

In [1]:

```
!pip install bs4
!pip install requests
```

```
Requirement already satisfied: bs4 in c:\users\rohan\anaconda3\lib\site-pa
ckages (0.0.1)
Requirement already satisfied: beautifulsoup4 in c:\users\rohan\anaconda3
\lib\site-packages (from bs4) (4.11.1)
Requirement already satisfied: soupsieve>1.2 in c:\users\rohan\anaconda3\l
ib\site-packages (from beautifulsoup4->bs4) (2.3.1)
Requirement already satisfied: requests in c:\users\rohan\anaconda3\lib\si
te-packages (2.27.1)
Requirement already satisfied: charset-normalizer~=2.0.0 in c:\users\rohan
\anaconda3\lib\site-packages (from requests) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\rohan\anaconda3\li
b\site-packages (from requests) (3.3)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\rohan\ana
conda3\lib\site-packages (from requests) (1.26.9)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\rohan\anacon
da3\lib\site-packages (from requests) (2021.10.8)
```

In [6]:

```
from bs4 import BeautifulSoup
import requests
```

In [7]:

```
page = requests.get('https://presidentofindia.nic.in/former-presidents.htm')
```

In [8]:

```
page
```

Out[8]:

```
<Response [200]>
```

In [9]:

```
soup = BeautifulSoup(page.content)
soup
```

Out[9]:

```
<!DOCTYPE html>
<html lang="en" xml:lang="en" xmlns="http://www.w3.org/1999/xhtml">
<head id="Head1"><title>
    Former Presidents - The President of India
</title><meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<!--<meta http-equiv="Content-Style-Type" content="text/css" /><meta http-equiv="Content-Script-Type" content="type" />-->
<meta content="telephone=no" name="format-detection"/><meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<!-- Start Favicon -->
<link href="favicon.ico" rel="shortcut icon" type="image/x-icon"/><link href="js/panorama_viewer.css" rel="stylesheet" type="text/css"/>
<!-- Start Javascript Library -->
<script language="JavaScript" type="text/javascript">
    if (top != self) top.location.replace(location);
    version = "3.6.0";
    nrettvPhoto = "3.1.6":
</script>
```

In [10]:

```
Name = soup.find('div', class_="presidentListing")
Name
```

Out[10]:

```
<div class="presidentListing">
<h3>Shri Ram Nath Kovind (birth - 1945)</h3>
<p><span class="terms">Term of Office:</span> 25 July, 2017 to 25 July, 2022 </p>
<p><a href="https://ramnathkovind.nic.in" target="_blank">https://ramnathkovind.nic.in</a></p>
</div>
```

In [18]:

```
Name.text.split('\n')
```

Out[18]:

```
['',
 'Shri Ram Nath Kovind (birth - 1945)',
 'Term of Office: 25 July, 2017 to 25 July, 2022 ',
 'https://ramnathkovind.nic.in',
 '']
```

In [35]:

```
name=[]  
for i in soup.find_all('div',class_="presidentListing"):  
    name.append(i.text.split('\n')[1])  
  
name
```

Out[35]:

```
['Shri Ram Nath Kovind (birth - 1945)',  
'Shri Pranab Mukherjee (1935-2020)',  
'Smt Pratibha Devisingh Patil (birth - 1934)',  
'DR. A.P.J. Abdul Kalam (1931-2015)',  
'Shri K. R. Narayanan (1920 - 2005)',  
'Dr Shankar Dayal Sharma (1918-1999)',  
'Shri R Venkataraman (1910-2009)',  
'Giani Zail Singh (1916-1994)',  
'Shri Neelam Sanjiva Reddy (1913-1996)',  
'Dr. Fakhruddin Ali Ahmed (1905-1977)',  
'Shri Varahagiri Venkata Giri (1894-1980)',  
'Dr. Zakir Husain (1897-1969)',  
'Dr. Sarvepalli Radhakrishnan (1888-1975)',  
'Dr. Rajendra Prasad (1884-1963) ']
```

In [40]:

```
term_of_office=[]  
for i in soup.find_all('div',class_="presidentListing"):  
    term_of_office.append(i.text.split('\n')[2])  
  
term_of_office
```

Out[40]:

```
['Term of Office: 25 July, 2017 to 25 July, 2022 ',  
'Term of Office: 25 July, 2012 to 25 July, 2017 ',  
'Term of Office: 25 July, 2007 to 25 July, 2012 ',  
'Term of Office: 25 July, 2002 to 25 July, 2007 ',  
'Term of Office: 25 July, 1997 to 25 July, 2002 ',  
'Term of Office: 25 July, 1992 to 25 July, 1997 ',  
'Term of Office: 25 July, 1987 to 25 July, 1992 ',  
'Term of Office: 25 July, 1982 to 25 July, 1987 ',  
'Term of Office: 25 July, 1977 to 25 July, 1982 ',  
'Term of Office: 24 August, 1974 to 11 February, 1977',  
'Term of Office: 3 May, 1969 to 20 July, 1969 and 24 August, 1969 to 24 A  
ugust, 1974',  
'Term of Office: 13 May, 1967 to 3 May, 1969',  
'Term of Office: 13 May, 1962 to 13 May, 1967',  
'Term of Office: 26 January, 1950 to 13 May, 1962']
```

In [41]:

```
print(len(name),len(term_of_office))
```

14 14

In [42]:

```
import pandas as pd
df=pd.DataFrame({'Name':name, 'Term_of_office':term_of_office})
df
```

Out[42]:

	Name	Term_of_office
0	Shri Ram Nath Kovind (birth - 1945)	Term of Office: 25 July, 2017 to 25 July, 2022
1	Shri Pranab Mukherjee (1935-2020)	Term of Office: 25 July, 2012 to 25 July, 2017
2	Smt Pratibha Devisingh Patil (birth - 1934)	Term of Office: 25 July, 2007 to 25 July, 2012
3	DR. A.P.J. Abdul Kalam (1931-2015)	Term of Office: 25 July, 2002 to 25 July, 2007
4	Shri K. R. Narayanan (1920 - 2005)	Term of Office: 25 July, 1997 to 25 July, 2002
5	Dr Shankar Dayal Sharma (1918-1999)	Term of Office: 25 July, 1992 to 25 July, 1997
6	Shri R Venkataraman (1910-2009)	Term of Office: 25 July, 1987 to 25 July, 1992
7	Giani Zail Singh (1916-1994)	Term of Office: 25 July, 1982 to 25 July, 1987
8	Shri Neelam Sanjiva Reddy (1913-1996)	Term of Office: 25 July, 1977 to 25 July, 1982
9	Dr. Fakhruddin Ali Ahmed (1905-1977)	Term of Office: 24 August, 1974 to 11 February...
10	Shri Varahaqiri Venkata Giri (1894-1980)	Term of Office: 3 May 1969 to 20 July 1969 a

In []:

In []:

In []:

In []:

In []: