

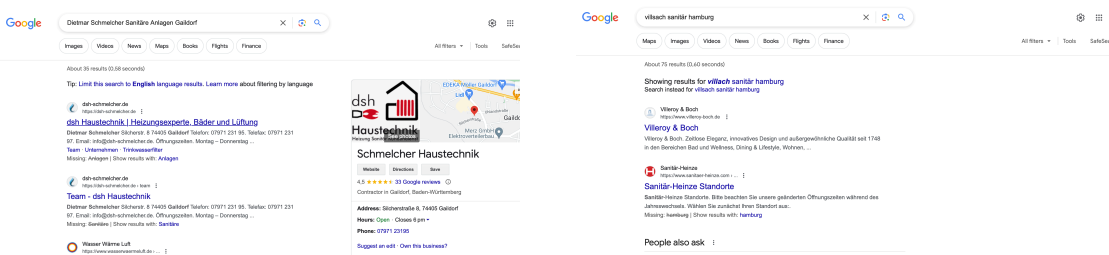
Input

ca. 80k-100k combinations of name + city of companies in Germany

e.g. :

Company_name	city
Dietmar Schmelcher Sanitäre Anlagen	Gaildorf
...	...

It serves as input to Google and optimally returns a result with a specific company-entry:



The first result would be the simple question:

- Does a specific company entry exist (like on the left) or not (like on the right)?

Data Scraping

For most, the specific company entry exists. From these company entries, the following information would be of interest:

- Name: *Schmelcher Haustechnik*
- Address (if it exists): *Silcherstraße 8, 74405 Gaildorf*
- Phone number (if it exists): *07971 23195*
- Avg. Customer Review stars: *4.5*
- Number of Customer Reviews: *33*

Review Scraping

Furthermore, it would be really cool to get the actual reviews (click on “33 reviews” below, then sorted by “newest”):

#	Stars	Time	Positive	Negative	Review (in original language)	Service
1	5	a month ago	Professionalism, Punctuality, Quality, Responsiveness, Value		Alles vorbildlich und topp,super Handwerker	General construction
2	5	4 months ago	Professionalism, Punctuality, Quality, Value		Sehr zufrieden!!!	
3	...					
4	...					

5	5	A year ago	Professionalism, Punctuality, Quality, Responsiveness, Value		Heizungserneuerung und Sanierungen im sanitären Bereich im Oktober 2022 von der Beratung bis zur Ausführung zu unserer vollsten Zufriedenheit. Immer wieder gerne	Bathroom remodeling, Installation, Repair water fixtures, Waterproofing, Plumbing services
6	...					

URL-Link Scraping

Finally, many company entries direct to the company website (clicking “Website”), but the URL behind it is a Google-one that redirects to the actual URL:

- URL from Google:
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiL-tTxpdqCAxULSvEDHaUSCqcQ_Bd6BQilARAE&url=https%3A%2F%2Fdsh-schmelcher.de%2F&usg=AOvVaw0CtmT4RdzCMJRbqOOrMjkt&opi=89978449
- Correct URL after redirection: <https://dsh-schmelcher.de>

If we could also get this correct URL after redirection, that would be great.

Output Format & Priority

Output would optimally be a dictionary format, such as:

```

example_dict
✓ 0.3s

{'Dietmar Schmelcher Sanitäre Anlagen Gaildorf': {'company_info': True,
'company_name': 'Schmelcher Haustechnik',
'address': '...',
'phone_number': '...',
'avg. customer reviews': 4.6,
'number of customer reviews': 33,
'reviews': [{'counter': 1,
'Stars': 5,
'Time': 'a month ago',
'Positive': '...',
'Negative': '...',
'Review': '...',
'Service': '...'},
{'counter': 2,
'Stars': 5,
'Time': '4 months ago',
'Positive': '...',
'Negative': '...',
'Review': '...',
'Service': '...'},
{'...'}],
'URL': 'https://dsh-schmelcher.de'},
'next_input_combination': {'...'}}

```