

```
pip install lxml html5lib beautifulsoup4
```

```
import pandas as pd
```

```
# Webpage url
```

```
url = 'https://en.wikipedia.org/wiki/History_of_Python'
```

```
# Extract tables
```

```
dfs = pd.read_html(url)
```

```
# Get first table
```

```
df = dfs[1]
```

```
# Extract columns
```

```
df2 = df[['Version', 'Release date']]
```

```
print(df2)
```

OUTPUT:

Version \	
0	0.9
1	1.0
2	1.1
3	1.2
4	1.3
5	1.4
6	1.5
7	1.6
8	2.0
9	2.1
10	2.2
11	2.3
12	2.4
13	2.5
14	2.6
15	2.7
16	3.0

17	3.1
18	3.2
19	3.3
20	3.4
21	3.5
22	3.6
23	3.7
24	3.8
25	3.9
26	3.10
27	3.11
28	3.12
29	3.13

	Release date
0	1991-02-20[2]
1	1994-01-26[2]
2	1994-10-11[2]
3	1995-04-13[2]
4	1995-10-13[2]
5	1996-10-25[2]
6	1998-01-03[2]
7	2000-09-05[43]
8	2000-10-16[45]
9	2001-04-15[46]
10	2001-12-21[47]
11	2003-06-29[48]
12	2004-11-30[49]
13	2006-09-19[50]
14	2008-10-01[27]
15	2010-07-03[32]
16	2008-12-03[27]
17	2009-06-27[52]
18	2011-02-20[54]
19	2012-09-29[55]
20	2014-03-16[56]
21	2015-09-13[58]
22	2016-12-23[60]
23	2018-06-27[61]
24	2019-10-14[62]
25	2020-10-05[63]
26	2021-10-04[65]
27	2022-10-24[66]
28	2023-10-02[67]
29	2024-10-01[68]

```
# Write to excel  
df2.to_excel('python.xlsx')
```

A	B	C	D	E	F	G
	Version	Release date				
0	0.9	1991-02-20[2]				
1	1.0	1994-01-26[2]				
2	1.1	1994-10-11[2]				
3	1.2	1995-04-13[2]				
4	1.3	1995-10-13[2]				
5	1.4	1996-10-25[2]				
6	1.5	1998-01-03[2]				
7	1.6	2000-09-05[43]				
8	2.0	2000-10-16[45]				
9	2.1	2001-04-15[46]				
10	2.2	2001-12-21[47]				
11	2.3	2003-06-29[48]				
12	2.4	2004-11-30[49]				
13	2.5	2006-09-19[50]				
14	2.6	2008-10-01[27]				
15	2.7	2010-07-03[32]				
16	3.0	2008-12-03[27]				
17	3.1	2009-06-27[52]				
18	3.2	2011-02-20[54]				
19	3.3	2012-09-29[55]				
20	3.4	2014-03-16[56]				
21	3.5	2015-09-13[58]				
22	3.6	2016-12-23[60]				
23	3.7	2018-06-27[61]				
24	3.8	2019-10-14[62]				
25	3.9	2020-10-05[63]				
26	3.10	2021-10-04[65]				
27	3.11	2022-10-24[66]				
28	3.12	2023-10-02[67]				
29	3.13	2024-10-01[68]				