

06a Google advanced search

The basic Google search is found at the site
www.google.co.za or www.google.com

The advanced search is found by appending "/advanced_search" e.g.
https://www.google.co.za/advanced_search

How the indexing of the web works

- The web consists of over 60 trillion pages
- Google has computers that crawl (search) the web. The program will read a page, index key words, and look for links to other pages.
- It goes to those links and repeats the process
- A site manager can determine if their pages are to be indexed. There are far more who want to get maximum exposure than there are who don't.
- The information is indexed and sorted. The index is massive !

When you enter a query Google searches its index and then applies a large number of criteria to try to ensure that the list of sites it provides is relevant to you. The more you use Google the more it is able to refine the searches for you.

Original internet searches

At the start of the Internet the searches were based on keywords. The search engine would apply some synonyms to the words you provided and try to match as many keywords as possible.

You could apply some logic to your search by using

AND : you want x AND y

OR : you want x OR y

You could also use "-" to indicate you did not want something.

Note these searches did not try to understand your input, they simply used keywords. Many people still search in this way. For example:

"caesar rubicon"

This might find your query. In this case it gets you close to an answer but in many cases you still have to refine your search.

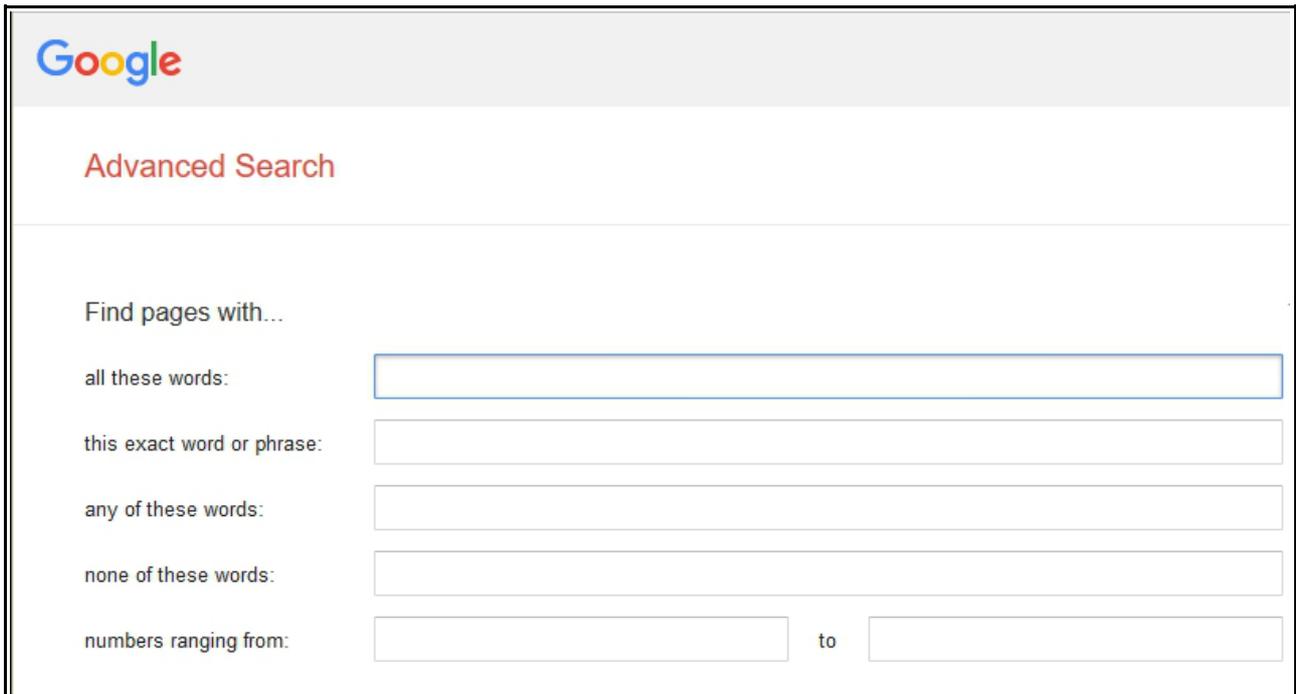
Current searches

You can now use English statements and in most cases Google will be able to interpret your request.

"when did caesar cross the rubicon"

The problem with many searches is that Google searches the whole web and, if you are looking for something in South Africa you might find you are flooded with irrelevant sites. To reduce this you need to consider using the Advanced Search facilities at the URL given at the top.

Advanced Search



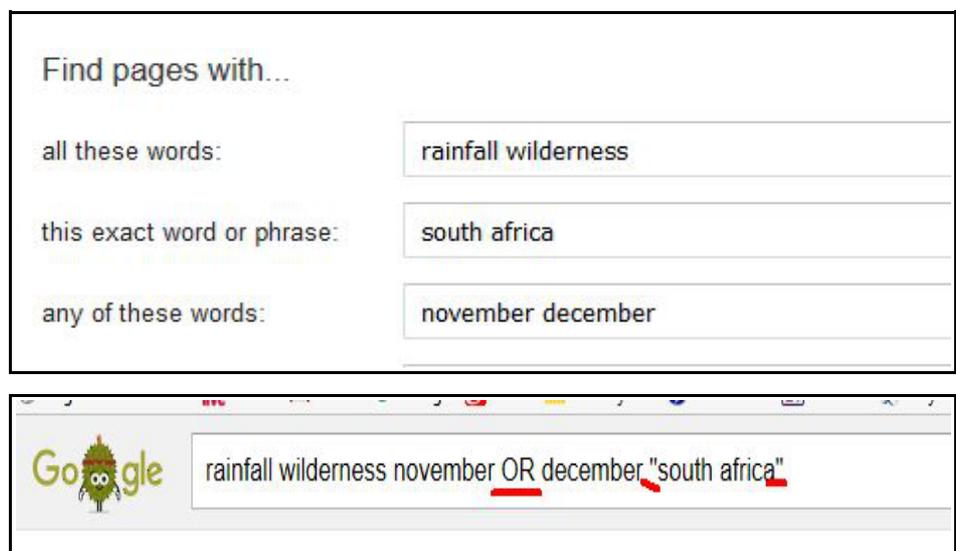
The screenshot shows the Google Advanced Search page. At the top left is the Google logo. Below it, the text "Advanced Search" is displayed in red. Underneath, the heading "Find pages with..." is followed by five input categories, each with an empty text box:

- all these words:
- this exact word or phrase:
- any of these words:
- none of these words:
- numbers ranging from: [] to []

You can refine your search by using the Advanced Search. After entering the required search criteria your input will be shown at the top of the search results page. Please look at it to understand how your input has been interpreted.

1. **"all these words"** : the search will find items containing all of the words in no specific order. It may also bring back some items where it eliminates some of the words it considers less significant. It will also look at synonyms of the words you use. Note that this is the default if you do not refine the search criteria as described below.
2. **"this exact word or phrase"** : You can also do this in the standard search by enclosing the word or phrase in double quotes. A single word in quotes will eliminate synonyms.
3. **"Any of these words"** : Each word does not have to appear. For example, a search green blue will show on the search results page as green OR blue.
4. **"none of these words"** : allows you to eliminate certain items
5. **"numbers ranging from"** : very useful when you are searching, for example, in a date range say 2001 to 2010

Example of an advanced search and how it will be displayed on the results page.



The screenshot shows the Google Advanced Search page with the following input fields filled:

- all these words: rainfall wilderness
- this exact word or phrase: south africa
- any of these words: november december

Below the input fields, the search query is displayed in the search bar: `rainfall wilderness november OR december "south africa"`. The words "OR" and "south africa" are underlined in red.

The list of web sites can then be further refined with the additional criteria. The one you should consider using more frequently is when the site was last updated.

Then narrow your results by...	
language:	any language
region:	any region
last update:	anytime
site or domain:	
terms appearing:	anywhere in the page
SafeSearch:	Show most relevant results
file type:	any format
usage rights:	not filtered by licence

1. **Language** : Self explanatory
2. **Region** : You can select the country in which the site is based
3. **"Last Update"** : Probably the one you should use most frequently. Since most of the web pages never disappear you can use this to eliminate old sites and get the most recently updated.
4. **"site or domain"** : you can search a specific site or restrict to specific sites such as ".co.za". This would then not retrieve a ".com" site located in South Africa.
5. **"file type"** : can be useful if you know, for example, you want a "pdf".

You can only learn to surf the net effectively by experimenting with your searches.