

# Web Scraping and Social Media Scraping Project

Designed by : Zhoashuai Wang 436661  
Weida Pan 369868

## Description

Imagine you are a League of Legends player. And you would like to know the win rate about the champion you are playing VS other champions, then you could choose your counter champion to win more games.

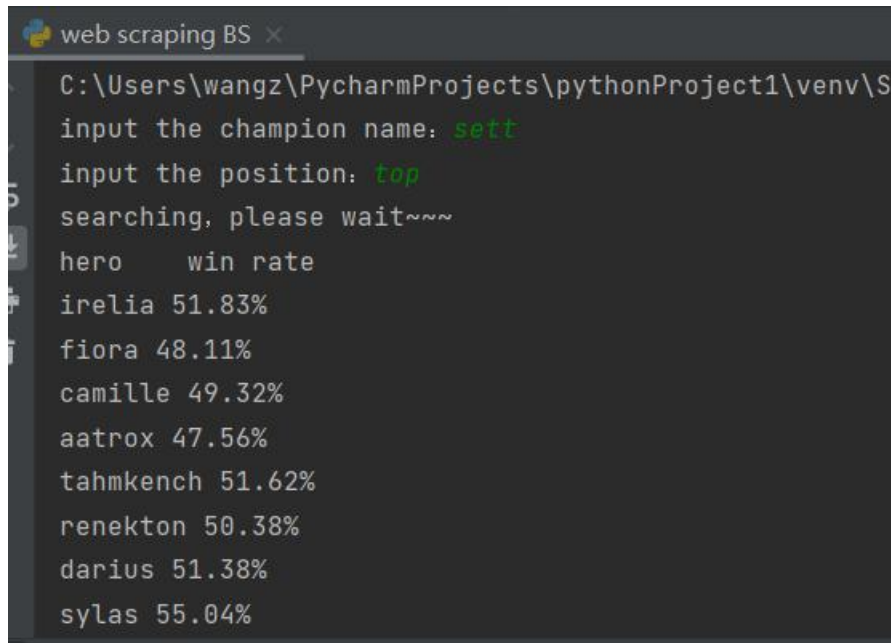
So this project is going to crawl the famous LOL data website: <https://euw.op.gg>

Our goal is : Giving one champion and the position, the output will be this champion's match-up winrate.

For all 3 parts, we will use champion 'sett', position 'top' as example:

### 1. For beautiful soup part

Just run it, and input champion name 'sett', position 'top' as example, result will be like that :



```
web scraping BS x
C:\Users\wangz\PycharmProjects\pythonProject1\venv\S
input the champion name: sett
input the position: top
searching, please wait~~~
hero    win rate
irelia  51.83%
fiora   48.11%
camille 49.32%
aatrox  47.56%
tahmkench 51.62%
renekton 50.38%
darius  51.38%
sylas   55.04%
```

### 2. For scrapy part

Our spider name is gamelol\_spider, so we used the command to run :

**scrapy crawl gamelol\_spider -o gamelol.csv**

And we open the gamelol.csv, the result will be like that:

	A	B	C	D	E	F	G	H	I
1	hero_name	win_rate							
2	irelia	0.5183							
3	fiora	0.4811							
4	camille	0.4932							
5	aatrox	0.4756							
6	tahmkench	0.5162							
7	renekton	0.5038							
8	darius	0.5138							
9	sylas	0.5504							
10	malphite	0.4935							
11	jayce	0.5481							
12	ornn	0.4836							
13	jax	0.4706							
14	mordekaiser	0.4722							
15	vayne	0.4376							
16	riven	0.5135							
17	gnar	0.4685							
18	garen	0.4909							
19	kennen	0.4919							
20	trundle	0.5069							
21	shen	0.5322							
22	volibear	0.4807							
23	wukong	0.4993							
24	teemo	0.4868							
25	akali	0.5022							
26	ryze	0.5215							
27	gwen	0.5654							
28	illaoi	0.4763							
29	lucian	0.5227							
30	tryndamere	0.4751							

PS: But in here, if we want to change champion name and position, we have to change the start\_urls to: <https://euw.op.gg/champion/{hero}/statistics/{position}/matchup>  
 Replace {hero} and {position} to your hero name and position.

### 3. For selenium part

Just run the code, it will open the website and giving result like that:

Open Files

F:\mypyppython\_files

venv

seleproject.py

web scraping BS.py

seleproject

```
C:\Users\wangz\PycharmProjects\python3\seleproject>python seleproject.py
input the hero:se
input the position:top
E:\mypyppython_files\seleproject.py:19: DeprecationWarning: The arguments to webdriver.Chrome() have changed. The
  browser = webdriver.Chrome(chrome_
hero win rate
Irelia 51.83%
Fiora 48.11%
Camille 49.32%
Aatrox 47.56%
Tahm Kench 51.62%
Renekton 50.38%
Darius 51.38%
Sylas 55.84%
```

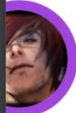
Sett Counter (Top, Season 11)

euw.op.gg/champion/sett/statistics/top/matchup#


Chrome is being controlled by automated test software.

CHAMPION

MATCH UP

Sett

VS

Gragas

55.81%

44.19%

Matchup Stats		Matchup Stats	
50.50%	Lane Kill Rate	49.50%	
74 : 1	KDA	2.15 : 1	
6.0%	Kill Participation	36.6%	
3,279	Damage Dealt to Champions	17,299	
3'51"	First Tower Kill	16'30"	

Waiting for lilqkd.net...

Search a Champion

#	Win Rate	Games Played
1	Irelia	51.83% 9,746 4,634
2	Fiora	48.11% 5,606 2,665
3	Camille	49.32% 5,126 2,435
4	Aatrox	47.56% 4,406 2,092
5	Tahm Kench	51.62% 4,096 1,939
6	Renekton	50.38% 3,926 1,866
7	Darius	51.38% 3,696 1,742