

LISTLY

User Guide

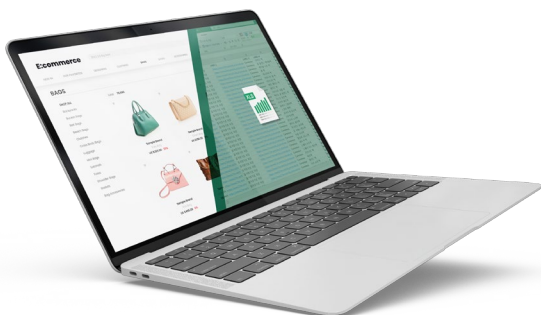
Listly is a free web scraping extension that can be used in Google Chrome, Microsoft Edge, and Naver Whale. Click Listly's extension icon on a destination website, and you'll instantly have structured data at your fingertips!

Listly Product Description Ver. Aug-10-2023
WWW.LISTLY.IO
Copyright 2023 LISTLY All rights reserved

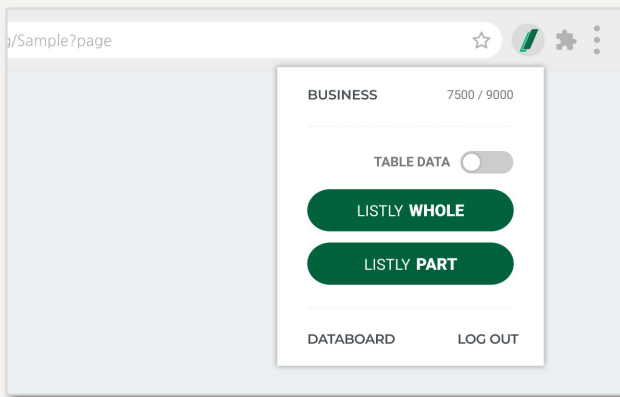


LISTLY Product Description

INDEX

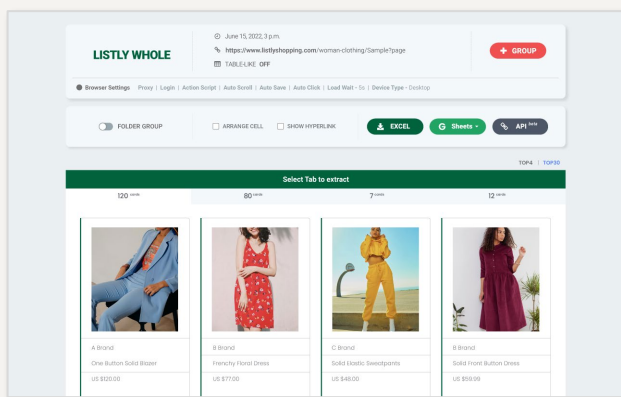


- 01 Listly Services
- 02 Listly Glossary
- 03 Extension Popup
- 04 Result Page
- 05 Data Board
- 06 Settings
- 07 Troubleshooting



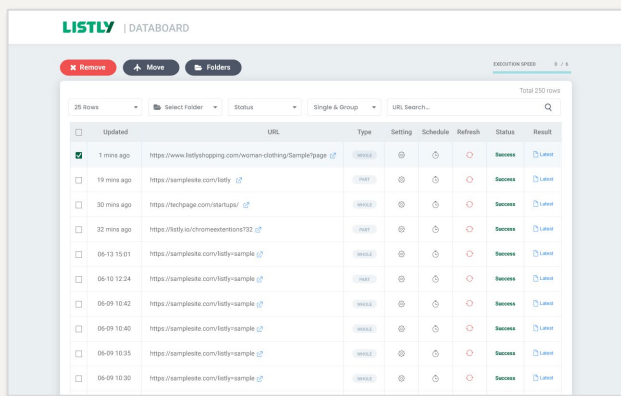
Extension Popup

From the extraction popup window, you can select whether to scrape an entire website or extract specific information. Click Listly's extension icon to get started and explore other features.



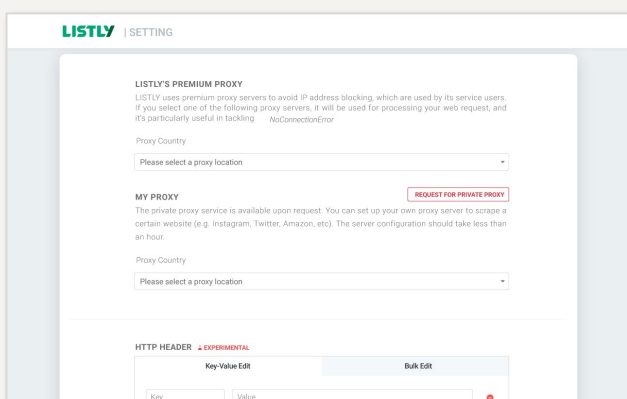
Result Page

After your web scraping is completed, you'll be taken to a page labeled "Listly Result." You can preview your web scraping results and download them in several different formats.



Data Board

The data board is fundamental to browsing data in Listly, and it provides quick access to recently-extracted web pages. Use the data board to see more detailed information about your web scraping tasks and customize settings.



Settings

Click the settings icon to access a variety of web scraping options: my proxy, device type, login, action script, auto-save, auto-scroll, auto-click, and more.

www.sampleurl.com/page1

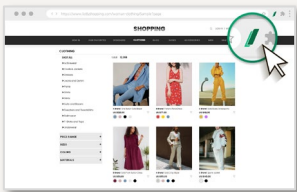
URL

A web address pointing to a target website, referring to a single web scraping task in Listly.

10-13 00:46	https://sampleite.com/listly	START	Success	100
10-11 17:45	Sample Site	WHILE	FAIL	0
06-13 15:01	https://sampleite.com/listly/sample	WHILE	DURING	0
06-10 12:24	https://sampleite.com/listly/sample	START	Success	0

Task

A task unit for web scraping, often used to refer to either a single task or a group task.



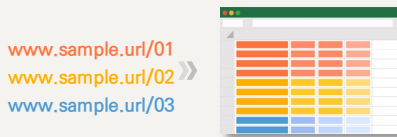
Target Page

A destination website to scrape data from.



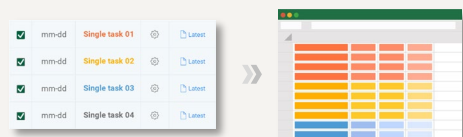
Source Page, Group Source

A web page that can be used as the start URL for group extraction.



Group Extraction

You can scrape data from multiple web pages at once and download them as a single file.



Folder Download

You can select **multiple tasks** to group them together as a folder and export data to a single file.

1,000 cards

Card

A content container and single unit of information extracted from a website, and the number refers to the container counts.

Select Tab to extract

Tab

The tab field is where you view many different cards. The cards will be presented in order of relevance.



Listing Page

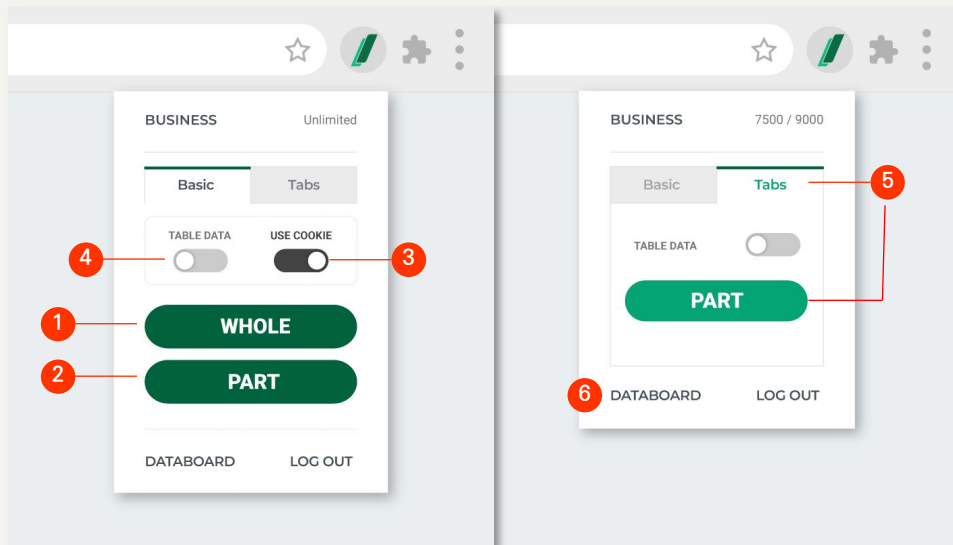
A single web page that presents a list of items within a certain category, such as a product listing page.



Detail Page

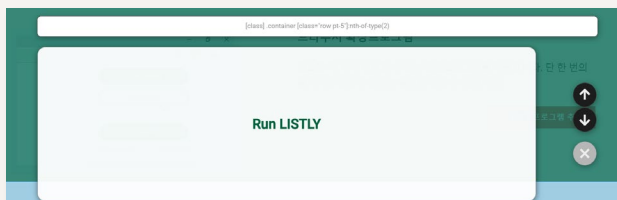
An individual web page that can be accessed from a listing page, which contains information about a specific item.

Extension Popup



- 1 **WHOLE** Collect data from the entire web page by automatically analyzing the website structure. This button will take you to the results page which recommends tabs that are most likely to include the information that needs to be extracted.

- 2 **PART**



LIExtract data by allowing you to select a single element at a time. Once clicked, the above layer will appear to scrape the page.

↑ : Click the up arrow to select the larger portion of the page.

↓ : Click the down arrow to select the smaller portion of the page.

Run Listly : Extract data from the selected part of the page.

- 3 **USE COOKIE**

Save information about your visit and use cookies that expire in one hour.

- 4 **TABLE DATA**

Use this feature to extract tabular data which consists of rows and columns. By default, it is set to OFF. You can turn on the toggle button and redo the extraction if you are not satisfied with the results.

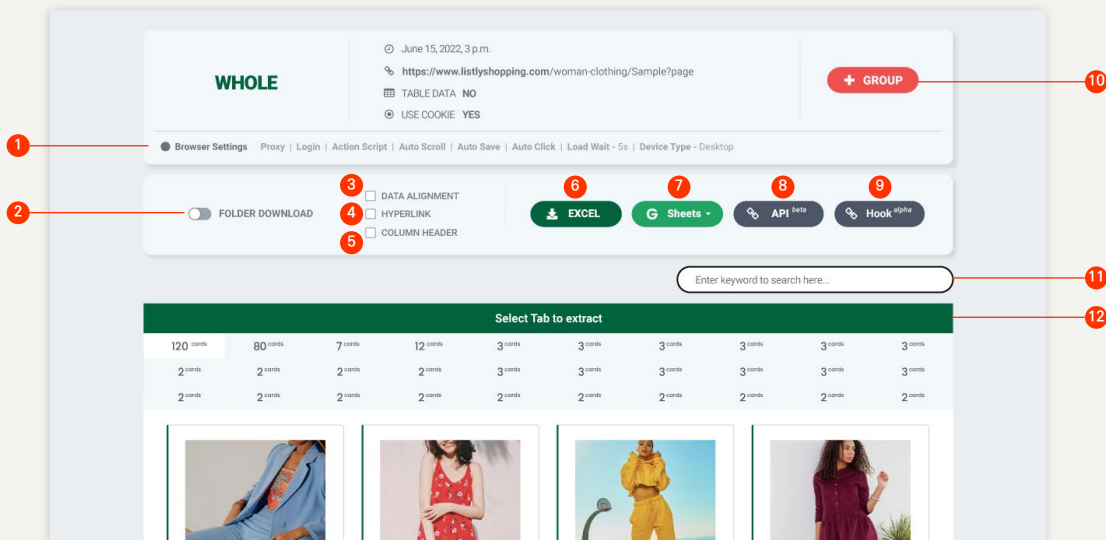
- 5 **Tabs-PART**

Extract pattern-based data that contain the same elements from any open browser tabs.

- 6 **DASHBOARD**

Redirect you to your Listly account databoard.

Result Page



- | | | |
|----|------------------------|---|
| 1 | Browser setting | Click on each button to access the settings page and change the browser settings. |
| 2 | Folder Download | If the toggle button is enabled, folder groups will pop up. Use this feature to access the information extracted and saved in a specific folder before. If you gather data using Tabs, the information from each web page will be automatically grouped and saved in a single folder. |
| 3 | Data Alignment | Automatically sort your web scraping results based on the patterns in data. |
| 4 | Hyperlink | Extract all URLs from hyperlinks. |
| 5 | Column Header | If enabled, create the variable names and labels as headers and export them to the first row in Excel or Google Spreadsheets. |
| 6 | Excel | Export data into Excel. |
| 7 | G Sheets | Export data into Google Sheets. |
| 8 | API | Use the Listly API to extract data in '.json' and '.csv' formats. |
| 9 | HOOK | Enables you to get or refresh the cached result after your web scraping tasks are done. |
| 10 | + GROUP | Group multiple web pages together to scrape by allowing you to enter sequential URLs starting from the current page URL. |
| 11 | Search Bar | Quickly access cards that contain specific content by searching for keywords. |
| 12 | Tab | Select a tab that contains the data you want. By default, a group of four tabs will be presented in order of relevance. |

*Click on each glossary term for more information.

Group Extraction Results Page

The screenshot shows the 'Group Extraction Results Page' interface. At the top, there is a 'Sample site' link (1). Below it are buttons for 'GROUP EXCEL' (2), 'Sheets' (3), and 'LOGS' (3). There are also checkboxes for 'ARRANGE CELL' and 'SHOW HYPERLINK'. A 'Group Source' section (4) includes 'REMOVE' (5) and 'REFRESH' (6) buttons, and a '+ URL' button (7). The main table (8) has columns for 'Updated', 'URL', 'Status', 'Tab Matched', 'Screenshot', and 'Result'. The table contains 8 rows of data, all with a 'Success' status. A search bar and filters for '100 Rows', 'Status', and 'Tab Matched' are located above the table.

Updated	URL	Status	Tab Matched	Screenshot	Result
09-27 17:37	https://samplesite.com/listly/page=01	Success	Y	-	Latest
09-27 17:42	https://samplesite.com/listly/page=02	Success	N	Screenshot	Latest
09-27 17:43	https://samplesite.com/listly/page=03	Success	N	Screenshot	Latest
09-27 17:40	https://samplesite.com/listly/page=04	Success	Y	-	Latest
09-27 17:39	https://samplesite.com/listly/page=05	Success	Y	-	Latest
09-27 17:41	https://samplesite.com/listly/page=06	Success	Y	-	Latest
09-27 17:43	https://samplesite.com/listly/page=07	Success	Y	-	Latest

- 1 **Group Name** Rename your extraction group

- 2 **Group Excel** You can use the Group Excel button to download the results of your group extraction task in Excel.

- 3 **Logs** Click Logs to go back and forth in your group extraction history.

- 4 **Group Source** Preview the information about the web page used as the start URL for group extraction.

- 5 **Remove** Remove the selected tasks from the group extraction results or the data board page.

- 6 **Refresh** Redo your web scraping task.

- 7 **+ URL** Enter the URL(s) you want to add to the current group.

- 8 **Tab Matching** Display the structure and consistency of the results of your web scraping task.

- 9 **Screenshot** Show a screenshot of your web scraping error message.

*Click on each glossary term for more information.

Data Board

The screenshot shows the LISTLY DASHBOARD interface. At the top, there are several action buttons: Remove (1), Move (2), Copy (3), Folders (4), Add, and Remove. On the right, there is a Reset button (5) and an Execution Speed indicator (6). Below these are filters for Rows (25 Rows), Select Folder, Status, Single & Group, Setting, and Schedule. A search bar is labeled 'URL, Group Search...'. The main table has columns: Updated, URL, Type, Setting, Schedule, Refresh, Status, and Result. The table contains several rows of tasks with various URLs and statuses. Red numbered callouts (7-11) point to specific columns: 7 to Type, 8 to Setting, 9 to Schedule, 10 to Status, and 11 to Result.

- | | | |
|----|-------------------------------|---|
| 1 | Remove | Remove selected tasks from your data board. |
| 2 | Move | Move selected tasks to a new folder. |
| 3 | Copy | Duplicate or copy selected tasks. Use this feature to change settings in part for single extraction or update URLs for group extraction with the same settings. |
| 4 | Folders | Create, edit, and remove folders. |
| 5 | Reset | If your web scraping job failed for long hours due to extraction speed errors, you can force quit all queued jobs. |
| 6 | <u>Execution Speed</u> | Adjust the extraction speed by setting the number of web pages to be extracted simultaneously. |
| 7 | Settings | Configure data extraction settings by task or by folder. |
| 8 | <u>Schedule</u> | Automatically extract data at a scheduled time. |
| 9 | Refresh | Redo data extraction. If you updated the extraction settings, you may want to click on the Refresh button to scrape data again. |
| 10 | Status | Check the status of your data extraction. The status provides information on the results of your web scraping task: SUCCESS / FAIL / DOING / WAITING. |
| 11 | Result | Redirect you to the results page. |

*Click on each glossary term for more information.

Settings

LISTLY'S PREMIUM PROXY

LISTLY uses premium proxy servers to avoid IP address blocking, which are used by its service users. If you select one of the following proxy servers, it will be used for processing your web request, and it's particularly useful in tackling `NoConnectionError`.

Proxy Country
Please select a proxy location

MY PROXY [REQUEST FOR PRIVATE PROXY](#)

The private proxy service is available upon request. You can set up your own proxy server to scrape a certain website (e.g. Instagram, Twitter, Amazon, etc). The server configuration should take less than an hour.

Proxy Country
No Proxy

HTTP HEADER [EXPERIMENTAL](#)


Key-Value Edit Bulk Edit

Key Value

[+ Add](#)

DEVICE TYPE

Desktop Tablet Mobile



LOGIN [BETA](#)

Warning This is the experimental feature. We store the ID and Password as a plain text without any secured encryption. Please remove the information after your extraction finishes. We will delete all the information at the next update for better security.

Please select a login account [Edit](#)

Use cookies instead of ID and Password

ACTION SCRIPT [VIEW EXAMPLE](#)

Enter JavaScript code here.

For any inquiries, please contact the LISTLY team.

WAIT FOR PAGE LOAD (SECONDS) [VIEW EXAMPLE](#)

0 min 60 sec

0 15 30 45 60

Proxy

Use a proxy setting to avoid getting blocked by the destination website when scraping from it. You can select either public or private proxies.

HTTP HEADER

HTTP headers allow you to avoid your web scraping from getting blocked and scrape the destination website by utilizing additional information during HTTP requests or responses.

Device Type

You can select a device type in mobile or tablet views.

Auto Login

Keep your login credentials safe with Listly's advanced encryption algorithm so that you can web scrape without signing in each time.

Action Script

Automate and/or combine repetitive mouse actions including clicks, movements, and scrolling.

Start Timer(Wait for page load)

Set a countdown timer to start scraping, allowing web pages to completely load and avoid empty data scraping.

*Click on each glossary term for more information.

Settings

AUTO SCROLL TO LOAD MORE DATA [VIEW EXAMPLE](#)

N Times to repeat: Wait for:

Auto save while scrolling *(beta)* [VIEW EXAMPLE](#)

* Turning on this allows you to save your data automatically while scrolling.

Auto Scroll

Automatically scrolls down on the destination website to load more data.

AUTO CLICK TO LOAD MORE DATA [VIEW EXAMPLE](#)

Element to auto click (CSS Selector):

N Times to repeat: Wait for:

Auto Click

Makes mouse clicks on a specific button repeatedly to load more data.

Schedule

Current Time

Asia/Seoul
2022-12-22 17:16 (Thursday)

1 **Timezone**

2 **Period**

3 **Time**

4 **E-mail Alert**

1 **Timezone**

Select a time zone for your web scraping task.

2 **Period**

Specify a set of frequencies and repeat intervals on a daily, weekly, or monthly basis.

3 **Time**

Set up a time for your web scraping task.

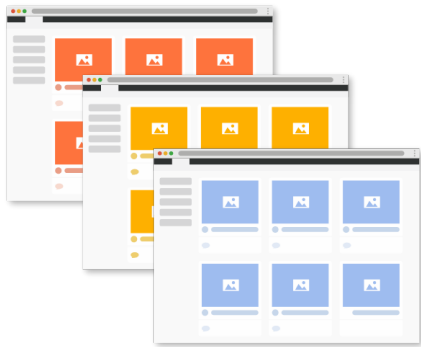
4 **E-mail Alert**

Send email notifications to inform you about the web scraping results if enabled.

*Click on each glossary term for more information.

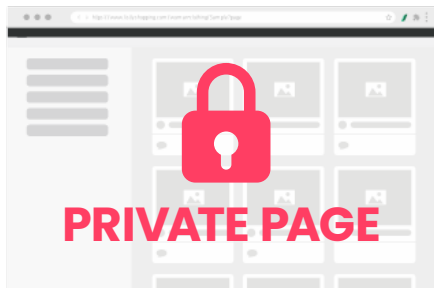
Troubleshooting

Check out the most common challenges of web scraping and how to address them. Browse our collection of helpful articles and find the answers to your question on [Listly's Help Center](#).



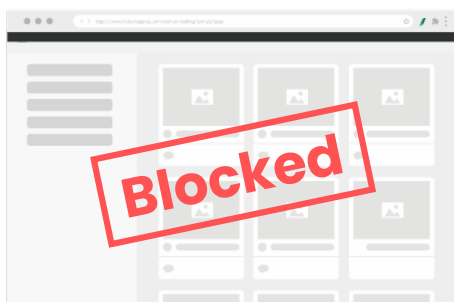
Scraping multiple web pages

Unsure where to start scraping multiple pages? Check out Listly's data extraction flowchart on the following page.



Configuring login settings

Wondering how to scrape data without having to sign in each time? Find out how to configure settings for websites that require authentication.

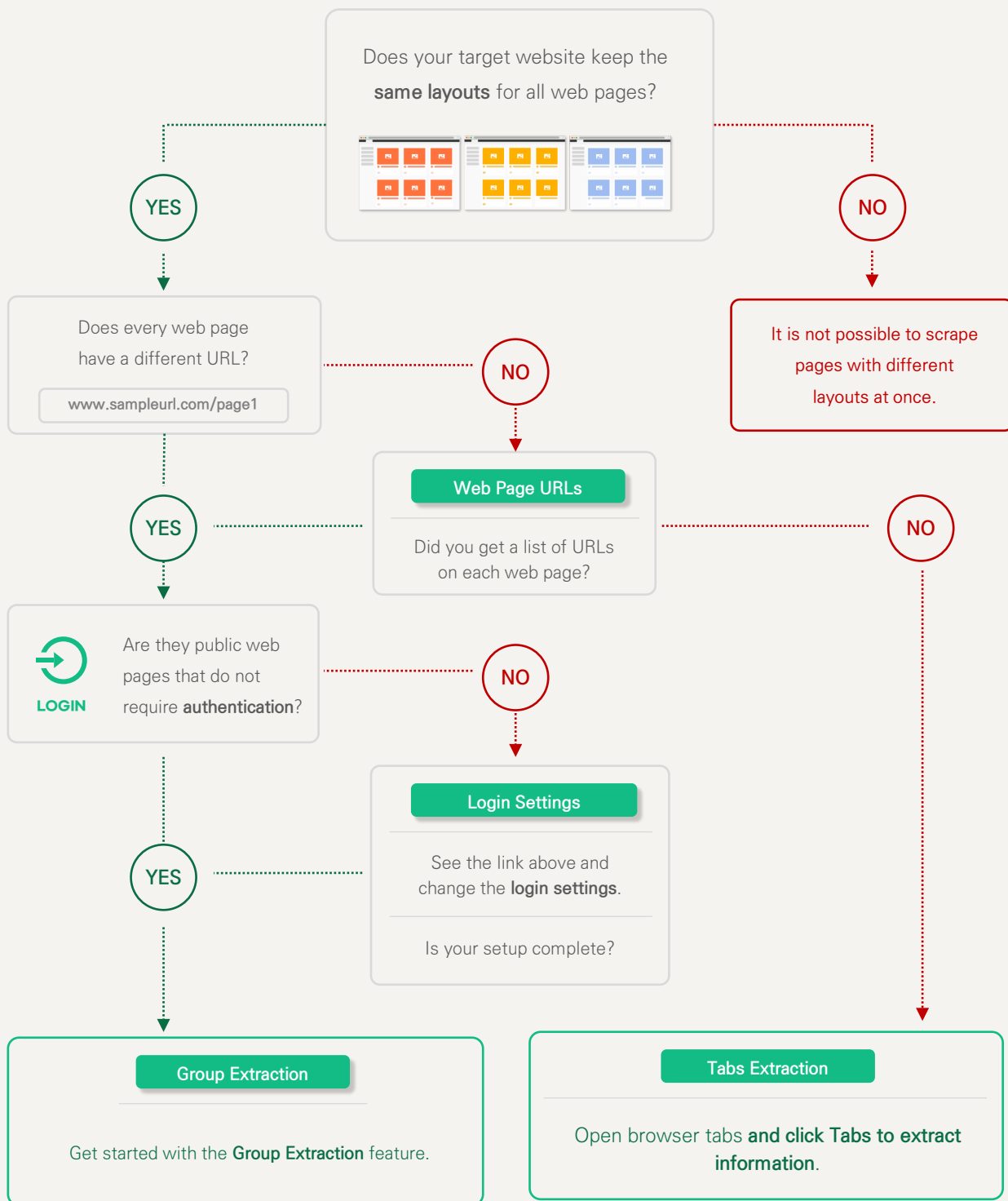


Web scraping without getting blocked

How do I avoid getting blocked? Learn more about simple tips to handle IP blocks and errors.

Scraping multiple web pages

You can scrape data from multiple web pages in two ways: **Group Extraction** and **Tabs Extraction**. Please continue to check the following flowchart that can help you get started in Listly.



Configuring login settings

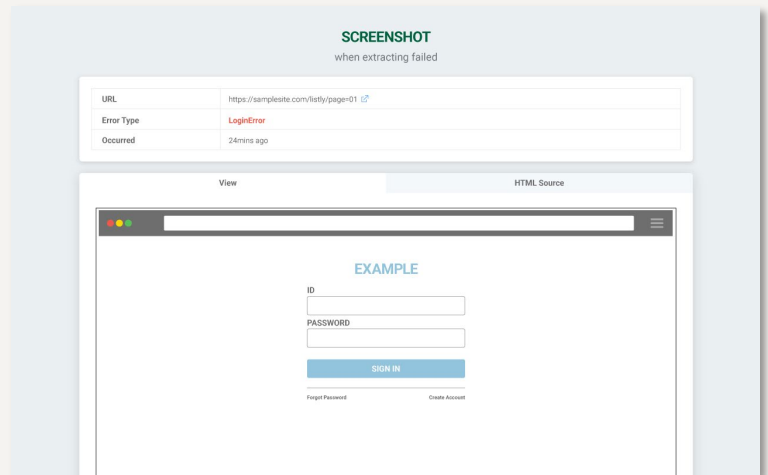
Listly keeps your login credentials safe with advanced encryption algorithms so that you can web scrape without signing in each time. However, if login cookies that are deleted after a specific period of time, you may want to set up your login settings as follows:

1. Any web scraping tasks scheduled at a specific time
2. Any group tasks that scrape data from multiple web pages using URL inputs

e.g. Login Error Screenshot

A screenshot of the login error captured from any previously viewed pages will show up if there is an error logging you in.

Login Settings



Unfortunately, the login feature of this beta version does not support the following websites:

1. Any websites that require two-step authentication or region-restricted websites (most websites that offer email services)
2. Any websites where users are prompted to input their username and password on separate web pages
3. Any websites that use a login modal popup where a unique URL is not specified
4. Any websites whose login URL contains complex string parameters

You may want to check out the following option if you want to scrape data from multiple open browser tabs without logging in multiple times.

TABS

Open browser tabs to scrape and click the Tabs button after signing into a website. With this feature, you can export data into a single folder, available to download into a spreadsheet. You may expect the same results from the Folder Download feature. Please note that this job cannot be redone.

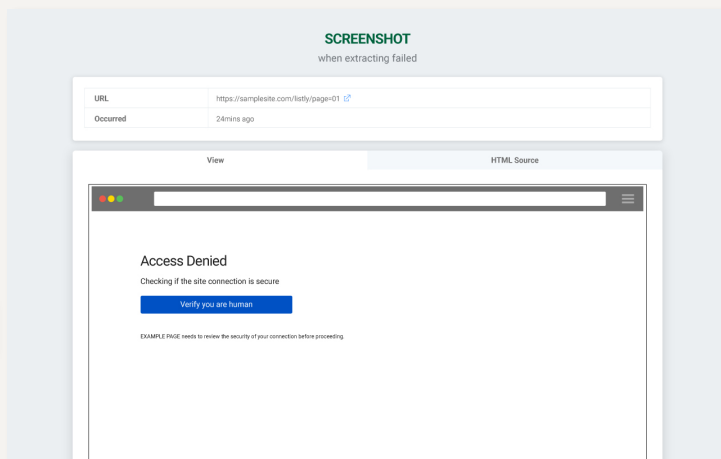
Web scraping without getting blocked

There may be cases where your IP address is getting blocked in the data collection process if the destination website detects unusual traffic to restrict access from a specific IP. You can simply solve this problem by changing proxy settings.

e.g. Blocked website screenshot

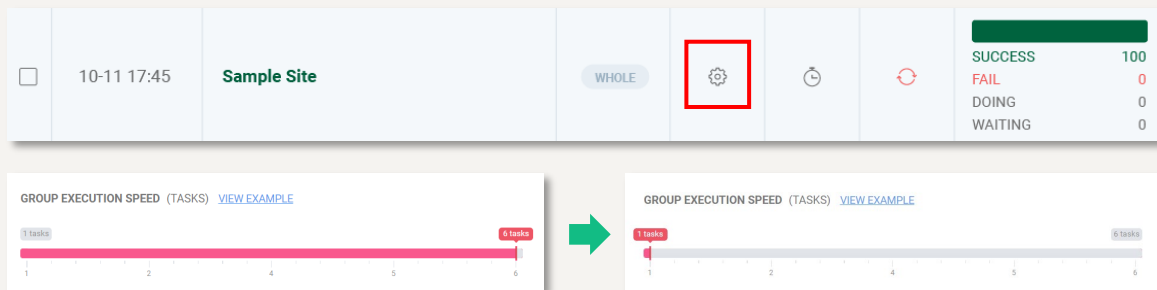
A screenshot of an error message, "Access Denied", will show up.

Proxy Settings



+) How to change extraction speed

Your destination website may restrict or block access by your IP address if it detects 'bot traffic' over a short period of time. To securely collect data from the website, you may want to lower the extraction speed to 1 by configuring **Settings → Group Execution Speed**.



If you want to adjust the speed of your web scraping task, change the execution speed to the upper right corner.

