrape-airbnb-listing-data/">https://www.parsehub.com/blog/scrape-airbnb-listing-data/</a>	How to Scrape Airbnb Listing Data   Pars In this tutorial, we will show you how to easily scrape Airbnb listings data. Pars is a powerful web scraper that can help you collect data from any website. You might be trying to find the perfect vacation spot or trying to scope out your own	
2023 Guide: How to Scrape Data from Multiple Websites	<a href="https://www.parsehub.com/blog/web-scraping-multiple-urls/">https://www.parsehub.com/blog/web-scraping-multiple-urls/</a>	How to Scrape Data from Multiple Websites In this 2023 updated scraping tutorial, we will show you how to scrape data from multiple websites using ParseHub, our free web scraper ( <a href="https://www.parsehub.com/blog/web-scraping-multiple-urls/">https://www.parsehub.com/blog/web-scraping-multiple-urls/</a> )	
How to Scrape Dell Products	<a href="https://www.parsehub.com/blog/scrape-dell-products/">https://www.parsehub.com/blog/scrape-dell-products/</a>	How to Scrape Dell Products In this tutorial, we will show you how to scrape Dell products from their website using ParseHub, our free web scraper ( <a href="https://www.parsehub.com/blog/scrape-dell-products/">https://www.parsehub.com/blog/scrape-dell-products/</a> )	
Microsoft Teams or Cisco Webex? & 2020 vs. 2021	<a href="https://www.parsehub.com/blog/microsoft-teams-cisco-webex-2020-vs-2021/">https://www.parsehub.com/blog/microsoft-teams-cisco-webex-2020-vs-2021/</a>	Microsoft Teams or Cisco Webex? & 2020 vs. 2021 Cisco Webex and Microsoft Teams are the two main players when it comes to business video conferencing and they both offer a wide variety of options for	
How to Scrape Walgreens	<a href="https://www.parsehub.com/blog/scrape-walgreens/">https://www.parsehub.com/blog/scrape-walgreens/</a>	How to Scrape Walgreens Walgreens was first founded in 1901 in Chicago, originally as a family-owned grocery store. It has since grown to become USA's second-largest pharmacy chain	
[2023 Guide] How to Scrape and Download Images from a Website	<a href="https://www.parsehub.com/blog/scrape-images-website/">https://www.parsehub.com/blog/scrape-images-website/</a>	How to Scrape and Download Images from a Website Downloading lots of images from a website can be quite time-consuming. Right-click, Save Image As..., repeat ad nauseam. You may want to scrape images	

~The End~