

```
In [2]: !pip install bs4
!pip install requests
```

```
Requirement already satisfied: bs4 in c:\users\user\anaconda3\lib\site-packages (0.0.1)
Requirement already satisfied: beautifulsoup4 in c:\users\user\anaconda3\lib\site-packages (from bs4) (4.12.2)
Requirement already satisfied: soupsieve>1.2 in c:\users\user\anaconda3\lib\site-packages (from beautifulsoup4->bs4) (2.4)
Requirement already satisfied: requests in c:\users\user\anaconda3\lib\site-packages (2.31.0)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\user\anaconda3\lib\site-packages (from requests) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\user\anaconda3\lib\site-packages (from requests) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\user\anaconda3\lib\site-packages (from requests) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\user\anaconda3\lib\site-packages (from requests) (2023.7.22)
```

```
In [3]: from bs4 import BeautifulSoup
import requests
```

```
In [5]: page = requests.get('https://en.wikipedia.org/wiki/Main_Page')
page
```

```
Out[5]: <Response [200]>
```

```
In [6]: soup = BeautifulSoup(page.content)
soup
```

```
e,"wgCentralAuthMobileDomain":false,"wgULSPosition":"interlanguage","wgULS
isCompactLinksEnabled":true,"wgULSisLanguageSelectorEmpty":false,"wgWikiba
seItemId":"Q5296","wgCheckUserClientHintsHeadersJsApi":["architecture","bi
tness","brands","fullVersionList","mobile","model","platform","platformVer
sion"],"GEHomepageSuggestedEditsEnableTopics":true,"wgGETopicsMatchModeEna
bled":false,"wgGESTructuredTaskRejectionReasonTextInputEnabled":false,"wgG
ELevelingUpEnabledForUser":false};RLSTATE={"skins.vector.user.styles":"rea
dy","ext.globalCssJs.user.styles":"ready","site.styles":"ready","user.styl
es":"ready","skins.vector.user":"ready","ext.globalCssJs.user":"ready","us
er":"ready","user.options":"loading","codex-search-styles":"ready","skins.
vector.styles":"ready",
"skins.vector.icons":"ready","ext.visualEditor.desktopArticleTarget.noscrip
t":"ready","ext.wikimediaBadges":"ready","ext.uls.interlanguage":"read
y"};RLPAGEMODULES=["site","mediawiki.page.ready","skins.vector.js","ext.vi
sualEditor.desktopArticleTarget.init","ext.visualEditor.targetLoader","ex
t.eventLogging","ext.wikimediaEvents","ext.navigationTiming","ext.cx.event
logging.campaigns","ext.centralNotice.geoIP","ext.centralNotice.startUp
","ext.gadget.ReferenceTooltips","ext.gadget.charinsert","ext.gadget.extr
a-toolbar-buttons","ext.gadget.switcher","ext.centralauth.centralautologi
n".mmv.head".mmv.bootstrap.autostart".ext.nonuns".ext.echo.centralaut
```

```
In [13]: first_titles = soup.find('span',class_="mw-headline")
first_titles
```

```
Out[13]: <span class="mw-headline" id="Welcome_to_Wikipedia">Welcome to <a href="/wiki/Wikipedia" title="Wikipedia">Wikipedia</a></span>
```

```
In [14]: first_titles.text
```

```
Out[14]: 'Welcome to Wikipedia'
```

```
In [16]: titles = []
for i in soup.find_all('span',class_="mw-headline"):
    titles.append(i.text)

titles
```

```
Out[16]: ['Welcome to Wikipedia',
'From today's featured article',
'Did you know\xa0...',
'In the news',
'On this day',
'Today's featured picture',
'Other areas of Wikipedia',
'Wikipedia's sister projects',
'Wikipedia languages']
```

```
In [17]: print(len (titles))
```

9

```
In [19]: import pandas as pd
df = pd.DataFrame({'Titels':titles})
df
```

```
Out[19]:
```

	Titels
0	Welcome to Wikipedia
1	From today's featured article
2	Did you know ...
3	In the news
4	On this day
5	Today's featured picture
6	Other areas of Wikipedia
7	Wikipedia's sister projects
8	Wikipedia languages

