

```
In [ ]: import csv
import requests
from bs4 import BeautifulSoup
from urllib.request import urlopen
import numpy as np
import pandas as pd
```

```
In [ ]: # loop that returns data in a list.
x = []
for i in range(3):
    res = requests.get(f'https://secure.runescape.com/m=hiscore/ranking?category_type=0&table=0&time_filter=0&date=1581438205530&page={i}')
    soup = BeautifulSoup(res.content, 'lxml')
    table = soup.find_all('table')[1]
    df = pd.read_html(str(table))
    x.append(eval(df[0].to_json(orient='records'))))
print(x)
```

```
In [ ]: type(x)
```

```
In [ ]: # Best loop for right now. Its in a dataframe and I changed the column names in R.
```

```
import pandas as pd
dfs = []
for page in range(3):
    url = f'https://secure.runescape.com/m=hiscore/ranking?category_type=0&table=0&time_filter=0&date=1581438205530&page={page}'
    dflist = pd.read_html(url)
    for df in dflist:
        if df.size > 0:
            dfs.append(df)
            ds = pd.concat(dfs)
            print(ds)
```

```
In [ ]: type(ds)
```

```
In [ ]: len(ds)
```

```
In [ ]: # Exporting the data set.
dssf = pd.DataFrame((ds))
dssf.to_csv('Runescape_leaderboards')
```

```
In [ ]: # bad loop: Runs multiple pages but only grabs last page of info

aa = []
page_soup = BeautifulSoup(page_repsone.text, 'html.parser')
for page_number in range(1, 4):
    page_repsone = requests.get(f'https://secure.runescape.com/m=hiscore/ranking?category_type=0&table=0&time_filter=0&date=1581438205530&page={page_number}')
    page_html = page_soup.prettify()
```

```
In [ ]: # Single page web scrape
art = requests.get('https://secure.runescape.com/m=hiscore/ranking#_ga=2.186985771.2064614990.1581437905-26754626.1581437905')
```

```
In [ ]: print(art.text[0:500])
```

```
In [ ]: soup = BeautifulSoup(art.text, 'html.parser')
```

```
In [ ]: ss = soup.find_all('td', {'class':'col3 align'})
qq = soup.find_all('td', {'class':'col4 align'})
```

```
In [ ]: # loop for rank
a = []
for tr in soup.find_all('td', {'class':'col1 align'}):
    rank = tr.a.text.strip()

    a.append((rank))
    print(rank)
```

```
In [ ]: # loop for name
b = []
for tr in soup.find_all('td', {'class':'col2'}):
    name = tr.text[2:7].strip()

    print(name)

    b.append((name))
```

```
In [ ]: # loop for level
c = []
for tr in soup.find_all('td', {'class':'col3 align'}):
    level = ss[0].a.text.strip()

    print(level)

    c.append((level))
```

```
In [ ]: # loop for xp
d = []
for tr in soup.find_all('td', {'class': 'col4 align'}):
    xp = tr.a.text.strip()

    d.append((xp))
print(xp)
```

```
In [ ]: # loop for odd rows

records = []
for tr in soup.find_all('tr', {'class': 'oddRow'}):
    name = tr.find_all('a')[1].text[2:7].strip()
    rank = tr.find_all('td')[0].find_all('a')[0].text.strip()
    xp = tr.find_all('a')[3].text[1:14]
    newlevel = tr.find_all('a')[2].text[1:6]

    records.append((name, rank, xp, newlevel))

print(name, rank, xp, newlevel)
```

```
In [ ]: # Joining each column
df1 = pd.DataFrame(a, columns=['Rank'])
df2 = pd.DataFrame(b, columns=['Player'])
df3 = pd.DataFrame(c, columns=['Level'])
df4 = pd.DataFrame(d, columns=['Xp'])
```

```
In [ ]: # Renaming
df1['Player'] = df2
df1['Level'] = df3
df1['Xp'] = df4
```

```
In [ ]: df1
```

```
In [ ]:
```