

✓ Beautiful Soap for web Scraping

```
import pandas as pd
import requests
from bs4 import BeautifulSoup

# Define user agent
headers = {'User-Agent': 'Mozilla/5.0 (Windows NT 6.3; Win 64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.162 Safa'}

# Make the request
response = requests.get('https://www.ambitionbox.com/list-of-companies?page=1', headers=headers)

# Check if the request was successful
if response.status_code == 200:
    # Parse the HTML content
    soup = BeautifulSoup(response.content, 'lxml')

    # Now you can work with the soup object
    # (e.g., find elements, extract data)
else:
    print("Failed to fetch webpage. Status code:", response.status_code)

print(soup.prettify())
```

```
<!DOCTYPE html>
<html data-n-head="%7B%22lang%22:%7B%22ssr%22:%22en%22%7D%7D" data-n-head-ssr="" lang="en">
<head>
<meta charset="utf-8"/>
<meta content="width=device-width,initial-scale=1,minimum-scale=1" name="viewport"/>
<meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<style>
@media only screen and (min-width:767px){.trp-img{width:400px!important;max-width:400px!important}}#ab-body{pointer-events:none}
</style>
<script>
window.dataLayer=window.dataLayer||[],window.gtag=window.gtag||function(){window.dataLayer.push(arguments)},gtag("js",new Date)
</script>
<title>
List of companies in India | AmbitionBox
</title>
<meta content="2024 AmbitionBox" data-n-head="ssr" name="copyright"/>
<meta content="1 day" data-n-head="ssr" name="revisit-after"/>
<meta content="AmbitionBox" data-n-head="ssr" name="application-name"/>
<meta content="EN" data-n-head="ssr" name="content-language"/>
<meta content="462822053404-hphug4pkahqljh2tc96g35at47o4isv2.apps.googleusercontent.com" data-n-head="ssr" name="google-signin-clie">
<meta content="712617688793459" data-n-head="ssr" property="fb:app_id"/>
<meta content="#ffffff" data-n-head="ssr" name="theme-color"/>
<meta content="#ffffff" data-n-head="ssr" name="msapplication-navbutton-color"/>
<meta content="default" data-n-head="ssr" name="apple-mobile-web-app-status-bar-style"/>
<meta content="100001438127755,100000444923785" data-n-head="ssr" property="fb:admins"/>
<meta content="AmbitionBox" data-n-head="ssr" property="og:site_name"/>
<meta content="website" data-n-head="ssr" property="og:type"/>
<meta content="summary_large_image" data-n-head="ssr" name="twitter:card"/>
<meta content="ambitionbox.com" data-n-head="ssr" name="twitter:domain"/>
<meta content="www.ambitionbox.com" data-n-head="ssr" name="author"/>
<meta content="document" data-n-head="ssr" name="resource-type"/>
<meta content="GLOBAL" data-n-head="ssr" name="distribution"/>
<meta content="general" data-n-head="ssr" name="rating"/>
<meta content="no-cache" data-n-head="ssr" name="pragma"/>
<meta content="Company:Reviews,Salaries,Benefits,Overview,Interview Questions" data-n-head="ssr" name="classification"/>
<meta content="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_192x192.png" data-n-head="ssr" name="msappli">
<meta content="yes" data-n-head="ssr" name="mobile-web-app-capable"/>
<meta content="AmbitionBox" data-n-head="ssr" name="apple-mobile-web-app-title"/>
<meta content="ANupXnJc0OGVOEFKYteoaBHgOPipYkAdjm3iRKvcdgI" data-n-head="ssr" name="google-site-verification"/>
<meta content="List of companies in India | AmbitionBox" data-hid="og:title" data-n-head="ssr" property="og:title"/>
<meta content="List of companies in India | AmbitionBox" data-hid="twitter:title" data-n-head="ssr" name="twitter:title"/>
<meta content="View list of companies in India with in-depth information on reviews, salaries, interviews, employee benefits, and m">
<meta content="View list of companies in India with in-depth information on reviews, salaries, interviews, employee benefits, and m">
<meta content="View list of companies in India with in-depth information on reviews, salaries, interviews, employee benefits, and m">
<meta content="https://static.ambitionbox.com/static/list-of-companies.jpg" data-hid="og:image" data-n-head="ssr" property="og:imag">
<meta content="https://static.ambitionbox.com/static/list-of-companies.jpg" data-hid="twitter:image" data-n-head="ssr" name="twitte">
<meta content="ALL" data-n-head="ssr" name="robots"/>
<meta content="https://www.ambitionbox.com/list-of-companies" data-n-head="ssr" property="og:url"/>
```

```

<meta content="https://www.ambitionbox.com/list-of-companies" data-n-head="ssr" property="twitter:url"/>
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_16x16.png" rel="icon" sizes="
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_32x32.png" rel="icon" sizes="
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_48x48.png" rel="icon" sizes="
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_96x96.png" rel="icon" sizes="
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_144x144.png" rel="icon" siz
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_192x192.png" rel="icon" siz
<link data-n-head="ssr" href="https://static.ambitionbox.com/alpha/img/supporting-images/browser_icon_192x192.png" rel="apple-touch
<link crossorigin="anonymous" data-n-head="ssr" href="https://static.ambitionbox.com" rel="preconnect"/>

```

```

# Extract the Company Name
for n in soup.find_all('h1'):
    print(n.text.strip())

```

List of companies in India

```

#Extract the Company Name
for i in soup.find_all('h2'):
    print(i.text.strip())

```

```

TCS
Accenture
Cognizant
Wipro
Capgemini
HDFC Bank
ICICI Bank
Infosys
HCLTech
Tech Mahindra
Genpact
Axis Bank
Teleperformance
Concentrix Corporation
Jio
Amazon
IBM
Reliance Retail
HDB Financial Services
Larsen & Toubro Limited
Companies by Industry
Companies by Locations
Companies by Type
Companies by Badges

```

```

#Extract the rating value
for k in soup.find_all('span',class_='companyCardWrapper__companyRatingValue'):
    print(k.text.strip())

```

```

3.8
4.0
3.9
3.8
3.9
4.0
4.0
3.8
3.7
3.7
3.9
3.9
3.6
3.9
4.0
4.1
4.1
3.9
4.0
4.1

```

```

#Extract the all card details
for j in soup.find_all('span',class_='companyCardWrapper__interLinking'):
    print(j.text.strip())

```

```

IT Services & Consulting | 1 Lakh+ Employees | Public | 56 years old | Mumbai +339 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 35 years old | Dublin +173 more
IT Services & Consulting | 1 Lakh+ Employees | Forbes Global 2000 | 30 years old | Teaneck, New Jersey. +156 more

```

IT Services & Consulting | 1 Lakh+ Employees | Public | 79 years old | Bangalore/Bengaluru +276 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 57 years old | Paris +140 more
Banking | 1 Lakh+ Employees | Public | 30 years old | Mumbai +1533 more
Banking | 1 Lakh+ Employees | Public | 30 years old | Mumbai +1276 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 43 years old | Bengaluru/Bangalore +179 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 33 years old | Noida +182 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 38 years old | Pune +262 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 27 years old | New York +107 more
Banking | 50k-1 Lakh Employees | Public | 31 years old | Mumbai +1295 more
BPO | 50k-1 Lakh Employees | Forbes Global 2000 | 46 years old | Paris +180 more
BPO | 50k-1 Lakh Employees | Public | 41 years old | Fremont +114 more
Telecom | 50k-1 Lakh Employees | Public | 17 years old | Navi Mumbai +1172 more
Internet | 1 Lakh+ Employees | Public | 30 years old | Seattle +438 more
IT Services & Consulting | 1 Lakh+ Employees | Public | 113 years old | Armonk +144 more
Retail | 1 Lakh+ Employees | 18 years old | Navi Mumbai +990 more
NBFC | 1 Lakh+ Employees | 17 years old | Ahmedabad +941 more
Engineering & Construction | 10k-50k Employees | Public | 86 years old | Mumbai +617 more