

# Extract Domain From a URL (with Google Sheets)

---

 [lupagedigital.com/blog/extract-domain-from-url-google-sheets](https://lupagedigital.com/blog/extract-domain-from-url-google-sheets)

April 2, 2021

Extracting a domain from a URL is often tedious and time-consuming. You might have come across this tasks from your managers. When working in SEO and digital analytics, stumbling upon URLs is a common occurrence. What's dull is cleaning those URLs to simple domain for data reporting.

If you have a URL list that an web analyst needs to parse through, it takes time and energy to manually type in the domain. Whether [adding columns and rows](#), Google Sheets is a great tool to make things easier.

**That being said, let's start. Create a new Google Sheets file on your computer.**

- 1 – Go to your [Google Drive](#) workspace.
- 2 – Click '+ New' then create a new Google Sheets file.

## How To Extract Domains From URLs with Google Sheets

---

There are many ways on how to extract domains from URLs. Some are still tedious and some are functions built into Google Sheets. There's no ultimate solution to this and the one below is the one I've used. No need for coding knowledge but knowing how to read syntax helps.

### Method 1: Find and Replace (Manual)

---

The first method to extract domains is "Find and replace" in Google Sheets. This is the easiest method because it doesn't need any code syntax in the spreadsheet. It's pretty intuitive and you're relying on the user interface only.

- 1 – Open a Google Sheets file.
- 2 – Paste a sample URL or a list of URLs to Column A.

Extract Domain From URL ☆ 📄 📁

File Edit View Insert Format Data Tools Add-ons Help Last edit was 2 mir

150% \$ % .0 .00 123 Default (Ari... 10 B I

	A	
1	URL	
2	https://www.google.com/search?q=test	
3	https://www.youtube.com/results?search_query=test	
4	https://blog.hubspot.com/marketing/how-to-start-a-blog	
5	https://domain.com/product?size=xl	
6		
7		

3 – Highlight all the URLs inside the column.

4 – Press CTRL + H for PC or Shift + Command + H for Mac. This is where you “find and replace” characters.

5- To start, type in https:// then leave “Replace with” empty.

### Find and replace ×

Find

Replace with

Search

Match case

Match entire cell contents

Search using regular expressions [Help](#)

Also search within formulas

6 – Click “Replace all”.

Then you need to repeat the entire process with the rest of the characters. Type in http://, click then click Replace all. Type in /, then click Replace all.

### Downsides:

It’s obviously not scalable. Imagine doing this with thousands of URLs. Apart from human error, it’s a nightmare once there are new batches of URLs added to the data. It’s not automatic and it requires more hours into it. On top of that, it doesn’t delete generated UTMs.

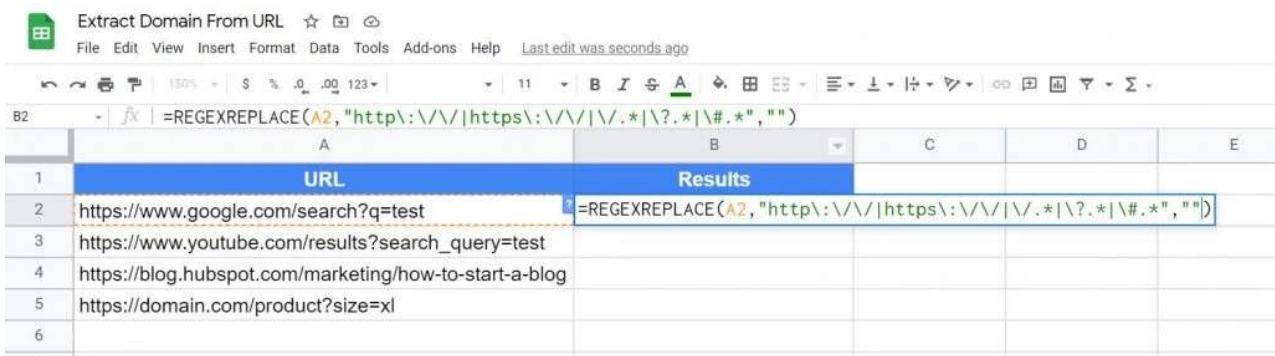


## Method 2: Regex Replace Function (Automated)

The `=REGEXREPLACE()` function is built-in to Google Sheets and it extracts domains from URLs. What’s great about is it’s only a simple line of code that you can paste into your cell. The function is not super technical and you can change it any way you see fit.

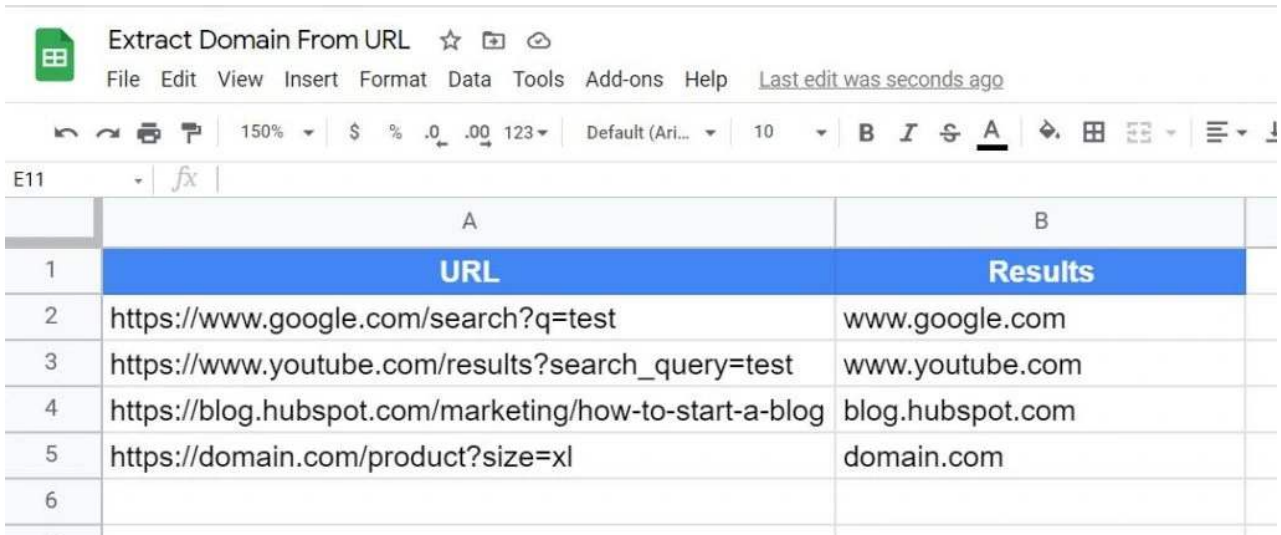
- 1 – Open a Google Sheets file.
- 2 – Paste a sample URL or a list of URLs to Column A.
- 3 – Add this function on column B beside the URLs.

```
=REGEXREPLACE(A2, "http://|https://|\\.|\\?|\\#|\\.", "")
```



Three parameters are used in the function: `=REGEXREPLACE(x, y, z)`  
x = the needed text that is replaced. In your case, the spreadsheet cell.  
y = the regular expression. The pattern that matches the characters.  
z = the new text that replaces x.

4 – The function now populates column B with the extracted domains from the URLs.



	A	B
1	URL	Results
2	https://www.google.com/search?q=test	www.google.com
3	https://www.youtube.com/results?search_query=test	www.youtube.com
4	https://blog.hubspot.com/marketing/how-to-start-a-blog	blog.hubspot.com
5	https://domain.com/product?size=xl	domain.com
6		

### Let me further explain:

Regexreplace is a regular expression that matches specific characters. And then replaces them with new characters. Hence, the name. This is a built-in function in Google Sheets

With the above, this will remove **HTTPS://**, **HTTP://**, and queries like generated UTM parameters. It also removes the forward slash after the top-level domain (**.com**, **.org**, **.co**). And it removes all the characters that come after.

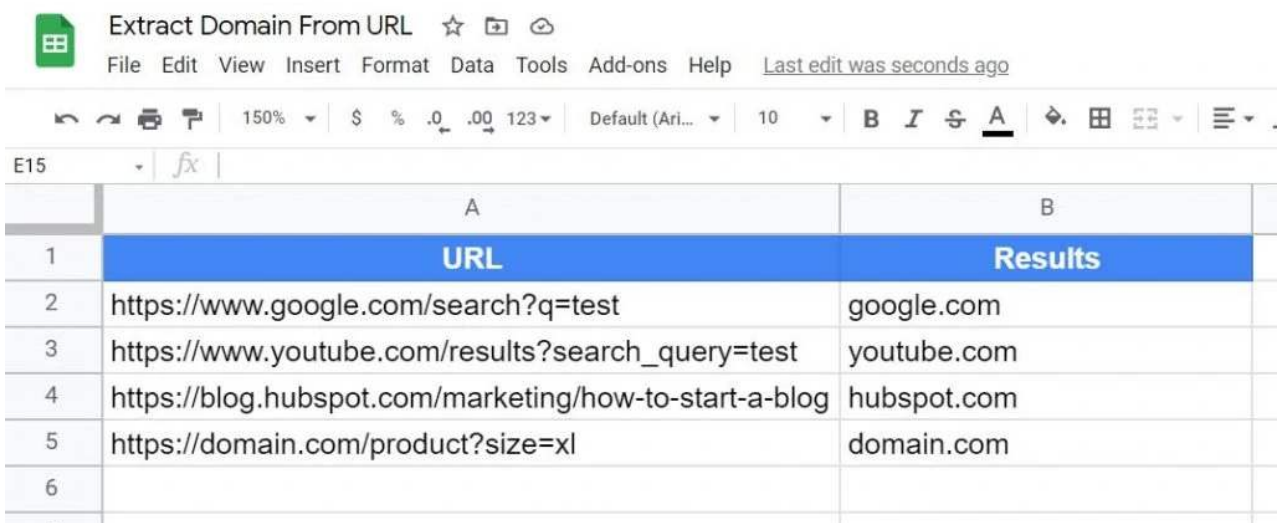
See this example:

https://www.domain.com/product?size=xl. The red colors are removed.

### Extract The Root Domain From a URL

Once you get a hang of regex replace, you can tweak the regular expression and remove subdomains. You can remove **www.** or **blog.** and extract the root domain. Add these in the regular expression: **www\.|blog\.**

```
=REGEXREPLACE(A2, "http\\:\\\\|https\\:\\\\|www\\.|blog\\.|\\/.*|\\?.*|\\#.*", "")
```



	A	B
1	URL	Results
2	https://www.google.com/search?q=test	google.com
3	https://www.youtube.com/results?search_query=test	youtube.com
4	https://blog.hubspot.com/marketing/how-to-start-a-blog	hubspot.com
5	https://domain.com/product?size=xl	domain.com
6		

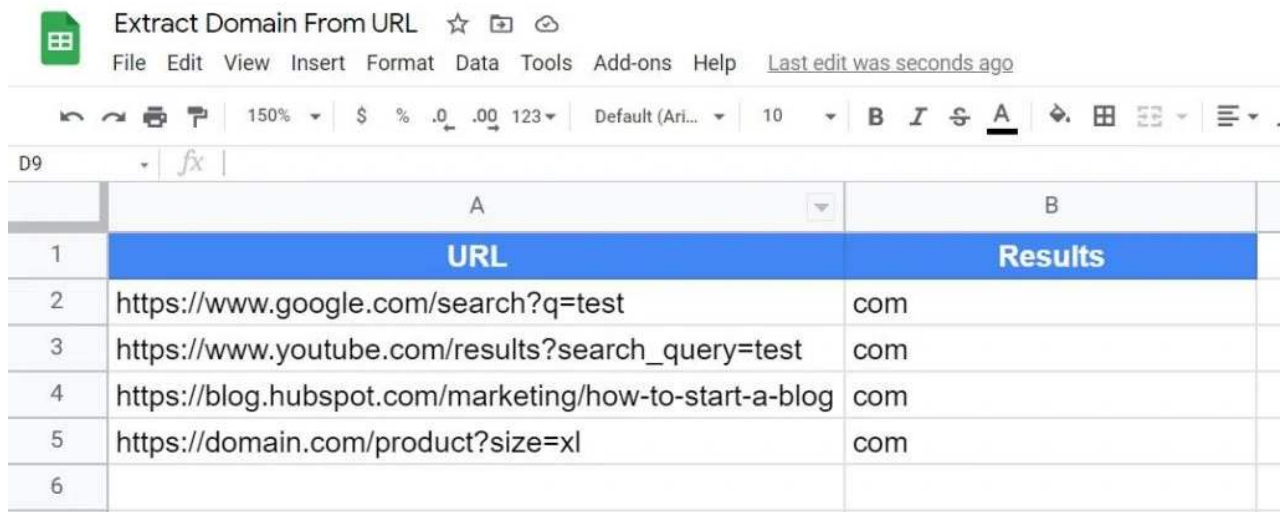
You can add more characters and remove subdomains if you want depending on your websites. It could be `api.`, `docs.`, `uat.`, `dev.`, `test.`, `news.`, and many more.

## Extract The Top-Level Domain From a URL

---

To take it even further, you can tweak the regular expression to extract the top-level domain from a URL. This is the `.com`, `.org`, `.co`, and all the others. This syntax is much simpler and shorter. This removes the subdomain, root domain, and query parameters.

```
=REGEXREPLACE(A2, ".*\.|\/.*", "")
```



	A	B
1	URL	Results
2	https://www.google.com/search?q=test	com
3	https://www.youtube.com/results?search_query=test	com
4	https://blog.hubspot.com/marketing/how-to-start-a-blog	com
5	https://domain.com/product?size=xl	com
6		

Viola! Try to find ways to improve this function. Regex replace isn't only for domain extraction but also for any data in spreadsheets. Work smarter, not harder. And especially work faster. Learn more about [regular expressions](#).

## Conclusion: Extract Domains from URLs with One Line of Code

---

Any analyst has ad-hoc reports that are dull and boring to do. If there's an opportunity to automate, take it. Don't waste time doing unimportant things. Focus on your more important tasks. Be patient in setting up automation. It's scalable in the end and the benefits are great.