

# Python Web Scrapping Project

**WEB SCRAPING OF THE JOB OPENING DETAILS  
BY- SWAPNALI YASHWANT KHOPATKAR.**

## **Project Aim:**

**Web Scrapping of Job openings on the website Freshersworld.com.**

## **Introduction:**

- Web scraping is a technique to fetch data from websites. While surfing on the web, many websites don't allow the user to save data for personal use.
- Web scraping can help companies gather the correct contact information from their target market—including names, job titles, email addresses, and cellphone numbers. Then, they can reach out to these contacts and generate more leads and sales for their business.
- In this Project we are gathering the job opening details on the Freshersworld.com.

## **Objectives:**

- The details related to the job profile can be gathered easily by web scraping.
- We have collected four fields in our dataframe.
- Our dataframe consists of **Job titles, Job location, Company name and Qualifications.**

## Step 1:

First we import the required libraries for scraping the data. And declare the website to be scraped.

```
In [1]: import requests
        from bs4 import BeautifulSoup
        import pandas as pd
        URL = "https://www.freshersworld.com/premium-jobs?src=homeheader"
```

```
In [4]: import requests
r = requests.get(URL)
htmlcontent = r.content
htmlcontent
```

```
Out[4]: b'<!DOCTYPE html><html lang=en><head><!-- Google Tag Manager --><script>(function(w,d,s,l,i){w[l]=w[l]||[];w[l].push({'gtm.start':\nnew Date().getTime(),event:'gtm.js'});var f=d.getElementsByTagName(s)[0],\nj=d.createElement(s),dl=l!='dataLayer'?'\&l=' + l:'';j.async=true;j.src=\n'https://www.googletagmanager.com/gtm.js?id= '+id+dl;f.parentNode.insertBefore(j,f);\n})(window,document,'\nscrip\t','dataLayer','GTM-5537RST');</script><!-- End Google Tag Manager --><meta name="robots" content="index,follow" /><title>Premium jobs - Apply for premium jobs from top MNCs</title><meta name="viewport" content="width=device-width, initial-scale=1.0"><meta http-equiv="X-UA-Compatible" content="IE=edge,chrome=1"><meta name="HandheldFriendly" content="true"><meta name="description" content="Premium jobs are top most MNC\'s hiring for a salary ranging anywhere in between 4-12 lakhs per anum. These jobs are posted by companies exclusively on Freshersworld.com." /><meta http-equiv="Content-Type" content="text/html; charset=utf-8" /><link property="\stylesheet" media="all" rel="shortcut icon" type="image/x-icon" href="https://s3.amazonaws.com/fws3/favicon.ico" /><link property="\stylesheet" media="all" rel="icon" type="image/icon" href="https://s3.amazonaws.com/fws3/favicon.ico" /><link property="\stylesheet" media="all" rel="image_src" href="https://s3.amazonaws.com/fws3/static/common/images/freshersworld.png" /><link href="/manifest.json" rel="manifest"><!-- RSS FEED--><link property="\stylesheet" rel="alternate" type="application/rss+xml" title="Company Jobs - RSS" href="/feed" /><!-- RSS FEED--><!-- Open Graph data --><meta property="fb:admins" content="114456545282851" /><meta property="og:type" content="website" /><meta property="og:url" content="https://www.freshersworld.com/premium-jobs" /><meta property="og:title" content="Premium jobs - Apply for premium jobs from top MNCs" /><meta property="og:description" content="Premium jobs are top most MNC\'s hiring for a salary ranging anywhere in between 4-12 lakhs per anum. These jobs are posted by companies exclusively on Freshersworld.com." /><meta property="og:image" content="https://s3.amazonaws.com/fws3/static/common/images/freshersworld.png" /><meta property="og:site_name" content="freshersworld.com" /><meta property="og:update time" content="2023-10-03 11:28:34" /><!-- Open G
```

## Step 2:

Using BeautifulSoup library we will be able to extract the information from web page.

```
In [ ]: soup = BeautifulSoup(htmlcontent,"html.parser")
```

## Step 3:

In this step we have used find\_all() function to find the information in the given class.

```
In [ ]: soup = BeautifulSoup(htmlcontent,"html.parser")

In [ ]: Job_titles= soup.find_all("span",class_="wrap-title seo_title")
print(Job_titles)

In [ ]: list_Job_Titles=[]
for i in range(0,len(Job_titles)):
    list_Job_Titles.append(Job_titles[i].get_text())
print(list_Job_Titles)

In [ ]: company_names= soup.find_all(class_="latest-jobs-title font-16 margin-none inline-block company-name")
print(company_names)

In [ ]: list_company_names=[]
for i in range(0,len(company_names)):
    list_company_names.append(company_names[i].get_text())
print(list_company_names)

In [ ]: Location= soup.find_all(class_="bold_font")
print(Location)

In [ ]: list_Location=[]
for i in range(0,len(Location)):
    list_Location.append(Location[i].get_text())
print(list_Location)
list_Location.pop()
print(list_Location)

In [ ]: Qualifications= soup.find_all(class_="qualifications display-block modal-open pull-left job-details-span")
print(Qualifications)

In [ ]: list_Qualifications=[]
for i in range(0,len(Qualifications)):
    list_Qualifications.append(Qualifications[i].get_text())
print(list_Qualifications)
```

We have created all the field lists.

## Step 4:

Now using pandas, we have created dataframe of our fields.

```
In [12]: import pandas as pd
df=pd.DataFrame({"Job Titles":list_Job_Titles,"Company Names":list_company_names,"Location":list_Location,
                "Qualifications":list_Qualifications})
```

**Step 5:** Create a csv file of the dataframe in current working directory. And read it in python.

```
In [12]: import pandas as pd
df=pd.DataFrame({"Job Titles":list_Job_Titles,"Company Names":list_company_names,"Location":list_Location,
                "Qualifications":list_Qualifications})
```

```
In [13]: df.to_csv("JobReading data.csv",index=False)
a=pd.read_csv("JobReading data.csv")
a
```

# Output:

	Job Titles	Company Names	Location	Qualifications
0	C/C Developers Jobs Opening in LG Soft India a...	LG Soft India	Bangalore	BE/B.Tech(CSE, ECE, IT...),ME/M.Tech(CSE, IT, ...
1	Java Developer Appxtend Jobs Opening in Zycus ...	Zycus Infotech	Bangalore	BE/B.Tech(CSE, ECE, IT), MCA, ME/M.Tech(CSE, E...
2	Java Developers / Manual Testing Jobs Opening ...	LG Soft India	Mumbai	BCA, BE/B.Tech(CSE, ECE, IT...),BSc(CS, IT, En...
3	Software Engineer Jobs Opening in CARBYNETECH ...	CARBYNETECH	Pune	BCA, BE/B.Tech(CSE, IT), MCA, ME/M.Tech(CSE, IT)
4	Front-End Web Developer Jobs Opening in Calipe...	Caliper Business Solutions Pvt. Ltd.	Bangalore	BCA, BE/B.Tech
5	Technical Support Engineer Jobs Opening in ECl...	EClinicalWorks	Hyderabad	BCA, BE/B.Tech(CSE, IT), BSc(IT), MCA
6	Software Engineering Intern/Associate Software...	CelebAI Technologies Pvt Ltd	Bangalore	BE/B.Tech(CSE, IT), ME/M.Tech(CSE, IT)
7	Testing QA Jobs Opening in Micro Focus Softwar...	Micro Focus Software India Pvt. Ltd	Ahmedabad	BE/B.Tech(CSE)
8	Engineer Trainee Jobs Opening in Successive Di...	Successive Digital	Delhi	BCA, BE/B.Tech(CSE, IT), MCA, ME/M.Tech(CSE, IT)
9	Linux C Developer Jobs Opening in IQnext at Ba...	IQnext	Bangalore	BE/B.Tech, ME/M.Tech
10	Software Developer Jobs Opening in Welspun Tra...	Welspun Transformation Services Limited	Noida	BCA, BE/B.Tech, MCA, ME/M.Tech
11	Java Developer Jobs Opening in Agrosperity Tec...	Agrosperity Tech Solutions Private Limited KIVI	Bangalore	BCA, BE/B.Tech, MCA, ME/M.Tech
12	QA Engineer Intern Jobs Opening in Nimesa Tech...	Nimesa Technologies	Ahmedabad	BCA, BE/B.Tech, MCA, ME/M.Tech
13	Full Stack Engineer Jobs Opening in Humanitics...	Humanitics Dimensions Software Pvt. Ltd	Chennai	BE/B.Tech(CSE, EEE, ECE...),MCA
14	Trainee Engineer Jobs Opening in Toshiba at Ba...	Toshiba	Bangalore	BE/B.Tech(CSE, ECE, IT), MCA, ME/M.Tech(CSE, E...
15	Associate Software Engineer Jobs Opening in Br...	Briskminds Software Solutions Pvt. Ltd.	Bangalore	BCA, BE/B.Tech(CSE, EEE, ECE...),MCA
16	Software Engineer Intern Jobs Opening in Medio...	Mediology Software Pvt. Ltd.	Bangalore	BE/B.Tech(CSE, IT), MCA
17	Software Quality Engineer Jobs Opening in Hexa...	Hexaview Technologies	Jaipur	BCA, BE/B.Tech, MCA, ME/M.Tech
18	Junior Software Developer Jobs Opening in IESo...	IESoft Technologies Private Limited.	Delhi	BCA, BE/B.Tech(CSE), ME/M.Tech(CSE)
19	Software Engineer Trainee Jobs Opening in Moun...	MountBlue Technologies Pvt Ltd.	Gurgaon	BE/B.Tech, MCA, ME/M.Tech
20	Python Developer Jobs Opening in Odisoft Techn...	Odisoft Technology Private Limited	Noida	BCA, BE/B.Tech, MCA, ME/M.Tech
21	React JS Developer Jobs Opening in Maitretech ...	Maitretech	Noida	BE/B.Tech(CSE, IT), MCA, ME/M.Tech(CSE, IT)

## **Conclusion:**

web scraping is a powerful method for obtaining information from websites and has a wide range of uses in business, research, and other industries.

Most of this data is unstructured data in an HTML format which is then converted into structured data in a spreadsheet or a database so that it can be used in various applications.

As in our project we have converted the dataframe into csv file for further use.

**THANK YOU !**