

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

```

In [ ]: art = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?
cat=1')
soup = BeautifulSoup(art.text, 'html.parser')

# Data pages 1-41 first pages on each cateogery in the catalog

import pandas as pd
dfs = []
for page in range(0,42):
    url = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat={page}'
    dflist = pd.read_html(url)
    for df in dflist:
        if df.size > 0:
            dfs.append(df)
            ds = pd.concat(dfs)
            print(ds)

# Cooking page 2
aart = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?cat=6&page=2')
ssoup = BeautifulSoup(aart.text, 'html.parser')

a = []
for roop in ssoup.find_all('span')[9:]:
    Item = roop.text
    print(Item)
    a.append((Item))

b = []
for make in ssoup.find_all('tr')[1:]:
    Price = make.find_all('td')[2].text
    print(Price)

    b.append((Price))

c = []
for rake in ssoup.find_all('tr')[1:]:
    Change = rake.find_all('a')[1].text
    print(Change)
    c.append((Change))

df1 = pd.DataFrame(a, columns=['Item'])
df2 = pd.DataFrame(b, columns=['Price'])
df3 = pd.DataFrame(c, columns=['Change'])

df1['Price'] = df2
df1['Change'] = df3

# Costumes 2-4
dfss = []
for page in range(2, 5):
    url1 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=7&page={page}'

```

```
dflistt = pd.read_html(urll)
for dff in dflistt:
    if dff.size > 0:
        dfss.append(dff)
        dss = pd.concat(dfss)
        print(dss)

# Crafting materials page 2
req = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?cat=8&page=2')
craftmat = BeautifulSoup(req.text, 'html.parser')

aa = []
for k in craftmat.find_all('span')[9:]:
    Item = k.text
    print(Item)
    aa.append((Item))

bb = []
for l in craftmat.find_all('tr')[1:]:
    Price = l.find_all('td')[2].text
    print(Price)

    bb.append((Price))

cc = []
for m in craftmat.find_all('tr')[1:]:
    Change = m.find_all('a')[1].text
    print(Change)
    cc.append((Change))

df11 = pd.DataFrame(aa, columns=['Item'])
df22 = pd.DataFrame(bb, columns=['Price'])
df33 = pd.DataFrame(cc, columns=['Change'])

df11['Price'] = df22
df11['Change'] = df33

df11

# food and drink

dfsss = []
for page in range(2, 5):
    urlll = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=12&page={page}'
    dflisttt = pd.read_html(urlll)
    for dfff in dflisttt:
        if dfff.size > 0:
            dfsss.append(dfff)
            dsss = pd.concat(dfsss)
            print(dsss)

# Herb page 2
reqq = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?cat=13&page=2')
```

```
herbmat = BeautifulSoup(reqq.text, 'html.parser')

aaa = []
for kk in herbmat.find_all('span')[9:]:
    Item = kk.text
    print(Item)
    aaa.append((Item))

bbb = []
for ll in herbmat.find_all('tr')[1:]:
    Price = ll.find_all('td')[2].text
    print(Price)

    bbb.append((Price))

ccc = []
for mm in herbmat.find_all('tr')[1:]:
    Change = mm.find_all('a')[1].text
    print(Change)
    ccc.append((Change))

df111 = pd.DataFrame(aaa, columns=['Item'])
df222 = pd.DataFrame(bbb, columns=['Price'])
df333 = pd.DataFrame(ccc, columns=['Change'])

df111['Price'] = df222
df111['Change'] = df333

# Jewel
reqqq = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?cat=16&page=2')
jewel = BeautifulSoup(reqqq.text, 'html.parser')

aaaa = []
for kkk in jewel.find_all('span')[9:]:
    Item = kkk.text
    print(Item)
    aaaa.append((Item))

bbbb = []
for lll in jewel.find_all('tr')[1:]:
    Price = lll.find_all('td')[2].text
    print(Price)

    bbbb.append((Price))

cccc = []
for mmm in jewel.find_all('tr')[1:]:
    Change = mmm.find_all('a')[1].text
    print(Change)
    cccc.append((Change))

df1111 = pd.DataFrame(aaaa, columns=['Item'])
df2222 = pd.DataFrame(bbbb, columns=['Price'])
df3333 = pd.DataFrame(cccc, columns=['Change'])
```

```

df11111['Price'] = df2222
df11111['Change'] = df3333

#Mage
reqqqq = requests.get('http://services.runescape.com/m=itemdb_rs/catalogue?cat=17&page=2')
mage = BeautifulSoup(reqqqq.text, 'html.parser')

aaaaa = []
for kkkk in mage.find_all('span')[9:]:
    Item = kkkk.text
    print(Item)
    aaaaa.append((Item))

bbbbbb = []
for llll in mage.find_all('tr')[1:]:
    Price = llll.find_all('td')[2].text
    print(Price)

    bbbbbb.append((Price))

ccccc = []
for mmmm in mage.find_all('tr')[1:]:
    Change = mmmm.find_all('a')[1].text
    print(Change)
    ccccc.append((Change))

df11111 = pd.DataFrame(aaaaa, columns=['Item'])
df22222 = pd.DataFrame(bbbbbb, columns=['Price'])
df33333 = pd.DataFrame(ccccc, columns=['Change'])

df11111['Price'] = df22222
df11111['Change'] = df33333

df11111

# Melee high level
dfssss = []
for page in range(2, 4):
    url1111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=21&page={page}'
    df11111 = pd.read_html(url1111)
    for dffff in df11111:
        if dffff.size > 0:
            dfssss.append(dffff)
    dssss = pd.concat(dfssss)
    print(dssss)

# Melee low level
dfsssss = []
for page in range(2, 4):
    url11111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=19&page={page}'
    df111111 = pd.read_html(url11111)

```

```

for dfffff in dflisttttt:
    if dfffff.size > 0:
        dfsssss.append(dfffff)
        dsssss = pd.concat(dfsssss)
        print(dsssss)

# Melee arm mid
dfsssssss = []
for page in range(2, 4):
    url111111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=20&page={page}'
    dflistttttt = pd.read_html(url111111)
    for dffffff in dflistttttt:
        if dffffff.size > 0:
            dfsssss.append(dffffff)
            dsssss = pd.concat(dfsssss)
            print(dsssss)

# Melee wep high
dfsssssssss = []
for page in range(2,3):
    url11111111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=24&page={page}'
    dflistttttttt = pd.read_html(url11111111)
    for dfffffff in dflistttttttt:
        if dfffffff.size > 0:
            dfsssssssss.append(dfffffff)
            dsssssssss = pd.concat(dfsssssssss)
            print(dsssssssss)

# Melee wep low
dfsssssssssss = []
for page in range(2,4):
    url111111111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=22&page={page}'
    dflistttttttttt = pd.read_html(url111111111)
    for dfffffff in dflistttttttttt:
        if dfffffff.size > 0:
            dfsssssssssss.append(dfffffff)
            dsssssssssss = pd.concat(dfsssssssssss)
            print(dsssssssssss)

# Melee wep mid
dfsssssssssss = []
for page in range(2,3):
    url1111111111 = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=23&page={page}'
    dflistttttttttt = pd.read_html(url1111111111)
    for dfffffff in dflistttttttttt:
        if dfffffff.size > 0:
            dfsssssssssss.append(dfffffff)
            dsssssssssss = pd.concat(dfsssssssssss)
            print(dsssssssssss)

# Misc 2-9
dfsssssssssssss = []
for page in range(2,10):

```

```

urllllllllllll = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=0&page={page}'
dfllstttttttttt = pd.read_html(urllllllllllll)
for dffffffffff in dfllstttttttttt:
    if dffffffffff.size > 0:
        dfsssssssssss.append(dffffffffff)
        dsssssssssss = pd.concat(dfsssssssssss)
        print(dsssssssssss)

# potions 2-5
emp = []
for page in range(2,6):
    link = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=26&page={page}'
    listt = pd.read_html(link)
    for dframe in listt:
        if dframe.size > 0:
            emp.append(dframe)
            dset = pd.concat(emp)
            print(dset)

# range arm
empp = []
for page in range(2,3):
    linkk = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=29&page={page}'
    listtt = pd.read_html(linkk)
    for dframee in listtt:
        if dframee.size > 0:
            empp.append(dframee)
            dsett = pd.concat(empp)
            print(dsett)

# Salvage
emppp = []
for page in range(2,3):
    linkkk = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=39&page={page}'
    listttt = pd.read_html(linkkk)
    for dframeee in listttt:
        if dframeee.size > 0:
            emppp.append(dframeee)
            dsettt = pd.concat(emppp)
            print(dsettt)

#tools and con 2-3
emppppp = []
for page in range(2,4):
    linkkkk = f'http://services.runescape.com/m=itemdb_rs/catalogue?cat=35&page={page}'
    listtttt = pd.read_html(linkkkk)
    for dframeeee in listtttt:
        if dframeeee.size > 0:
            emppppp.append(dframeeee)
            dsettttt = pd.concat(emppppp)
            print(dsettttt)

```

```
dataseta = ds.append(df1)
datasetb = dataseta.append(dss)
datasetc = datasetb.append(df11)
datasetd = datasetc.append(dfsss)
datasete = datasetd.append(df111)
datasetf = datasete.append(df1111)
datasetg = datasetf.append(df11111)
dataseth = datasetg.append(dfsssss)
dataseti = dataseth.append(dfsssss)
datasetj = dataseti.append(dfsssssss)
datasetk = datasetj.append(dfsssssssss)
datasetl = datasetk.append(dfsssssssssss)
datasetm = datasetl.append(dfssssssssssss)
datasetn = datasetm.append(dset)
dataseto = datasetn.append(dsett)
datasetp = dataseto.append(dsettt)
datasetq = datasetp.append(dsettttt)
```

```
In [ ]: datasetq.head()
```

```
In [ ]: datasetq.applymap(str)
```

```
In [ ]: datasetq.Price
```

```
In [ ]:
```

```
In [ ]:
```



```
In [ ]: for item in datasetq.Price:

    try:

        right = item.split(".")[1]

        left = item.split(".")[0]

        numsOnRight = len(right)-1

        if "k" in item:

            numOfZeroes = 3 - numsOnRight

        elif "m" in item:

            numOfZeroes = 6 - numsOnRight

        if numOfZeroes < 0:

            item = left + right[0:numsOnRight+numOfZeroes] + "." + right[numOfZeroes-1:numsOnRight]

        else:

            item = int(left + right[0:numsOnRight] + str(0)*numOfZeroes)

    except:

        if "k" in item:

            item = item[0:-1] + str(0)*3

        elif "m" in item:

            item = item[0:-1] + str(0)*6

    print(item)
```

```
In [ ]: datasetq
```

```
In [ ]: for item in datasetq.Price:

    item = item.replace('.', '')

    if "k" in item:

        item = int(item.replace("k", "00"))

    elif "m" in item:

        item = int(item.replace("m", "00000"))
```

In [ ]: