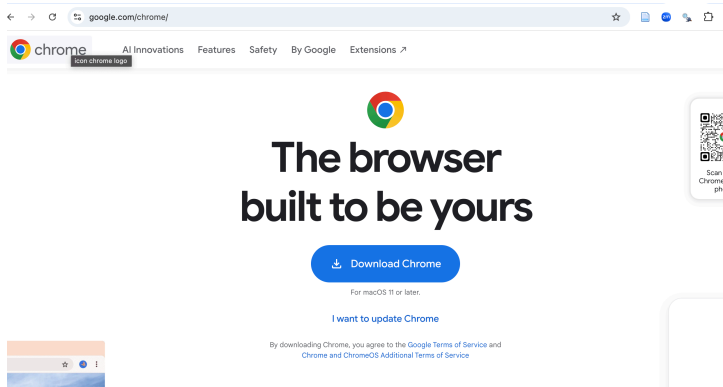
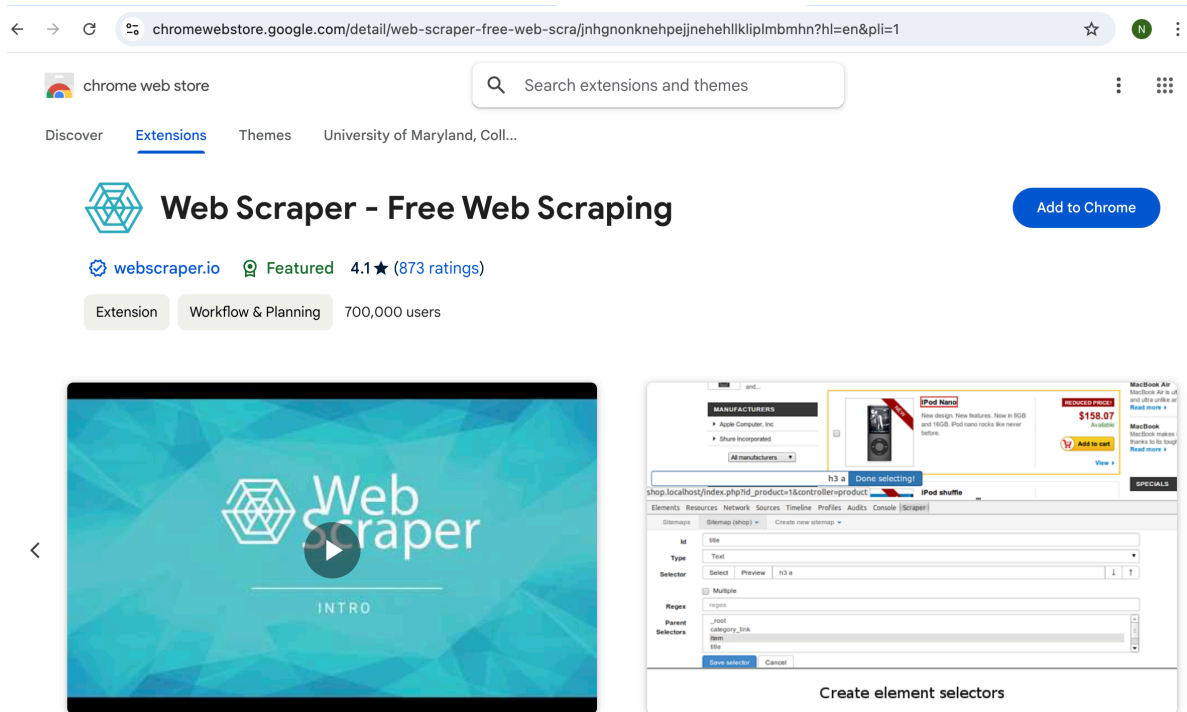


webscraping for a static webpage

Step 1: Install Chrome and find the webscraper extension



Step 2: Add the webscraper extension to your Chrome. Once added, it should appear in the right sidebar of your browser.



chrome web store

Search extensions and themes

Discover **Extensions** Themes University of Maryland, Coll...

Web Scraper - Free Web Scraping [Add to Chrome](#)

[webscraper.io](#) [Featured](#) 4.1★ (873 ratings)

Extension Workflow & Planning 700,000 users

Web Scraper
INTRO

MANUFACTURERS
Apple Computer, Inc.
Shure Incorporated
All manufacturers

Done selecting!

Create element selectors







Step 3: Open the website (or use keywords “infopemilu dpr indonesia”):
https://infopemilu.kpu.go.id/Pemilu/Dct_dpr. Click on “pilih dapil” and select “Aceh I”

← → ↻ infopemilu.kpu.go.id/Pemilu/Dct_dpr ☆ | N ⋮

← DAFTAR CALON TETAP DPR

Data ini berdasarkan nomorurut partai
 Pilih Nama Dapil:
 ACEH I

Show entries Search:

Partai Kebangkitan Bangsa	Partai Kebangkitan Bangsa	Partai Kebangkitan Bangsa	Partai Kebangkitan Bangsa	Partai Kebangkitan Bangsa	Partai Kebangkitan Bangsa
Dapil ACEH I district	Dapil ACEH I	Dapil ACEH I	Dapil ACEH I	Dapil ACEH I	Dapil ACEH I
nomorurut 1 ballot list	nomorurut 2	nomorurut 3	nomorurut 4	nomorurut 5	nomorurut 6
					
name H. IRMAWAN, S.Sos., M.M.	name M. RAMADHANA RUSLI BINTANG	name MAHDALENA, S.Hut.	name TEUKU HAMZAH HUSEN, S.E., M.M.	name Tgk. MUJLISAL, S.Ag.	name drh. HASNIDAR
gender LAKI - LAKI	gender LAKI - LAKI	gender PEREMPUAN	gender LAKI - LAKI	gender LAKI - LAKI	gender PEREMPUAN
origins KOTA BANDA ACEH	origins ACEH BESAR	origins ACEH BESAR	origins KOTA BATAM	origins PIDIE JAYA	origins ACEH BESAR
PROFIL	PROFIL	PROFIL	PROFIL	PROFIL	PROFIL

additional informations on candidates

Step 4: Go to 'More Tools' in the right sidebar of Chrome (the three dots symbol), then select 'Developer Tools.' You should be able to see the Web Scraper extension in the right pane of your Developer Tools.

The screenshot shows a web browser window displaying a list of candidates for the DPR. The URL is infopemilu.kpu.go.id/Pemilu/Dct_dpr. The page title is "DAFTAR CALON TETAP DPR". The page content shows a table of candidates, with the first column being the party name "Partai Kebangkitan Bangsa" and the second column being the candidate's name "Dapil ACEH I". The third column shows the candidate's number (1 through 6). The fourth column shows the candidate's photo. The browser's developer tools are open, showing the "Elements" pane on the left and the "Web Scraper" pane on the right. The "Web Scraper" pane is highlighted with a red circle. The "Web Scraper" pane shows a list of elements to be scraped, including the candidate's name and number. The "Web Scraper" pane also shows a search bar and a "Show" button.

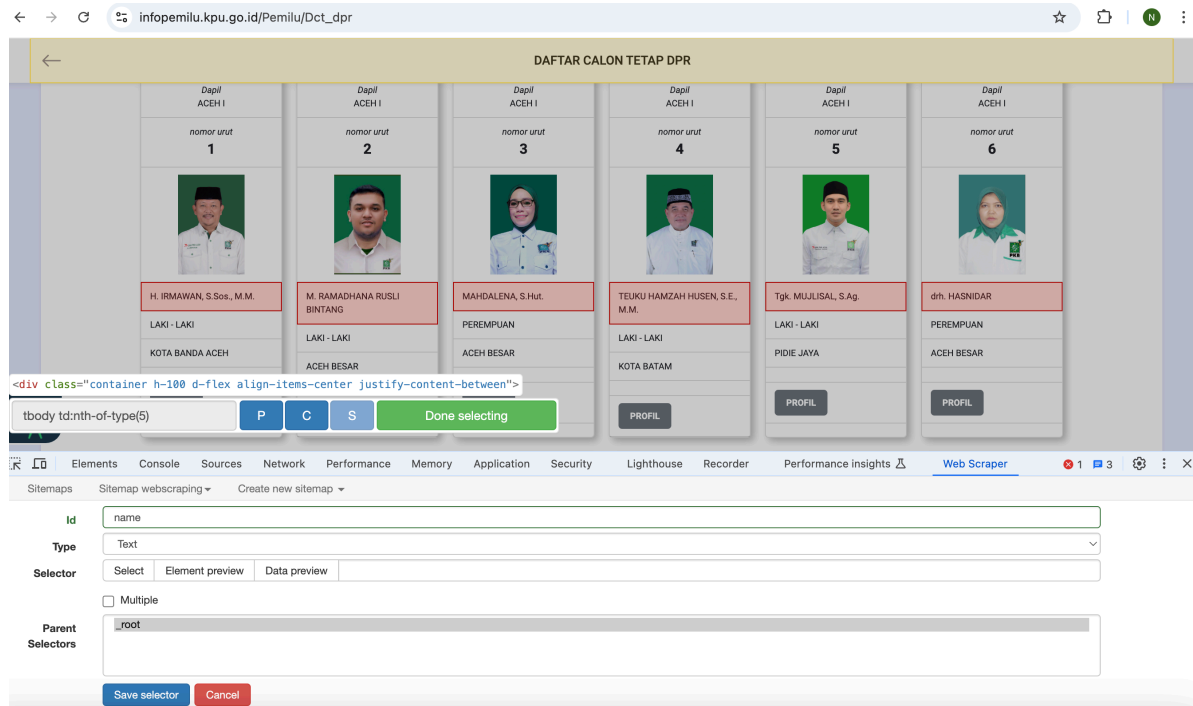
Partai Kebangkitan Bangsa	Dapil ACEH I	nomor urut	Photo
Partai Kebangkitan Bangsa	Dapil ACEH I	1	[Photo]
Partai Kebangkitan Bangsa	Dapil ACEH I	2	[Photo]
Partai Kebangkitan Bangsa	Dapil ACEH I	3	[Photo]
Partai Kebangkitan Bangsa	Dapil ACEH I	4	[Photo]
Partai Kebangkitan Bangsa	Dapil ACEH I	5	[Photo]
Partai Kebangkitan Bangsa	Dapil ACEH I	6	[Photo]

Step 5: Let's start scraping! In the webscraper tab, select 'create new sitemap' and pick 'create sitemap'

The screenshot shows a web browser at the URL `infopemilu.kpu.go.id/Pemilu/Dct_dpr`. The page title is "DAFTAR CALON TETAP DPR". The main content area is titled "Pilih Nama Dapil: ACEH I" and displays a list of candidates for the Partai Kebangkitan Bangsa (PKB) in Aceh I. The candidates are listed in a grid with their names, order numbers, and photos. Below the browser window, the 'Web Scraper' extension interface is visible. The 'Sitemaps' tab is selected, and the 'Create new sitemap' dropdown is open, showing 'Create Sitemap' as an option. Annotations with pink arrows point to the 'Sitemap name' field (containing 'webscraping') and the 'Start URL 1' field (containing 'https://infopemilu.kpu.go.id/Pemilu/Dct_dpr'), with text indicating to 'pick whatever name' and 'copy and paste the url' respectively.

- After you put the name and paste the url, select 'create sitemap'. You should be able to see your ID and domain in the 'sitemaps' pane.
- Click on you ID, select 'add new selectors'.

Step 6: This is the MOST IMPORTANT step. We will start selecting elements on the webpage that we want to scrape.















- Let's start with the name of the candidates. Select id and write 'name'. This is going to be the name in our variable.
- In the 'selectors' tab, click on 'select'. Your cursor will now allow you to select the Java/HTML elements that you want. Be careful not to highlight any other elements. Select only the 'name'.
- Now if you click the name 'H. IRMAWAN, S.Sos., M.M.' once, the scraper will automatically sort only him. But if you click the name next to him, 'M. RAMADHANA RUSLI BINTANG', the scraper will automatically sort ALL THE NAMES on the webpage.
- So, whenever you click the second item in your selected element, the scraper will automatically include all similar elements on the webpage.
- Click 'done selecting', check 'multiple', and 'save selector'. Now you should be able to see all the scraped names in the 'data preview'.

Step 7: We want to scrape 'party' for the second variable. Follow the same steps as we did for 'name', and write 'party' as the ID.

← → ↻ infopemilu.kpu.go.id/Pemilu/Dct_dpr ☆ ⓘ 11 ⋮

← DAFTAR CALON TETAP DPR

Show ▾ entries Search:

 Partai Kebangkitan Bangsa	 Partai Kebangkitan Bangsa	 Partai Kebangkitan Bangsa	 Partai Kebangkitan Bangsa	 Partai Kebangkitan Bangsa	 Partai Kebangkitan Bangsa
Dapil ACEH II	Dapil ACEH II	Dapil ACEH II	Dapil ACEH II	Dapil ACEH II	Dapil ACEH II
nomor urut 1	nomor urut 2	nomor urut 3	nomor urut 4	nomor urut 5	nomor urut 6
					
MUHAMMAD ADAM M. Ed.	ANA FITRIA SARI PINEM, S.H.	H. DARKASYI, S.H.	SUMADI ARSYAH	ARINI ARISKA, A.Md.Keb	ARINI ARISKA, A.Md.Keb
PEREMPUAN	PEREMPUAN	LAKI - LAKI	LAKI - LAKI	PEREMPUAN	PEREMPUAN

`<div class="card-body" data-select2-id="95">`

`tbody td:nth-of-type(1)` P C S Done selecting

Elements Console Sources Network Performance Memory Application Security Lighthouse Recorder Performance insights Web Scraper

Sitemaps Sitemap webscraping Create new sitemap

Id party

Type Text

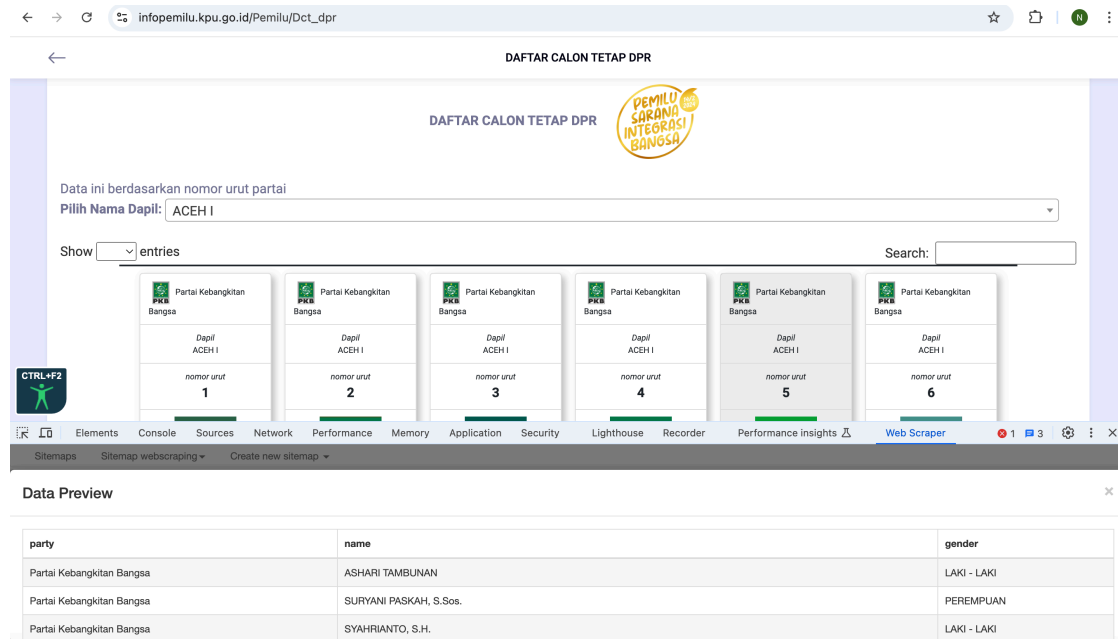
Selector Select Element preview Data preview

☐ Multiple

Parent Selectors `root`

Save selector Cancel

Step 8: Let's do it once again for 'gender' variable. Once done, you should be able to gather data that looks like this in your 'data preview' tab.



DAFTAR CALON TETAP DPR

Data ini berdasarkan nomor urut partai

Pilih Nama Dapil: ACEH I

Show entries Search:

nomor urut	nama	gender
1	ASHARI TAMBUNAN	LAKI - LAKI
2	SURYANI PASKAH, S.Sos.	PEREMPUAN
3	SYAHRIANTO, S.H.	LAKI - LAKI

Web Scraper

Step 9: You can now pull in any other districts and get the same data without having to set up the variables again!

- All you have to do is select any district, and click on 'data preview'.
- Once you're done, you import this into excel/csv and ready to do your own analysis.

