In [3]:
```
from urllib.request import urlopen
from bs4 import BeautifulSoup
html = urlopen('https://en.wikipedia.org/wiki/Ajith_Kumar')
bs = BeautifulSoup(html, "html.parser")
titles = bs.find_all(['h1', 'h2','h3','h4','h5','h6'])
print('List all the header tags :', *titles, sep='\n\n')
```

```
List all the header tags :

<h2 class="vector-pinnable-header-label">Contents</h2>

<h1 class="firstHeading mw-first-heading" id="firstHeading"><span class="mw-p
age-title-main">Ajith Kumar</span></h1>

<h2><span class="mw-headline" id="Early_life">Early life</span></h2>

<h2><span class="mw-headline" id="Acting_career">Acting career</span></h2>

<h3><span id="1990.E2.80.931998"></span><span class="mw-headline" id="1990–19
98">1990–1998</span></h3>

<h3><span id="1999.E2.80.932008"></span><span class="mw-headline" id="1999–20
08">1999–2008</span></h3>

<h3><span id="2010.E2.80.932019"></span><span class="mw-headline" id="2010–20
19">2010–2019</span></h3>

<h3><span class="mw-headline" id="2022-present">2022-present</span></h3>

<h2><span class="mw-headline" id="Other_works">Other works</span></h2>

<h2><span class="mw-headline" id="Racing_career">Racing career</span></h2>

<h3><span id="Formula_BMW_Asia_.282003.29"></span><span class="mw-headline" i
d="Formula_BMW_Asia_(2003)">Formula BMW Asia (2003)</span></h3>

<h3><span id="Formula_2_.282010.29"></span><span class="mw-headline" id="Form
ula_2_(2010)">Formula 2 (2010)</span></h3>

<h2><span class="mw-headline" id="Personal_life">Personal life</span></h2>

<h2><span class="mw-headline" id="Filmography">Filmography</span></h2>

<h2><span class="mw-headline" id="Awards">Awards</span></h2>

<h2><span class="mw-headline" id="Notes">Notes</span></h2>

<h2><span class="mw-headline" id="References">References</span></h2>

<h2><span class="mw-headline" id="External_links">External links</span></h2>
```

In [14]:
```
import requests
from bs4 import BeautifulSoup
import pandas as pd

# URL of the website
url = "https://presidentofindia.nic.in/former-presidents.htm"

# Send a GET request to the URL
response = requests.get(url)

# Check if the request was successful (status code 200)
if response.status_code == 200:
    # Parse the HTML content of the page
    soup = BeautifulSoup(response.text, "html.parser")

    # Find the table containing the list of former presidents
    table = soup.find("table", {"class": "table table-bordered table-striped t

    # Initialize lists to store the data
    names = []
    terms = []

    # Extract data from the table
    for row in table.find_all("tr")[1:]:  # Skip the header row
        columns = row.find_all("td")
        name = columns[0].text.strip()
        term = columns[1].text.strip()
        names.append(name)
        terms.append(term)

    # Create a DataFrame
    data = {"Name": names, "Term of Office": terms}
    df = pd.DataFrame(data)

    # Display the DataFrame
    print(df)
else:
    print("Failed to retrieve the web page.")
```

```
Failed to retrieve the web page.
```

In [15]:
```
import requests
from bs4 import BeautifulSoup

# Define the URL of the web page to scrape
url = "https://presidentofindia.nic.in/former-presidents"

# Send an HTTP GET request to the URL
response = requests.get(url)

# Check if the request was successful (status code 200)
if response.status_code == 200:
    # Get the HTML content of the page
    html_content = response.text

    # Parse the HTML content using BeautifulSoup
    soup = BeautifulSoup(html_content, 'html.parser')

    # Find all <h3> tags and print their text content
    h3_tags = soup.find_all('h3')

    for h3 in h3_tags:
        print(h3.text)
else:
    print("Failed to retrieve the web page.")
```

```
Shri Ram Nath Kovind
Shri Pranab Mukherjee
Smt Pratibha Devisingh Patil
DR. A.P.J. Abdul Kalam
Shri K. R. Narayanan
Dr Shankar Dayal Sharma
Shri R Venkataraman
Giani Zail Singh
Shri Neelam Sanjiva Reddy
Dr. Fakhruddin Ali Ahmed
Shri Varahagiri Venkata Giri
Dr. Zakir Husain
Dr. Sarvepalli Radhakrishnan
Dr. Rajendra Prasad
```

In [20]:
```
import requests
from bs4 import BeautifulSoup
import pandas as pd

# Define the URL for the most downloaded articles page
url = "https://www.journals.elsevier.com/artificial-intelligence/most-download

# Send a GET request to the URL
response = requests.get(url)

# Check if the request was successful (status code 200)
if response.status_code == 200:
    soup = BeautifulSoup(response.text, "html.parser")

    # Find the container that holds the article titles
    article_titles = soup.find_all("div", class_="pod-listing-title")

    # Initialize a list to store the paper titles
    paper_titles = []

    # Extract data from the container
    for title in article_titles:
        paper_title = title.find("a").text.strip()
        paper_titles.append(paper_title)

    # Create a DataFrame
    data = {"Paper Title": paper_titles}
    df = pd.DataFrame(data)

    # Display the DataFrame
    print(df)
else:
    print("Failed to retrieve the web page.")
```

```
Empty DataFrame
Columns: [Paper Title]
Index: []
```

In [22]:
```
import requests
from bs4 import BeautifulSoup
import pandas as pd

# Define the URL for dineout.co.in
url = "https://www.dineout.co.in/"

# Send a GET request to the URL
response = requests.get(url)

# Check if the request was successful (status code 200)
if response.status_code == 200:
    soup = BeautifulSoup(response.text, "html.parser")

    # Find the container that holds the restaurant names
    restaurant_containers = soup.find_all("div", class_="restnt-info")

    # Initialize a list to store the restaurant names
    restaurant_names = []

    # Extract data from the containers
    for container in restaurant_containers:
        name = container.find("h3").text.strip()
        restaurant_names.append(name)

    # Create a DataFrame
    data = {"Restaurant Name": restaurant_names}
    df = pd.DataFrame(data)

    # Display the DataFrame
    print(df)
else:
    print("Failed to retrieve the web page.")
```

```
Empty DataFrame
Columns: [Restaurant Name]
Index: []
```

In [ ]:
```

```