WEB SCRAPING Assignment

In all the following questions, you have to use BeautifulSoup to scrape different websites and collect data as per the requirement of the question.

Print Data Frame in every Answers of the questions

Use Jupyter Notebooks to program, upload it on your GitHub and send the link of the jupyter notebook to your SME.

1. Write a python program to display all the header tags from‘en.wikipedia.org/wiki/Main_Page’.

In [1]:
```python
from urllib.request import urlopen
from bs4 import BeautifulSoup
import requests
page = requests.get('https://en.wikipedia.org/wiki/Main_Page')
page.content
soup = BeautifulSoup(page.content, "html.parser")
titles = soup.find_all(['h1', 'h2','h3','h4','h5','h6'])
print('List all the header tags :', *titles, sep='\n\n')
```

```
List all the header tags :

<h1 class="firstHeading" id="firstHeading">Main Page</h1>

<h2 class="mp-h2" id="mp-tfa-h2"><span id="From_today.27s_featured_article"></span><span class="mw-headline" id="From_t
oday's_featured_article">From today's featured article</span></h2>

<h2 class="mp-h2" id="mp-dyk-h2"><span class="mw-headline" id="Did_you_know_...">Did you know ...</span></h2>

<h2 class="mp-h2" id="mp-itn-h2"><span class="mw-headline" id="In_the_news">In the news</span></h2>

<h2 class="mp-h2" id="mp-otd-h2"><span class="mw-headline" id="On_this_day">On this day</span></h2>

<h2 class="mp-h2" id="mp-tfp-h2"><span id="Today.27s_featured_picture"></span><span class="mw-headline" id="Today's_fea
tured_picture">Today's featured picture</span></h2>

<h2 class="mp-h2" id="mp-other"><span class="mw-headline" id="Other_areas_of_Wikipedia">Other areas of Wikipedia</span>
</h2>

<h2 class="mp-h2" id="mp-sister"><span id="Wikipedia.27s_sister_projects"></span><span class="mw-headline" id="Wikipedi
a's_sister_projects">Wikipedia's sister projects</span></h2>

<h2 class="mp-h2" id="mp-lang"><span class="mw-headline" id="Wikipedia_languages">Wikipedia languages</span></h2>

<h2>Navigation menu</h2>

<h3 class="vector-menu-heading" id="p-personal-label"><span>Personal tools</span>
</h3>

<h3 class="vector-menu-heading" id="p-namespaces-label"><span>Namespaces</span>
</h3>
```

```
<h3 class="vector-menu-heading" id="p-variants-label"><span>Variants</span>
</h3>

<h3 class="vector-menu-heading" id="p-views-label"><span>Views</span>
</h3>

<h3 class="vector-menu-heading" id="p-cactions-label"><span>More</span>
</h3>

<h3>
<label for="searchInput">Search</label>
</h3>

<h3 class="vector-menu-heading" id="p-navigation-label"><span>Navigation</span>
</h3>

<h3 class="vector-menu-heading" id="p-interaction-label"><span>Contribute</span>
</h3>

<h3 class="vector-menu-heading" id="p-tb-label"><span>Tools</span>
</h3>

<h3 class="vector-menu-heading" id="p-coll-print_export-label"><span>Print/export</span>
</h3>

<h3 class="vector-menu-heading" id="p-wikibase-otherprojects-label"><span>In other projects</span>
</h3>

<h3 class="vector-menu-heading" id="p-lang-label"><span>Languages</span>
</h3>
```

In [9]:

```
page.status_code
```

Out[9]:

```
200
```

In [3]:
```
all_headers=[]
for i in range(0,len(titles)):
    all_headers.append(titles[i].get_text().replace("\n"," "))
all_headers
```

Out[3]:
```
['Main Page',
 "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',
 'Navigation menu',
 'Personal tools ',
 'Namespaces ',
 'Variants ',
 'Views ',
 'More ',
 ' Search ',
 'Navigation ',
 'Contribute ',
 'Tools ',
 'Print/export ',
 'In other projects ',
 'Languages ']
```

In [2]:
```
import pandas as pd
```

In [16]:

```
df=pd.DataFrame(all_headers)
df
```

Out[16]:

	0
0	Main Page
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
9	Navigation menu
10	Personal tools
11	Namespaces
12	Variants
13	Views
14	More
15	Search
16	Navigation
17	Contribute
18	Tools
19	Print/export
20	In other projects
21	Languages

In [17]:
```
len(df)
```

Out[17]:
```
22
```

2. Write a python program to display IMDB’s Top rated 100 movies’ data (i.e. Name, IMDB rating, Year of release).

In [19]:
```
page=requests.get("https://www.imdb.com/chart/top/")
```

In [20]:
```
page
```

Out[20]:
```
<Response [200]>
```

In [21]:
```
soup=BeautifulSoup(page.content,"html.parser")
soup
```

Out[21]:
```

<!DOCTYPE html>

<html xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://ogp.me/ns#">
<head>
<meta charset="utf-8"/>
<meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<style>
                body#styleguide-v2 {
                    background: no-repeat fixed center top #000;
                }
            </style>
<script type="text/javascript">var IMDbTimer={starttime: new Date().getTime(),pt:'java'};</script>
<script>
    if (typeof uet == 'function') {
      uet("bb", "LoadTitle", {wb: 1});
    }
</script>
<script>(function(t){ (t.events = t.events || {})["csm_head_pre_title"] = new Date().getTime(); })(IMDbTimer);</scri
pt>
```

In [22]:
```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://ogp.me/ns#">
 <head>
  <meta charset="utf-8"/>
  <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
  <style>
   body#styleguide-v2 {
                    background: no-repeat fixed center top #000;
                }
  </style>
  <script type="text/javascript">
   var IMDbTimer={starttime: new Date().getTime(),pt:'java'};
  </script>
  <script>
   if (typeof uet == 'function') {
      uet("bb", "LoadTitle", {wb: 1});
    }
  </script>
  <script>
```

In [23]:
```
#top 100 movies name
name=soup.find_all("td",class_="titleColumn")
name
```

Out[23]:
```
[<td class="titleColumn">
       1.
       <a href="/title/tt0111161/" title="Frank Darabont (dir.), Tim Robbins, Morgan Freeman">The Shawshank Redempti
on</a>
 <span class="secondaryInfo">(1994)</span>
 </td>,
 <td class="titleColumn">
       2.
       <a href="/title/tt0068646/" title="Francis Ford Coppola (dir.), Marlon Brando, Al Pacino">The Godfather</a>
 <span class="secondaryInfo">(1972)</span>
 </td>,
 <td class="titleColumn">
       3.
       <a href="/title/tt0071562/" title="Francis Ford Coppola (dir.), Al Pacino, Robert De Niro">The Godfather: Par
t II</a>
 <span class="secondaryInfo">(1974)</span>
 </td>,
 <td class="titleColumn">
       4.
       <a href="/title/tt0468569/" title="Christopher Nolan (dir.), Christian Bale, Heath Ledger">The Dark Knight</
```

In [24]:
```
movies_name=[]
for i in name:
    movies_name.append(i.text.replace("\n",""))
movies_name
```

Out[24]:
```
['       1.      The Shawshank Redemption(1994)',
 '       2.      The Godfather(1972)',
 '       3.      The Godfather: Part II(1974)',
 '       4.      The Dark Knight(2008)',
 '       5.      12 Angry Men(1957)',
 "       6.      Schindler's List(1993)",
 '       7.      The Lord of the Rings: The Return of the King(2003)',
 '       8.      Pulp Fiction(1994)',
 '       9.      Il buono, il brutto, il cattivo(1966)',
 '       10.      The Lord of the Rings: The Fellowship of the Ring(2001)',
 '       11.      Fight Club(1999)',
 '       12.      Forrest Gump(1994)',
 '       13.      Inception(2010)',
 '       14.      The Lord of the Rings: The Two Towers(2002)',
 '       15.      Star Wars: Episode V - The Empire Strikes Back(1980)',
 '       16.      The Matrix(1999)',
 '       17.      Goodfellas(1990)',
 "       18.      One Flew Over the Cuckoo's Nest(1975)",
 '       19.      Shichinin no samurai(1954)',
```

In [25]:
```
movies_name=[]
for i in name:
    for j in i.find_all("a"):
        movies_name.append(j.text.replace("\n",""))
movies_name
```

Out[25]:
```
['The Shawshank Redemption',
 'The Godfather',
 'The Godfather: Part II',
 'The Dark Knight',
 '12 Angry Men',
 "Schindler's List",
 'The Lord of the Rings: The Return of the King',
 'Pulp Fiction',
 'Il buono, il brutto, il cattivo',
 'The Lord of the Rings: The Fellowship of the Ring',
 'Fight Club',
 'Forrest Gump',
 'Inception',
 'The Lord of the Rings: The Two Towers',
 'Star Wars: Episode V - The Empire Strikes Back',
 'The Matrix',
 'Goodfellas',
 "One Flew Over the Cuckoo's Nest",
 'Shichinin no samurai',
```

In [45]:
```
# top 100 rating of movies
rating=soup.find_all("td",class_="ratingColumn imdbRating")
rating
```

Out[45]:
```
[<td class="ratingColumn imdbRating">
 <strong title="9.2 based on 2,450,181 user ratings">9.2</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.1 based on 1,695,979 user ratings">9.1</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.0 based on 1,178,004 user ratings">9.0</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.0 based on 2,405,903 user ratings">9.0</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="8.9 based on 725,178 user ratings">8.9</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="8.9 based on 1,260,498 user ratings">8.9</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="8.9 based on 1,703,181 user ratings">8.9</strong>
```

In [46]:
```
imdb_rating=[]
for i in range(0,len(rating)):
    imdb_rating.append(rating[i].get_text().replace("\n"," "))
imdb_rating
```

Out[46]:
```
[' 9.2 ',
 ' 9.1 ',
 ' 9.0 ',
 ' 9.0 ',
 ' 8.9 ',
 ' 8.9 ',
 ' 8.9 ',
 ' 8.8 ',
 ' 8.8 ',
 ' 8.8 ',
 ' 8.8 ',
 ' 8.7 ',
 ' 8.7 ',
 ' 8.7 ',
 ' 8.7 ',
 ' 8.6 ',
 ' 8.6 ',
 ' 8.6 ',
 ' 8.6 ',
In [47]: df=pd.DataFrame(imdb_rating)
         df
```

```
Out[47]:
           0
  0      9.2
  1      9.1
  2      9.0
  3      9.0
  4      8.9
...      ...
245      8.0
246      8.0
247      8.0
248      8.0
249      8.0

250 rows × 1 columns
```

```
In [ ]:
```

In [48]:

```
#top 100 movies year
year=soup.find_all("span",class_="secondaryInfo")
year
```

Out[48]:

```
[<span class="secondaryInfo">(1994)</span>,
 <span class="secondaryInfo">(1972)</span>,
 <span class="secondaryInfo">(1974)</span>,
 <span class="secondaryInfo">(2008)</span>,
 <span class="secondaryInfo">(1957)</span>,
 <span class="secondaryInfo">(1993)</span>,
 <span class="secondaryInfo">(2003)</span>,
 <span class="secondaryInfo">(1994)</span>,
 <span class="secondaryInfo">(1966)</span>,
 <span class="secondaryInfo">(2001)</span>,
 <span class="secondaryInfo">(1999)</span>,
 <span class="secondaryInfo">(1994)</span>,
 <span class="secondaryInfo">(2010)</span>,
 <span class="secondaryInfo">(2002)</span>,
 <span class="secondaryInfo">(1980)</span>,
 <span class="secondaryInfo">(1999)</span>,
 <span class="secondaryInfo">(1990)</span>,
 <span class="secondaryInfo">(1975)</span>,
 <span class="secondaryInfo">(1954)</span>,
```

In [55]:
```
imdb_year=[]
for i in range(0,len(year)):
    imdb_year.append(year[i].get_text().replace("\n",""))
imdb_year
```

Out[55]:
```
['(1994)',
 '(1972)',
 '(1974)',
 '(2008)',
 '(1957)',
 '(1993)',
 '(2003)',
 '(1994)',
 '(1966)',
 '(2001)',
 '(1999)',
 '(1994)',
 '(2010)',
 '(2002)',
 '(1980)',
 '(1999)',
 '(1990)',
 '(1975)',
 '(1954)',
```

In [56]:

```
df=pd.DataFrame(imdb_year)
df
```

Out[56]:

```
          0
  0  (1994)
  1  (1972)
  2  (1974)
  3  (2008)
  4  (1957)
...     ...
245  (1984)
246  (1927)
247  (1957)
248  (2018)
249  (2013)

250 rows × 1 columns
```

In [58]:

```
import pandas as pd
imdb_movies=pd.DataFrame({})
imdb_movies['movies_name']=movies_name[:100]
imdb_movies['imdb_rating']=imdb_rating[:100]
imdb_movies['imdb_year']=imdb_year[:100]
imdb_movies
```

Out[58]:

	movies_name	imdb_rating	imdb_year
0	The Shawshank Redemption	9.2	(1994)
1	The Godfather	9.1	(1972)
2	The Godfather: Part II	9.0	(1974)
3	The Dark Knight	9.0	(2008)
4	12 Angry Men	8.9	(1957)
...	...	...	...
95	Jagten	8.3	(2012)
96	Requiem for a Dream	8.3	(2000)
97	Singin' in the Rain	8.3	(1952)
98	North by Northwest	8.3	(1959)
99	Eternal Sunshine of the Spotless Mind	8.3	(2004)

100 rows × 3 columns

In [59]:

```
len(imdb_movies)
```

Out[59]:

```
100
```

3.Write a python program to display IMDB’s Top rated 100 Indian movies’ data (i.e. Name, IMDB rating, Yearof release).

In [60]:
```
page = requests.get("https://www.imdb.com/india/top-rated-indian-movies/?pf_rd_m=A2FGELUUNOQJNL&pf_rd_p=461131e5-5af0-4e
```

In [61]:
```
page
```

Out[61]:
```
<Response [200]>
```

In [62]:
```
soup=BeautifulSoup(page.content)
soup
```

Out[62]:
```
<!DOCTYPE html>
<html xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://ogp.me/ns#">
<head>
<meta charset="utf-8"/>
<meta content="IE=edge" http-equiv="X-UA-Compatible"/>
<style>
                body#styleguide-v2 {
                    background: no-repeat fixed center top #000;
                }
            </style>
<style>
            body#styleguide-v2 #root {
                box-shadow: none;
            }
        </style>
<script type="text/javascript">var IMDbTimer={starttime: new Date().getTime(),pt:'java'};</script>
<script>
    if (typeof uet == 'function') {
      uet("bb", "LoadTitle", {wb: 1});
```

In [63]:
```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://ogp.me/ns#">
 <head>
  <meta charset="utf-8"/>
  <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
  <style>
   body#styleguide-v2 {
                    background: no-repeat fixed center top #000;
                }
  </style>
  <style>
   body#styleguide-v2 #root {
                box-shadow: none;
            }
  </style>
  <script type="text/javascript">
   var IMDbTimer={starttime: new Date().getTime(),pt:'java'};
  </script>
  <script>
   if (typeof uet == 'function') {
```

In [64]:
```
#top 100 indian movies name
name=soup.find_all("td",class_="titleColumn")
name
```

Out[64]:
```
[<td class="titleColumn">
       1.
       <a href="/title/tt0093603/" title="Mani Ratnam (dir.), Kamal Haasan, Saranya Ponvannan">Nayakan</a>
 <span class="secondaryInfo">(1987)</span>
 </td>,
 <td class="titleColumn">
       2.
       <a href="/title/tt0367495/" title="Sundar C. (dir.), Kamal Haasan, Madhavan">Anbe Sivam</a>
 <span class="secondaryInfo">(2003)</span>
 </td>,
 <td class="titleColumn">
       3.
       <a href="/title/tt8176054/" title="Mari Selvaraj (dir.), Kathir, Anandhi">Pariyerum Perumal</a>
 <span class="secondaryInfo">(2018)</span>
 </td>,
 <td class="titleColumn">
       4.
       <a href="/title/tt0079221/" title="Hrishikesh Mukherjee (dir.), Amol Palekar, Bindiya Goswami">Golmaal</a>
 <span class="secondaryInfo">(1979)</span>
```

In [65]:

```
movies_name=[]
for i in name:
    movies_name.append(i.text.replace("\n",""))
movies_name
```

Out[65]:

```
['      1.      Nayakan(1987)',
 '      2.      Anbe Sivam(2003)',
 '      3.      Pariyerum Perumal(2018)',
 '      4.      Golmaal(1979)',
 '      5.      C/o Kancharapalem(2018)',
 '      6.      Apur Sansar(1959)',
 '      7.      Manichitrathazhu(1993)',
 '      8.      Kireedam(1989)',
 '      9.      Natsamrat(2016)',
 '      10.      Pather Panchali(1955)',
 '      11.      96(2018)',
 '      12.      Thevar Magan(1992)',
 '      13.      Black Friday(2004)',
 '      14.      Kumbalangi Nights(2019)',
 '      15.      Visaaranai(2015)',
 '      16.      3 Idiots(2009)',
 '      17.      Soorarai Pottru(2020)',
 '      18.      Jersey(2019)',
 '      19.      Taare Zameen Par(2007)',
```

In [66]:

```
movies_name=[]
for i in name:
    for j in i.find_all("a"):
        movies_name.append(j.text.replace("\n",""))
movies_name
```

Out[66]:

```
['Nayakan',
 'Anbe Sivam',
 'Pariyerum Perumal',
 'Golmaal',
 'C/o Kancharapalem',
 'Apur Sansar',
 'Manichitrathazhu',
 'Kireedam',
 'Natsamrat',
 'Pather Panchali',
 '96',
 'Thevar Magan',
 'Black Friday',
 'Kumbalangi Nights',
 'Visaaranai',
 '3 Idiots',
 'Soorarai Pottru',
 'Jersey',
 'Taare Zameen Par',
```

In [67]:

```
#top 100 indian ratings movies
rating=soup.find_all("td",class_="ratingColumn imdbRating")
```

In [69]:
```
rating
```

Out[69]:
```
[<td class="ratingColumn imdbRating">
 <strong title="9.2 based on 2,450,181 user ratings">9.2</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.1 based on 1,695,979 user ratings">9.1</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.0 based on 1,178,004 user ratings">9.0</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="9.0 based on 2,405,903 user ratings">9.0</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="8.9 based on 725,178 user ratings">8.9</strong>
 </td>,
 <td class="ratingColumn imdbRating">
 <strong title="8.9 based on 1,260,498 user ratings">8.9</strong>
 </td>,
 <td class="ratingColumn imdbRating">
```

In [70]:
```
imdb_rating=[]
for i in range(0,len(rating)):
    imdb_rating.append(rating[i].get_text().replace("\n",""))
imdb_rating
```

Out[70]:
```
['9.2',
 '9.1',
 '9.0',
 '9.0',
 '8.9',
 '8.9',
 '8.9',
 '8.8',
 '8.8',
 '8.8',
 '8.8',
 '8.7',
 '8.7',
 '8.7',
 '8.7',
 '8.6',
 '8.6',
 '8.6',
 '8.6',
```

In [71]:

```
df=pd.DataFrame(imdb_rating)
df
```

Out[71]:

```
       0
  0  9.2
  1  9.1
  2  9.0
  3  9.0
  4  8.9
...  ...
245  8.0
246  8.0
247  8.0
248  8.0
249  8.0

250 rows × 1 columns
```

In [31]:

```
#top 100 indian movies year
year=soup.find_all("span",class_="secondaryInfo")
year
```

Out[31]:

```
[<span class="secondaryInfo">(1987)</span>,
 <span class="secondaryInfo">(2003)</span>,
 <span class="secondaryInfo">(2018)</span>,
 <span class="secondaryInfo">(1979)</span>,
 <span class="secondaryInfo">(2018)</span>,
 <span class="secondaryInfo">(1959)</span>,
 <span class="secondaryInfo">(1993)</span>,
 <span class="secondaryInfo">(1989)</span>,
 <span class="secondaryInfo">(2016)</span>,
 <span class="secondaryInfo">(1955)</span>,
 <span class="secondaryInfo">(2018)</span>,
 <span class="secondaryInfo">(1992)</span>,
 <span class="secondaryInfo">(2004)</span>,
 <span class="secondaryInfo">(2019)</span>,
 <span class="secondaryInfo">(2015)</span>,
 <span class="secondaryInfo">(2009)</span>,
 <span class="secondaryInfo">(2020)</span>,
 <span class="secondaryInfo">(2019)</span>,
 <span class="secondaryInfo">(2007)</span>,
 <span class="secondaryInfo">(2021)</span>,
```

In [72]:

```
imdb_year=[]
for i in range(0,len(year)):
    imdb_year.append(year[i].get_text().replace("\n",""))
imdb_rating
```

Out[72]:

```
['9.2',
 '9.1',
 '9.0',
 '9.0',
 '8.9',
 '8.9',
 '8.9',
 '8.8',
 '8.8',
 '8.8',
 '8.8',
 '8.7',
 '8.7',
 '8.7',
 '8.7',
 '8.6',
 '8.6',
 '8.6',
 '8.6',
 '8.6'
```

In [74]:
```
df=pd.DataFrame(imdb_rating)
df
```

Out[74]:
```
       0
  0  9.2
  1  9.1
  2  9.0
  3  9.0
  4  8.9
...  ...
245  8.0
246  8.0
247  8.0
248  8.0
249  8.0

250 rows × 1 columns
```

In [78]:

```
indian_movies=pd.DataFrame({})
indian_movies['movies_name']=movies_name[0:100]
indian_movies['movies_ratings']=imdb_rating[0:100]
indian_movies["movies_year"]=imdb_year[0:100]
indian_movies
```

Out[78]:

	movies_name	movies_ratings	movies_year
0	Nayakan	9.2	(1994)
1	Anbe Sivam	9.1	(1972)
2	Pariyerum Perumal	9.0	(1974)
3	Golmaal	9.0	(2008)
4	C/o Kancharapalem	8.9	(1957)
...	...	...	...
95	Virumandi	8.3	(2012)
96	Rang De Basanti	8.3	(2000)
97	Lagaan: Once Upon a Time in India	8.3	(1952)
98	Uri: The Surgical Strike	8.3	(1959)
99	Andaz Apna Apna	8.3	(2004)

100 rows × 3 columns

In [ ]:

```

```

In [35]:

```
len(indian_movies)
```

Out[35]:

```
100
```

4. Write a python program to scrap book name, author name, genre and book review of any 5 books from'www.bookpage.com' (http://www.bookpage.com%E2%80%99)

In [213]:
```
page=requests.get("https://bookpage.com/reviews/26593-winfred-rembert-chasing-me-to-my-grave-nonfiction#.YSEBeflR1PY")
page
```

Out[213]:
```
<Response [200]>
```

In [214]:
```
soup=BeautifulSoup(page.content)
soup
```

Out[214]:
```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - Chasing Me to My Grave by Winfred Rembert | BookPage</title>
<meta content="Winfred Rembert recounts gripping, often harrowing stories of growing up in Georgia, surviving a lync
hing and discovering art while imprisoned in a chain gang." name="description"/>
<link href="https://bookpage.com/reviews/26593-winfred-rembert-chasing-me-to-my-grave-nonfiction" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - Chasing Me to My Grave by Winfred Rembert" property="og:title"/>
<meta content="Winfred Rembert recounts gripping, often harrowing stories of growing up in Georgia, surviving a lync
hing and discovering art while imprisoned in a chain gang." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26593-winfred-rembert-chasing-me-to-my-grave-nonfiction" property="og:ur
l"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - Chasing Me to My Grave by Winfred Rembert" name="twitter:title"/>
<meta content="Winfred Rembert recounts gripping, often harrowing stories of growing up in Georgia, surviving a lync
hing and discovering art while imprisoned in a chain gang." name="twitter:description"/>
```

In [215]:
```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
 <head>
  <title>
   Book Review - Chasing Me to My Grave by Winfred Rembert | BookPage
  </title>
  <meta content="Winfred Rembert recounts gripping, often harrowing stories of growing up in Georgia, surviving a ly
nching and discovering art while imprisoned in a chain gang." name="description"/>
  <link href="https://bookpage.com/reviews/26593-winfred-rembert-chasing-me-to-my-grave-nonfiction" rel="canonical"/
>
  <meta content="BookPage.com" property="og:site_name"/>
  <meta content="Book Review - Chasing Me to My Grave by Winfred Rembert" property="og:title"/>
  <meta content="Winfred Rembert recounts gripping, often harrowing stories of growing up in Georgia, surviving a ly
nching and discovering art while imprisoned in a chain gang." property="og:description"/>
  <meta content="article" property="og:type"/>
  <meta content="https://bookpage.com/reviews/26593-winfred-rembert-chasing-me-to-my-grave-nonfiction" property="og:
url"/>
  <meta content="/images/medium/missing.png" property="og:image"/>
  <meta content="summary" name="twitter:card"/>
```

In [250]:
```
book_name=soup.find("h1",class_="italic")
book_name
```

Out[250]:
```
<h1 class="italic">
<span style="font-style:normal;">★ </span>How to Find Your Way in the Dark
</h1>
```

In [251]:
```
book_name1=book_name.text.replace("\n","")
book_name1
```

Out[251]:
```
'★ How to Find Your Way in the Dark'
```

In [ ]:
```

```

In [253]:

```
author_name=soup.find("h4",class_="sans")
author_name
```

Out[253]:

```
<h4 class="sans">
Derek B. Miller
</h4>
```

In [254]:

```
author_name1=author_name.text.replace("\n","")
author_name1
```

Out[254]:

```
'Derek B. Miller'
```

In [260]:

```
genre=soup.find("p",class_="genre-links")
genre
```

Out[260]:

```
<p class="genre-links">
<a href="/search?book_genre=fiction">Fiction</a>
 /
<a href="/search?book_genre=coming_of_age">Coming of Age</a>
</p>
```

In [261]:

```
genre1=genre.text.replace("\n","")
genre1
```

Out[261]:

```
'Fiction / Coming of Age'
```

In [267]:

```
book_review=soup.find("div",class_="content-well article-show")
book_review
```

Out[267]:

```
<div class="content-well article-show">
<div class="container-fluid">
<div class="row-fluid">
<div class="span7">
<div class="article-body">
<p>
</p><p>Prepare for surprises galore in <b>How to Find Your Way in the Dark</b>, a rollicking novel that begins with a l
onely truck ride in New England in 1938 and follows its characters through a decade of fascinating history. Just when y
ou think the story is heading one way, it veers in another, completely unexpected direction.</p>
<p>Twelve-year-old Sheldon Horowitz and his father are driving home from Hartford, Connecticut, to Whately, Massachuset
ts, after honoring the one-year anniversary of Sheldon’s mother’s death. She and her sister died in a horrific movie th
eater fire in Hartford. And as if that isn’t enough tragedy for the novel’s first 13 pages, a truck purposely forces Sh
eldon and his father’s car off the road during their return trip, and Sheldon’s father dies.</p>
<p>Readers of Derek B. Miller’s award-winning thriller, <a href="https://bookpage.com/reviews/15318-derek-b-miller-shel
dons-sense-snow-fiction#.YN9gC5NKjUI"><i>Norwegian by Night</i></a>, will recognize Sheldon as that novel’s 82-year-old
protagonist. As a Tom Sawyer-like boy in <b>How to Find Your Way in the Dark</b>, Sheldon is determined to make sense o
f his double tragedies, and his attempts to do so take the reader on one hell of a ride. As he seeks out the leering, m
ustached truck driver who killed his father, his quest leads him straight into danger—think mobsters, guns and jewel th
efts.</p>
<p>Miller has crafted a wide-ranging, years-spanning yet tightly structured plot, and he excels at placing memorable ch
aracters in unusual circumstances. Sheldon is joined in his adventures by his two older cousins, Abe and Mirabelle, and
his best friend, Lenny, all of whom play pivotal roles. One summer, Lenny and Sheldon end up as bellhops at the famed G
rossinger’s Resort in the Catskills, where Lenny practices standup comedy amid the glamorous, bustling atmosphere.</p>
<p>An underlying seriousness lies at the heart of all of this intrigue, hilarity and fun. Sheldon, Abe, Mirabelle and L
enny, all Jewish, must confront the many faces of antisemitism during the turbulent years of World War II. Miller weave
s in a multitude of historical details, including reports of the horrors in Europe and America’s reluctance to interven
e.</p>
<p>The ending of <b>How to Find Your Way in the Dark</b> is nothing short of brilliant, tying up a variety of loose end
s while making a powerful statement about the need to fully recognize and address antisemitism. Readers are left with m
uch to ponder, including life’s many uncertainties and cruel twists of fate. Despite these unhappy truths, we are also
left with the uplifting wisdom of Lenny’s urgent prayer: “Dear God, give me the strength to be joyful.”</p>
</div>
</div>
<div class="span1"></div>
<div class="span4">
<div class="book-details">
<div class="book-cover">
<img alt="How to Find Your Way in the Dark by Miller, Derek B. book jacket image" src="https://t1.bookpage.com/books/im
```

```
ages/87fc22995c9de505ed6273a7934bd4f5/medium.jpg"/>
</div>
<div class="book-data">
<h4 class="italic">
<span style="font-style:normal;">★ </span>How to Find Your Way in the Dark
</h4>
<p class="sans bold">
By
<a href="/author/derek-b-miller?book_author_id=9475">Derek B. Miller</a>
</p>
<p>
HMH
<br/>
$26.00
<br/>
ISBN
9780358269601
<br/>
Published
07/27/2021
</p>
<p class="genre-links">
<a href="/search?book_genre=fiction">Fiction</a>
 /
<a href="/search?book_genre=coming_of_age">Coming of Age</a>
</p>
</div>
<div class="affiliate-links-small">
</div>
</div>
<div class="sq-ad">
<script language="JavaScript" type="text/javascript">
 aj_server = 'https://litbreaker.nui.media/pipeline/'; aj_tagver = '1.0';
 aj_zone = 'litbreaker'; aj_adspot = '519024'; aj_page = '0'; aj_dim ='514841'; aj_ch = ''; aj_ct = ''; aj_kw = '';
 aj_pv = true; aj_click = '';
</script><script language="JavaScript" src="https://img.nui.media/banners/ajtg.js" type="text/javascript"></script>
</div>
</div>
</div>
</div>
</div>
```

In [268]:
```
book_review1=book_review.text.replace("\n","")
book_review1
```

Out[268]:
```
'Prepare for surprises galore in How to Find Your Way in the Dark, a rollicking novel that begins with a lonely truck r
ide in New England in 1938 and follows its characters through a decade of fascinating history. Just when you think the
story is heading one way, it veers in another, completely unexpected direction.Twelve-year-old Sheldon Horowitz and his
father are driving home from Hartford, Connecticut, to Whately, Massachusetts, after honoring the one-year anniversary
of Sheldon’s mother’s death. She and her sister died in a horrific movie theater fire in Hartford. And as if that isn’t
enough tragedy for the novel’s first 13 pages, a truck purposely forces Sheldon and his father’s car off the road durin
g their return trip, and Sheldon’s father dies.Readers of Derek B. Miller’s award-winning thriller, Norwegian by Night,
will recognize Sheldon as that novel’s 82-year-old protagonist. As a Tom Sawyer-like boy in How to Find Your Way in the
Dark, Sheldon is determined to make sense of his double tragedies, and his attempts to do so take the reader on one hel
l of a ride. As he seeks out the leering, mustached truck driver who killed his father, his quest leads him straight in
to danger—think mobsters, guns and jewel thefts.Miller has crafted a wide-ranging, years-spanning yet tightly structure
d plot, and he excels at placing memorable characters in unusual circumstances. Sheldon is joined in his adventures by
his two older cousins, Abe and Mirabelle, and his best friend, Lenny, all of whom play pivotal roles. One summer, Lenny
and Sheldon end up as bellhops at the famed Grossinger’s Resort in the Catskills, where Lenny practices standup comedy
amid the glamorous, bustling atmosphere.An underlying seriousness lies at the heart of all of this intrigue, hilarity a
nd fun. Sheldon, Abe, Mirabelle and Lenny, all Jewish, must confront the many faces of antisemitism during the turbulen
t years of World War II. Miller weaves in a multitude of historical details, including reports of the horrors in Europe
and America’s reluctance to intervene.The ending of How to Find Your Way in the Dark is nothing short of brilliant, tyi
ng up a variety of loose ends while making a powerful statement about the need to fully recognize and address antisemit
ism. Readers are left with much to ponder, including life’s many uncertainties and cruel twists of fate. Despite these
unhappy truths, we are also left with the uplifting wisdom of Lenny’s urgent prayer: “Dear God, give me the strength to
be joyful.”★ How to Find Your Way in the DarkByDerek B. MillerHMH$26.00ISBN9780358269601Published07/27/2021Fiction / C
oming of Age'
```

In [ ]:
```

```

In [225]:
```
page=requests.get("https://bookpage.com/reviews/26469-derek-b-miller-how-to-find-your-way-dark-fiction#.YSD3iflR1PY")
page
```

Out[225]:
```
<Response [200]>
```

In [226]:
```
soup=BeautifulSoup(page.content)
soup
```

Out[226]:
```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - How to Find Your Way in the Dark by Derek B. Miller | BookPage</title>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
ment about the need to fully address antisemitism." name="description"/>
<link href="https://bookpage.com/reviews/26469-derek-b-miller-how-to-find-your-way-dark-fiction" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - How to Find Your Way in the Dark by Derek B. Miller" property="og:title"/>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
ment about the need to fully address antisemitism." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26469-derek-b-miller-how-to-find-your-way-dark-fiction" property="og:ur
l"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - How to Find Your Way in the Dark by Derek B. Miller" name="twitter:title"/>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
```

In [49]:
```
print(soup.prettify)
```

```
<bound method Tag.prettify of <!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - How to Find Your Way in the Dark by Derek B. Miller | BookPage</title>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
ment about the need to fully address antisemitism." name="description"/>
<link href="https://bookpage.com/reviews/26469-derek-b-miller-how-to-find-your-way-dark-fiction" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - How to Find Your Way in the Dark by Derek B. Miller" property="og:title"/>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
ment about the need to fully address antisemitism." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26469-derek-b-miller-how-to-find-your-way-dark-fiction" property="og:ur
l"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - How to Find Your Way in the Dark by Derek B. Miller" name="twitter:title"/>
<meta content="The ending of How to Find Your Way in the Dark is nothing short of brilliant, making a powerful state
```

In [50]:
```
book_name=soup.find("h1",class_="italic")
book_name
```

Out[50]:
```
<h1 class="italic">
<span style="font-style:normal;">★ </span>How to Find Your Way in the Dark
</h1>
```

In [51]:
```
book_name.text.replace("\n","")
```

Out[51]:
```
'★ How to Find Your Way in the Dark'
```

In [52]:

```
author_name=soup.find("h4",class_="sans")
author_name
```

Out[52]:

```
<h4 class="sans">
Derek B. Miller
</h4>
```

In [53]:

```
author_name.text.replace("\n","")
```

Out[53]:

```
'Derek B. Miller'
```

In [54]:

```
genre=soup.find("p",class_="genre-links")
genre
```

Out[54]:

```
<p class="genre-links">
<a href="/search?book_genre=fiction">Fiction</a>
 /
<a href="/search?book_genre=coming_of_age">Coming of Age</a>
</p>
```

In [55]:

```
genre.text.replace("\n","")
```

Out[55]:

```
'Fiction / Coming of Age'
```

In [56]:
```
book_review=soup.find("div",class_="article-body")
book_review
```

Out[56]:
```
<div class="article-body">
<p>
</p><p>Prepare for surprises galore in <b>How to Find Your Way in the Dark</b>, a rollicking novel that begins with a l
onely truck ride in New England in 1938 and follows its characters through a decade of fascinating history. Just when y
ou think the story is heading one way, it veers in another, completely unexpected direction.</p>
<p>Twelve-year-old Sheldon Horowitz and his father are driving home from Hartford, Connecticut, to Whately, Massachuset
ts, after honoring the one-year anniversary of Sheldon’s mother’s death. She and her sister died in a horrific movie th
eater fire in Hartford. And as if that isn’t enough tragedy for the novel’s first 13 pages, a truck purposely forces Sh
eldon and his father’s car off the road during their return trip, and Sheldon’s father dies.</p>
<p>Readers of Derek B. Miller’s award-winning thriller, <a href="https://bookpage.com/reviews/15318-derek-b-miller-shel
dons-sense-snow-fiction#.YN9gC5NKjUI"><i>Norwegian by Night</i></a>, will recognize Sheldon as that novel’s 82-year-old
protagonist. As a Tom Sawyer-like boy in <b>How to Find Your Way in the Dark</b>, Sheldon is determined to make sense o
f his double tragedies, and his attempts to do so take the reader on one hell of a ride. As he seeks out the leering, m
ustached truck driver who killed his father, his quest leads him straight into danger—think mobsters, guns and jewel th
efts.</p>
<p>Miller has crafted a wide-ranging, years-spanning yet tightly structured plot, and he excels at placing memorable ch
aracters in unusual circumstances. Sheldon is joined in his adventures by his two older cousins, Abe and Mirabelle, and
his best friend, Lenny, all of whom play pivotal roles. One summer, Lenny and Sheldon end up as bellhops at the famed G
rossinger’s Resort in the Catskills, where Lenny practices standup comedy amid the glamorous, bustling atmosphere.</p>
<p>An underlying seriousness lies at the heart of all of this intrigue, hilarity and fun. Sheldon, Abe, Mirabelle and L
enny, all Jewish, must confront the many faces of antisemitism during the turbulent years of World War II. Miller weave
s in a multitude of historical details, including reports of the horrors in Europe and America’s reluctance to interven
e.</p>
<p>The ending of <b>How to Find Your Way in the Dark</b> is nothing short of brilliant, tying up a variety of loose end
s while making a powerful statement about the need to fully recognize and address antisemitism. Readers are left with m
uch to ponder, including life’s many uncertainties and cruel twists of fate. Despite these unhappy truths, we are also
left with the uplifting wisdom of Lenny’s urgent prayer: “Dear God, give me the strength to be joyful.”</p>
</div>
```

In [57]:
```
book_review.text.replace("\n","")
```

Out[57]:
```
'Prepare for surprises galore in How to Find Your Way in the Dark, a rollicking novel that begins with a lonely truck r
ide in New England in 1938 and follows its characters through a decade of fascinating history. Just when you think the
story is heading one way, it veers in another, completely unexpected direction.Twelve-year-old Sheldon Horowitz and his
father are driving home from Hartford, Connecticut, to Whately, Massachusetts, after honoring the one-year anniversary
of Sheldon’s mother’s death. She and her sister died in a horrific movie theater fire in Hartford. And as if that isn’t
enough tragedy for the novel’s first 13 pages, a truck purposely forces Sheldon and his father’s car off the road durin
g their return trip, and Sheldon’s father dies.Readers of Derek B. Miller’s award-winning thriller, Norwegian by Night,
will recognize Sheldon as that novel’s 82-year-old protagonist. As a Tom Sawyer-like boy in How to Find Your Way in the
Dark, Sheldon is determined to make sense of his double tragedies, and his attempts to do so take the reader on one hel
l of a ride. As he seeks out the leering, mustached truck driver who killed his father, his quest leads him straight in
to danger—think mobsters, guns and jewel thefts.Miller has crafted a wide-ranging, years-spanning yet tightly structure
d plot, and he excels at placing memorable characters in unusual circumstances. Sheldon is joined in his adventures by
his two older cousins, Abe and Mirabelle, and his best friend, Lenny, all of whom play pivotal roles. One summer, Lenny
and Sheldon end up as bellhops at the famed Grossinger’s Resort in the Catskills, where Lenny practices standup comedy
amid the glamorous, bustling atmosphere.An underlying seriousness lies at the heart of all of this intrigue, hilarity a
nd fun. Sheldon, Abe, Mirabelle and Lenny, all Jewish, must confront the many faces of antisemitism during the turbulen
t years of World War II. Miller weaves in a multitude of historical details, including reports of the horrors in Europe
and America’s reluctance to intervene.The ending of How to Find Your Way in the Dark is nothing short of brilliant, tyi
ng up a variety of loose ends while making a powerful statement about the need to fully recognize and address antisemit
ism. Readers are left with much to ponder, including life’s many uncertainties and cruel twists of fate. Despite these
unhappy truths, we are also left with the uplifting wisdom of Lenny’s urgent prayer: “Dear God, give me the strength to
be joyful.”'
```

In [58]:
```
page=requests.get("https://bookpage.com/reviews/26490-rodrigo-garcia-farewell-to-gabo-mercedes-nonfiction#.YSEHg_lR1PY")
page
```

Out[58]:
```
<Response [200]>
```

In [59]:
```
soup=BeautifulSoup(page.content)
soup
```

Out[59]:
```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - A Farewell to Gabo and Mercedes by Rodrigo García | BookPage</title>
<meta content="When Gabriel García Márquez began his long slide toward dementia, his son began taking notes for this
intimate, endearing tribute to his late parents." name="description"/>
<link href="https://bookpage.com/reviews/26490-rodrigo-garcia-farewell-to-gabo-mercedes-nonfiction" rel="canonical"/
>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - A Farewell to Gabo and Mercedes by Rodrigo García" property="og:title"/>
<meta content="When Gabriel García Márquez began his long slide toward dementia, his son began taking notes for this
intimate, endearing tribute to his late parents." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26490-rodrigo-garcia-farewell-to-gabo-mercedes-nonfiction" property="og:
url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - A Farewell to Gabo and Mercedes by Rodrigo García" name="twitter:title"/>
```

In [60]:
```
print(soup.prettify)
```

```
<bound method Tag.prettify of <!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - A Farewell to Gabo and Mercedes by Rodrigo García | BookPage</title>
<meta content="When Gabriel García Márquez began his long slide toward dementia, his son began taking notes for this
intimate, endearing tribute to his late parents." name="description"/>
<link href="https://bookpage.com/reviews/26490-rodrigo-garcia-farewell-to-gabo-mercedes-nonfiction" rel="canonical"/
>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - A Farewell to Gabo and Mercedes by Rodrigo García" property="og:title"/>
<meta content="When Gabriel García Márquez began his long slide toward dementia, his son began taking notes for this
intimate, endearing tribute to his late parents." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26490-rodrigo-garcia-farewell-to-gabo-mercedes-nonfiction" property="og:
url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - A Farewell to Gabo and Mercedes by Rodrigo García" name="twitter:title"/>
<meta content="When Gabriel García Márquez began his long slide toward dementia, his son began taking notes for this
```

In [61]:
```
book_name=soup.find("h1",class_="italic")
book_name
```

Out[61]:
```
<h1 class="italic">
<span style="font-style:normal;">★ </span>A Farewell to Gabo and Mercedes
</h1>
```

In [62]:
```
book_name.text.replace("\n","")
```

Out[62]:
```
'★ A Farewell to Gabo and Mercedes'
```

In [63]:
```
author_name=soup.find("h4",class_="sans")
author_name
```

Out[63]:
```
<h4 class="sans">
Rodrigo García
</h4>
```

In [64]:
```
author_name.text.replace("\n","")
```

Out[64]:
```
'Rodrigo García'
```

In [65]:
```
genre=soup.find("p",class_="genre-links")
genre
```

Out[65]:
```
<p class="genre-links">
<a href="/search?book_genre=nonfiction">Nonfiction</a>
 /
<a href="/search?book_genre=memoir">Memoir</a>
 /
<a href="/search?book_genre=family_relationships">Family &amp; Relationships</a>
</p>
```

In [66]:
```
genre.text.replace("\n","")
```

Out[66]:
```
'Nonfiction / Memoir / Family & Relationships'
```

In [67]:

```
book_review=soup.find("div",class_="article-body")
book_review
```

Out[67]:

```
<div class="article-body">
<p>
</p><p>Rodrigo García is a film and television director, writer, cinematographer and son of the late Nobel winner Gabri
el García Márquez, affectionately known as Gabo, author of <i>One Hundred Years of Solitude</i> and <i>Love in the Time
of Cholera</i>. When García’s world-famous father began his long slide toward dementia, García began taking notes. “Wri
ting about the death of loved ones must be about as old as writing itself, and yet the inclination to do it instantly t
ies me up in knots,” he writes. “I am appalled that I am thinking of taking notes, ashamed as I take notes, disappointe
d in myself as I revise notes.”</p>
<p>All who have loved García Márquez’s works will rejoice that his son overcame that angst, dutifully waiting until aft
er his father’s death in 2014 and his mother’s death in 2020 to publish his intimate, endearing tribute, <b>A Farewell
to Gabo and Mercedes: A Son’s Memoir of Gabriel García Márquez and Mercedes Barcha</b>. García’s notes, acutely observa
tional, are simultaneously infused with love, respect and the pain of loss. He admits that his relationships with his p
arents were complicated. Their lives had public, private and even secret components, and García frets about crossing li
nes that might leave his parents helplessly exposed. Still, from his dying father’s bedside in Mexico City to his last
moment with his mother (shared digitally, as COVID-19 prevented him from traveling), García is a guardian of their dign
ity. </p>
<p>Yet this memoir’s details are indeed intimate. We're ushered into García Márquez's study as he works, until the reno
wned author slowly realizes he no longer can. García’s mother rises above her grief, insisting that she is a woman, not
a widow, as she entertains the flow of mourning guests from around the globe—even the complete stranger who manages to
con her out of quite a bit of cash. We follow García into the crematorium as he gazes upon his father for the last tim
e, tempering that blow with the thought that García Márquez might have enjoyed flirting with the funeral worker who gav
e his body a little makeup, a final flourish on his way out.</p>
<p>Fittingly, García begins each chapter with an excerpt from one of his father’s works, and it’s this connection betwe
en life and art that holds this intense memoir together. As one epigraph from <i>Love in the Time of Cholera</i> puts i
t, “he was overwhelmed by the belated suspicion that it is life, more than death, that has no limits.”</p>
</div>
```

In [68]:
```
book_review.text.replace("\n","")
```

Out[68]:
```
"Rodrigo García is a film and television director, writer, cinematographer and son of the late Nobel winner Gabriel Gar
cía Márquez, affectionately known as Gabo, author of One Hundred Years of Solitude and Love in the Time of Cholera. Whe
n García’s world-famous father began his long slide toward dementia, García began taking notes. “Writing about the deat
h of loved ones must be about as old as writing itself, and yet the inclination to do it instantly ties me up in knot
s,” he writes. “I am appalled that I am thinking of taking notes, ashamed as I take notes, disappointed in myself as I
revise notes.”All who have loved García Márquez’s works will rejoice that his son overcame that angst, dutifully waitin
g until after his father’s death in 2014 and his mother’s death in 2020 to publish his intimate, endearing tribute, A F
arewell to Gabo and Mercedes: A Son’s Memoir of Gabriel García Márquez and Mercedes Barcha. García’s notes, acutely obs
ervational, are simultaneously infused with love, respect and the pain of loss. He admits that his relationships with h
is parents were complicated. Their lives had public, private and even secret components, and García frets about crossin
g lines that might leave his parents helplessly exposed. Still, from his dying father’s bedside in Mexico City to his l
ast moment with his mother (shared digitally, as COVID-19 prevented him from traveling), García is a guardian of their
dignity.\xa0Yet this memoir’s details are indeed intimate. We're ushered into García Márquez's study as he works, until
the renowned author slowly realizes he no longer can. García’s mother rises above her grief, insisting that she is a wo
man, not a widow, as she entertains the flow of mourning guests from around the globe—even the complete stranger who ma
nages to con her out of quite a bit of cash. We follow García into the crematorium as he gazes upon his father for the
last time, tempering that blow with the thought that García Márquez might have enjoyed flirting with the funeral worker
who gave his body a little makeup, a final flourish on his way out.Fittingly, García begins each chapter with an excerp
t from one of his father’s works, and it’s this connection between life and art that holds this intense memoir togethe
r. As one epigraph from Love in the Time of Cholera puts it, “he was overwhelmed by the belated suspicion that it is li
fe, more than death, that has no limits.”"
```

In [69]:
```
page=requests.get("https://bookpage.com/reviews/26377-jason-mott-hell-book-fiction#.YSEK_flR1PY")
page
```

Out[69]:
```
<Response [200]>
```

In [70]:
```
soup=BeautifulSoup(page.content)
soup
```

Out[70]:
```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - Hell of a Book by Jason Mott | BookPage</title>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on
the same page." name="description"/>
<link href="https://bookpage.com/reviews/26377-jason-mott-hell-book-fiction" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - Hell of a Book by Jason Mott" property="og:title"/>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on
the same page." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26377-jason-mott-hell-book-fiction" property="og:url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - Hell of a Book by Jason Mott" name="twitter:title"/>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on
the same page." name="twitter:description"/>
```

In [71]:

```
print(soup.prettify)
```

```
<bound method Tag.prettify of <!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>Book Review - Hell of a Book by Jason Mott | BookPage</title>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on the same page." name="description"/>
<link href="https://bookpage.com/reviews/26377-jason-mott-hell-book-fiction" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="Book Review - Hell of a Book by Jason Mott" property="og:title"/>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on the same page." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/reviews/26377-jason-mott-hell-book-fiction" property="og:url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="Book Review - Hell of a Book by Jason Mott" name="twitter:title"/>
<meta content="Jason Mott’s fourth novel can break your heart and make you laugh out loud at the same time, often on the same page." name="twitter:description"/>
```

In [72]:

```
book_name=soup.find("h1",class_="italic")
book_name
```

Out[72]:

```
<h1 class="italic">
<span style="font-style:normal;">★ </span>Hell of a Book
</h1>
```

In [73]:

```
book_name.text.replace("\n","")
```

Out[73]:

```
'★ Hell of a Book'
```

In [74]:
```
author_name=soup.find("h4",class_="sans")
author_name
```

Out[74]:
```
<h4 class="sans">
Jason Mott
</h4>
```

In [75]:
```
author_name.text.replace("\n","")
```

Out[75]:
```
'Jason Mott'
```

In [76]:
```
genre=soup.find("p",class_="genre-links")
genre
```

Out[76]:
```
<p class="genre-links">
<a href="/search?book_genre=fiction">Fiction</a>
 /
<a href="/search?book_genre=literary_fiction">Literary Fiction</a>
</p>
```

In [77]:
```
genre.text.replace("\n","")
```

Out[77]:
```
'Fiction / Literary Fiction'
```

In [78]:

```
book_review=soup.find("div",class_="article-body")
book_review
```

Out[78]:

```
<div class="article-body">
<p>
</p><p>There’s an intimacy to Jason Mott’s fiction, retained even when the scope of his narrative widens. But even by t
hese standards, his fourth novel is a uniquely tight, personal story that digs into deeply emotional territory. Through
two interwoven storylines unfolding in a witty, often devastatingly incisive style, <b>Hell of a Book</b> is a journey
into the heart of a very particular American experience, one that far too many don’t live to tell.</p>
<p><b>Hell of a Book</b> is named for the novel written by Mott’s protagonist, an unnamed author embarking on a booze-f
ueled book tour across the United States, hopping from hotel room to hotel room and interview after interview. But the
author is less keen to talk about his book than about the Kid, a mysterious and possibly imaginary Black child who has
appeared by the author’s side and now follows him everywhere. </p>
<p>As the author and the Kid get to know each other, Mott intersperses their tale with that of Soot, a Black boy who en
dures bullying in his small town for the color of his skin, and whose childhood seems to be on a tragic and all-too-com
mon trajectory.</p>
<p>You may think you see where these two stories are headed, where they will converge and knit together, and what they
will have to say at the end, but you don’t. And even if you could, Mott’s bittersweet, remarkably nimble novel would st
ill keep you turning the pages. </p>
<p><b>Hell of a Book</b> is a masterwork of balance, as Mott navigates the two narratives and their delicate tonal dist
inctions. A surrealist feast of imagination that’s brimming with very real horrors, frustrations and sorrows, it can br
eak your heart and make you laugh out loud at the same time, often on the same page. This is an achievement of American
fiction that rises to meet this particular moment with charm, wisdom and truth.</p>
<p> </p>
<p><i>ALSO IN BOOKPAGE: </i><a href="http://bkpg.it/6Rs"><b>Hell of a Book</b><i> author Jason Mott discusses the new c
onfidence he’s found as a writer. “I hope that it leads to more creative exploration and new paths of storytelling in t
he future.”</i></a></p>
</div>
```

In [79]:
```
book_review.text.replace("\n","")
```

Out[79]:
```
'There’s an intimacy to Jason Mott’s fiction, retained even when the scope of his narrative widens. But even by these s
tandards, his fourth novel is a uniquely tight, personal story that digs into deeply emotional territory. Through two i
nterwoven storylines unfolding in a witty, often devastatingly incisive style, Hell of a Book is a journey into the hea
rt of a very particular American experience, one that far too many don’t live to tell.Hell of a Book is named for the n
ovel written by Mott’s protagonist, an unnamed author embarking on a booze-fueled book tour across the United States, h
opping from hotel room to hotel room and interview after interview. But the author is less keen to talk about his book
than about the Kid, a mysterious and possibly imaginary Black child who has appeared by the author’s side and now follo
ws him everywhere.\xa0As the author and the Kid get to know each other, Mott intersperses their tale with that of Soot,
a Black boy who endures bullying in his small town for the color of his skin, and whose childhood seems to be on a trag
ic and all-too-common trajectory.You may think you see where these two stories are headed, where they will converge and
knit together, and what they will have to say at the end, but you don’t. And even if you could, Mott’s bittersweet, rem
arkably nimble novel would still keep you turning the pages.\xa0Hell of a Book is a masterwork of balance, as Mott navi
gates the two narratives and their delicate tonal distinctions. A surrealist feast of imagination that’s brimming with
very real horrors, frustrations and sorrows, it can break your heart and make you laugh out loud at the same time, ofte
n on the same page. This is an achievement of American fiction that rises to meet this particular moment with charm, wi
sdom and truth.\xa0ALSO IN BOOKPAGE: Hell of a Book author Jason Mott discusses the new confidence he’s found as a writ
er. “I hope that it leads to more creative exploration and new paths of storytelling in the future.”'
```

In [80]:
```
page=requests.get("https://bookpage.com/features/26522-i-hate-how-much-i-love-you-romance#.YSENtPlR1PY")
page
```

Out[80]:
```
<Response [200]>
```

In [81]:

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

Out[81]:

```
<!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>I hate how much I love you | BookPage</title>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." name="description"/>
<link href="https://bookpage.com/features/26522-i-hate-how-much-i-love-you-romance" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="I hate how much I love you" property="og:title"/>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/features/26522-i-hate-how-much-i-love-you-romance" property="og:url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="I hate how much I love you" name="twitter:title"/>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." name="twitter:description"/>
```

In [82]:
```
print(soup.prettify)
```

```
<bound method Tag.prettify of <!DOCTYPE html>
<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://opengraphprotocol.org/schema/">
<head>
<title>I hate how much I love you | BookPage</title>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." name="description"/>
<link href="https://bookpage.com/features/26522-i-hate-how-much-i-love-you-romance" rel="canonical"/>
<meta content="BookPage.com" property="og:site_name"/>
<meta content="I hate how much I love you" property="og:title"/>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." property="og:description"/>
<meta content="article" property="og:type"/>
<meta content="https://bookpage.com/features/26522-i-hate-how-much-i-love-you-romance" property="og:url"/>
<meta content="/images/medium/missing.png" property="og:image"/>
<meta content="summary" name="twitter:card"/>
<meta content="@bookpage" name="twitter:site"/>
<meta content="I hate how much I love you" name="twitter:title"/>
<meta content="This pair of enemies-to-lovers romances takes readers on a bicoastal journey, from classic Los Angele
s showbiz glitz to sleek Manhattan high-rises." name="twitter:description"/>
```

In [83]:
```
book_name=soup.find("h1",class_="italic")
book_name
```

Out[83]:
```
<h1 class="italic">
I hate how much I love you
</h1>
```

In [84]:
```
book_name.text.replace("\n","")
```

Out[84]:
```
'I hate how much I love you'
```

In [85]:

```
author_name=soup.find("p",class_="sans bold author-info")
author_name
```

Out[85]:

```
<p class="sans bold author-info">
Feature by
<a href="/search?contributor_id=1124">Dolly R. Sickles</a>
</p>
```

In [86]:

```
author_name.text.replace("\n"," ")
```

Out[86]:

```
' Feature by Dolly R. Sickles '
```

In [87]:

```
genre=soup.find("p",class_="genre-links sans")
genre
```

Out[87]:

```
<p class="genre-links sans">
<a href="/search?book_genre=romance">Romance</a>
 /
<a href="/search?book_genre=contemporary_romance">Contemporary Romance</a>
</p>
```

In [88]:

```
genre.text.replace("\n","")
```

Out[88]:

```
'Romance / Contemporary Romance'
```

In [89]:

```
book_review=soup.find("div",class_="article-body")
book_review
```

Out[89]:

```
<div class="article-body">
<p>
</p><p>This pair of enemies-to-lovers romances takes readers on a bicoastal journey from classic Los Angeles showbiz gl
itz to sleek Manhattan high-rises.</p>
<p>LA-based YA author Bridget Morrissey hits the ground running with her first adult novel, <strong>Love Scenes</strong
>. Morrissey’s Angeleno bona fides shine through in this sweet and thoughtful rom-com that tackles real-world problems
of addiction, sobriety and second chances with wit and wisdom.</p>
<p>Told in first person, <strong>Love Scenes</strong> follows Sloane Ford, actor and second-generation Hollywood royalt
y, as she contemplates retiring from the industry following the disastrous demise of her longtime gig on a TV crime pro
cedural. But she’s quickly roped in to being the consulting producer of a new movie written by her stepfather, directed
by her sister, co-starring her mother and starring Joseph Donovan, a fellow actor she never wanted to work with again a
fter his difficult behavior during the one and only movie they made together.</p>
<p>Sloane’s gigantic family, with parents and siblings and steps galore, has the potential to overwhelm everything else
in this romance—particularly since, in true Hollywood fashion, everyone is involved in showbiz. But they’re well rounde
d and secondary to the real focus of the story: Sloane and Joe. Despite the professional, generational and financial pr
ivilege Sloane could easily fall back on, she works hard and recognizes all the ways she’s been given a leg up in her c
areer. She also recognizes, eventually, Joe's genuine effort to make amends. He was an emotional wreck and struggling w
ith alcoholism when they last worked together, but now he’s sober and dedicated to keeping it that way. What follows is
real emotional growth, true friendship and a satisfying love story.</p>
<p>Lauren Layne proves once again that she’s the queen of contemporary New York City romance with <strong>To Sir, With
Love</strong>. Her breezy dialogue and delightful characters will fully immerse readers in this dreamy and sophisticate
d love story.</p>
<p>It’s easy to connect with Gracie Cooper right out of the gate. She’s an earnest, hopeful character anyone would love
to call a friend. She’s crazy about her best friend’s baby, names the pigeons she feeds in the park, blurts out everyth
ing and blindly gives herself over to the attraction she feels to “Sir,” the mystery man she’s been chatting with under
the name “Lady” through a dating app called MysteryMate. All of that, and she’s set aside her own dreams to keep her la
te father’s champagne shop, Bubbles &amp; More, in business.</p>
<p>Gracie thinks she’s fallen in love with Sir. But if she has, how can she also be so drawn to Sebastian Andrews, the
man whose company wants to buy out Bubbles &amp; More’s lease? At first glance, Sebastian is the villain in this fairy
tale, but the more Gracie learns about the businessman, the more the goodness in her recognizes the goodness in him. Se
bastian is kind and supportive and recognizes the community impact his business decisions could cause. And the more tim
e they spend in one another’s company, the more he suspects the possibility of Gracie being the Lady to his Sir.</p>
<p>This hopeful, happy love story sparkles with fairy dust, even as Layne makes it clear how high the professional and
personal stakes are for her main couple. The superb characterization of Gracie and Sebastian and the parallel journeys
they take toward one another make <strong>To Sir, With Love</strong> a wonderfully satisfying romance.</p>
</div>
```

In [90]:
```
book_review.text.replace("\n","")
```

Out[90]:
```
"This pair of enemies-to-lovers romances takes readers on a bicoastal journey\xa0from classic Los Angeles showbiz glitz
to sleek Manhattan high-rises.LA-based YA author Bridget Morrissey hits the ground running with her first adult novel,
Love Scenes. Morrissey’s Angeleno bona fides shine through in this sweet and thoughtful rom-com that tackles real-world
problems of addiction, sobriety and second chances with wit and wisdom.Told in first person, Love Scenes follows Sloane
Ford, actor and second-generation Hollywood royalty, as she contemplates retiring from the industry following the disas
trous demise of her longtime gig on a TV crime procedural. But she’s quickly roped in to being the consulting producer
of\xa0a new movie written by her stepfather, directed by her sister, co-starring her mother and starring Joseph Donova
n, a fellow actor she never wanted to work with again after his difficult behavior during the one and only movie they m
ade together.Sloane’s gigantic family, with parents and siblings and steps galore, has the potential to overwhelm every
thing else in this romance—particularly since, in true Hollywood fashion, everyone is involved in showbiz. But they’re
well rounded and secondary to the real focus of the story: Sloane and Joe. Despite the professional, generational and f
inancial privilege Sloane could easily fall back on, she works hard and recognizes all the ways she’s been given\xa0a l
eg up in her career. She also recognizes, eventually, Joe's genuine effort to make\xa0amends. He was an emotional wreck
and struggling with alcoholism when they last worked together, but now he’s\xa0sober and dedicated to keeping it that w
ay. What follows is real emotional growth, true friendship and a satisfying love story.Lauren Layne proves once again t
hat she’s the queen of contemporary New York City romance with To Sir, With Love. Her breezy dialogue and delightful ch
aracters will fully immerse readers in\xa0this dreamy and sophisticated love story.It’s easy to connect with Gracie Coo
per right out of the gate. She’s an earnest, hopeful character\xa0anyone would love to call a friend. She’s crazy about
her best friend’s baby, names the pigeons she feeds in the park, blurts out everything and blindly gives herself over t
o the attraction she feels to “Sir,” the mystery man she’s been chatting with under the name\xa0“Lady” through a dating
app called MysteryMate.\xa0All of that, and she’s set aside her own dreams to keep her late father’s champagne shop, Bu
bbles & More, in business.Gracie thinks she’s fallen in love with Sir. But if she has, how can she also be so drawn to
Sebastian Andrews, the man whose company wants to buy out Bubbles & More’s lease? At first glance, Sebastian is the vil
lain in this fairy tale, but the more Gracie learns about the businessman, the more the goodness in her recognizes the
goodness in him. Sebastian is kind and supportive\xa0and recognizes the community impact his business decisions could\x
a0cause. And the more time they spend in one another’s company, the more he suspects the possibility of Gracie being th
e Lady to his Sir.This hopeful, happy love story sparkles with fairy dust, even as Layne makes it clear how high the pr
ofessional and personal stakes are for her main couple. The superb characterization of Gracie and Sebastian and the\xa0
parallel journeys they take toward\xa0one another make To Sir, With Love a wonderfully satisfying romance."
```

```
5. Write a python program to scrape cricket rankings from ‘www.icc-cricket.com’. You have to scrape:
i) Top 10 ODI teams in men’s cricket along with the records for matches, points andrating.
```

In [91]:

```
page=requests.get("https://www.icc-cricket.com/rankings/mens/team-rankings/odi")
page
```

Out[91]:

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

In [92]:

```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[92]:

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="ICC Ranking for ODI teams International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official International Cricket Council ranking for One Day International (ODI) cricket teams. Discove
r latest ICC rankings table, predict upcoming matches, see points and ratings for all teams." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official International Cricket Council ranking for One Day International (ODI) cricket teams. Discove
r latest ICC rankings table, predict upcoming matches, see points and ratings for all teams." name="twitter:descript
ion"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="ICC Ranking for ODI teams International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>ICC Ranking for ODI teams International Cricket Council</title>
<meta content="Official International Cricket Council ranking for One Day International (ODI) cricket teams. Discove
r latest ICC rankings table, predict upcoming matches, see points and ratings for all teams." property="og:descripti
on"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
```

In [93]:

```
table=soup.find("table",class_="table")
table
```

Out[93]:

```
<table class="table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell rankings-table__team">
<span class="u-hide-mobile">Team</span>
<span class="u-show-mobile">T</span>
</th>
<th class="table-head__cell u-center-text">
<span class="u-hide-mobile">Matches</span>
<span class="u-show-mobile">M</span>
</th>
<th class="table-head__cell u-center-text">
<span class="u-hide-mobile">Points</span>
<span class="u-show-mobile">P</span>
</th>
<th class="table-head__cell u-text-right rating">
<span class="u-hide-mobile">Rating</span>
<span class="u-show-mobile">R</span>
</th>
```

In [94]:

```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n","").strip()
    headers.append(title)
```

In [95]:

```
df= pd.DataFrame(columns= headers)
df
```

Out[95]:

Pos	TeamT	MatchesM	PointsP	RatingR

```
In [96]: for row in table.find_all('tr')[1:]:
             data=row.find_all("td")
             row_data=[td.text.replace("\n","").strip() for td in data]
             length=len(df)
             df.loc[length]=row_data
             #df.to_csv("row_data")
```

In [97]:
```
df=pd.read_csv("top_teams_data")
df
```

Out[97]:
```
    Unnamed: 0  Pos               TeamT  MatchesM  PointsP  RatingR
 0           0    1       New ZealandNZ        17    2,054      121
 1           1    2          EnglandENG        32    3,793      119
 2           2    3        AustraliaAUS        28    3,244      116
 3           3    4            IndiaIND        32    3,624      113
 4           4    5       South AfricaSA        22    2,267      103
 5           5    6         PakistanPAK        27    2,524       93
 6           6    7       BangladeshBAN        29    2,639       91
 7           7    8       West IndiesWI        30    2,523       84
 8           8    9         Sri LankaSL        29    2,303       79
 9           9   10      AfghanistanAFG        17    1,054       62
10          10   11      NetherlandsNED         7      336       48
11          11   12          IrelandIRE        23    1,067       46
12          12   13         ZimbabweZIM        17      646       38
13          13   14         ScotlandSCO         7      258       37
14          14   15             OmanOMA         7      240       34
15          15   16            NepalNEP         5      119       24
16          16   17              UAEUAE         9      190       21
17          17   18          NamibiaNAM         6       97       16
18          18   19     United StatesUSA         8       93       12
19          19   20  Papua New GuineaPNG         5        0        0
```

In [ ]:

In [98]:
```
df=pd.DataFrame(data=df,columns=headers)
df.head(10)
```

Out[98]:

	Pos	TeamT	MatchesM	PointsP	RatingR
0	1	New ZealandNZ	17	2,054	121
1	2	EnglandENG	32	3,793	119
2	3	AustraliaAUS	28	3,244	116
3	4	IndiaIND	32	3,624	113
4	5	South AfricaSA	22	2,267	103
5	6	PakistanPAK	27	2,524	93
6	7	BangladeshBAN	29	2,639	91
7	8	West IndiesWI	30	2,523	84
8	9	Sri LankaSL	29	2,303	79
9	10	AfghanistanAFG	17	1,054	62

ii) Top 10 ODI Batsmen in men along with the records of their team and rating.

In [99]:
```
page=requests.get("https://www.icc-cricket.com/rankings/mens/player-rankings/odi/batting")
page
```

Out[99]:
```
<Response [200]>
```

In [100]:
```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[100]:
```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="Live Cricket Scores &amp; News International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="twitter:description"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="Live Cricket Scores &amp; News International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>Live Cricket Scores &amp; News International Cricket Council</title>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." property="og:description"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
<script>
```

In [101]:

```
table=soup.find("table",class_="table rankings-table")
table
```

Out[101]:

```
<table class="table rankings-table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell">Player</th>
<th class="table-head__cell">Team</th>
<th class="table-head__cell table-head__cell--rating">Rating</th>
<th class="table-head__cell u-text-right u-hide-phablet">Career Best Rating</th>
</tr>
</thead>
<tr class="rankings-block__banner">
<td class="rankings-block__position">1</td>
<td class="rankings-block__top-player-container">
<div class="u-flex-center u-text-left">
<div class="rankings-block__player-image-container rankings-block__player-image-container--large">
<a href="/rankings/mens/player-rankings/2759">
<img alt="Player Image" class="rankings-block__player-image image-missing" data-error="true" data-player="2759" data
-size="130x150" data-widget="player-image" src="//resources.pulse.icc-cricket.com/players/130x150/Photo-Missing.pn
g"/>
```

In [102]:

```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n"," ").strip()
    headers.append(title)
```

In [103]:

```
df= pd.DataFrame(columns= headers)
df
```

Out[103]:

Pos	Player	Team	Rating	Career Best Rating

In [104]:

```
for row in table.find_all('tr')[1:]:
    data=row.find_all("td")
    row_data=[td.text.replace("\n","").strip() for td in data]
    length=len(df)
    df.loc[length]=row_data
    #df.to_csv("top_batsmen_data")
```

In [105]:

```
df=pd.read_csv("top_batsmen_data")
df
```

Out[105]:

	Unnamed: 0	Pos	Player	Team	Rating	Career Best Rating
0	0	1	Babar Azam	PAK	873	873 v England, 13/07/2021
1	1	2	Virat Kohli	IND	844	911 v England, 12/07/2018
2	2	3	Rohit Sharma	IND	813	885 v Sri Lanka, 06/07/2019
3	3	4	Ross Taylor	NZ	801	841 v Bangladesh, 05/06/2019
4	4	5	Aaron Finch	AUS	779	798 v England, 25/06/2019
...	...	...	...	...	...	...
95	95	96 This player has moved down in the ra...	Moeen Ali	ENG	404	522 v West Indies, 27/09/2017
96	96	97 This player has moved down in the ra...	Wanindu De Silva	SL	403	414 v England, 04/07/2021
97	97	This player has moved down in the rankings sin...	Chris Woakes	ENG	403	481 v Australia, 26/01/2018
98	98	99 This player has moved down in the ra...	Simi Singh	IRE	402	402 v South Africa, 16/07/2021
99	99	100 This player has moved down in the r...	Matthew Cross	SCO	398	427 v Oman, 18/08/2019

100 rows × 6 columns

In [106]:
```
df=pd.DataFrame(data=df,columns=headers)
df.head(10)
```

Out[106]:

	Pos	Player	Team	Rating	Career Best Rating
0	1	Babar Azam	PAK	873	873 v England, 13/07/2021
1	2	Virat Kohli	IND	844	911 v England, 12/07/2018
2	3	Rohit Sharma	IND	813	885 v Sri Lanka, 06/07/2019
3	4	Ross Taylor	NZ	801	841 v Bangladesh, 05/06/2019
4	5	Aaron Finch	AUS	779	798 v England, 25/06/2019
5	6	Jonny Bairstow	ENG	775	796 v India, 26/03/2021
6	7 This player has moved up in the ranki...	David Warner	AUS	762	880 v Pakistan, 26/01/2017
7	8 This player has moved up in the ranki...	Quinton de Kock	SA	758	813 v Sri Lanka, 10/03/2019
8	This player has moved down in the rankings sin...	Shai Hope	WI	758	808 v Bangladesh, 17/05/2019
9	10	Kane Williamson	NZ	754	799 v India, 09/07/2019

iii) Top 10 ODI bowlers along with the records of their team and rating.

In [107]:
```
page=requests.get("https://www.icc-cricket.com/rankings/mens/player-rankings/odi/bowling")
page
```

Out[107]:
```
<Response [200]>
```

In [108]:

```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[108]:

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="Live Cricket Scores &amp; News International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos and fixtures from the International Cricket Council." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos and fixtures from the International Cricket Council." name="twitter:description"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="Live Cricket Scores &amp; News International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>Live Cricket Scores &amp; News International Cricket Council</title>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos and fixtures from the International Cricket Council." property="og:description"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
<script>
```

In [109]:
```
table=soup.find("table",class_="table rankings-table")
table
```

Out[109]:
```
<table class="table rankings-table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell">Player</th>
<th class="table-head__cell">Team</th>
<th class="table-head__cell table-head__cell--rating">Rating</th>
<th class="table-head__cell u-text-right u-hide-phablet">Career Best Rating</th>
</tr>
</thead>
<tr class="rankings-block__banner">
<td class="rankings-block__position">1</td>
<td class="rankings-block__top-player-container">
<div class="u-flex-center u-text-left">
<div class="rankings-block__player-image-container rankings-block__player-image-container--large">
<a href="/rankings/mens/player-rankings/969">
<img alt="Player Image" class="rankings-block__player-image image-missing" data-error="true" data-player="969" data-
size="130x150" data-widget="player-image" src="//resources.pulse.icc-cricket.com/players/130x150/Photo-Missing.png"/
>
```

In [110]:
```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n"," ").strip()
    headers.append(title)
```

In [111]:
```
df=pd.DataFrame(columns= headers)
df
```

Out[111]:

Pos	Player	Team	Rating	Career Best Rating

In [112]:

```python
for row in table.find_all('tr')[1:]:
    data=row.find_all("td")
    row_data=[td.text.replace("\n","").strip() for td in data]
    length=len(df)
    df.loc[length]=row_data
    #df.to_csv("top_bowlers_data")
```

In [113]:

```python
df=pd.read_csv("top_bowlers_data")
df
```

Out[113]:

	Unnamed: 0	Pos	Player	Team	Rating	Career Best Rating
0	0	1	Trent Boult	NZ	737	770 v West Indies, 22/06/2019
1	1	2 This player has moved up in the ranki...	Josh Hazlewood	AUS	709	733 v England, 26/01/2018
2	2	3 This player has moved down in the ran...	Mujeeb Ur Rahman	AFG	708	712 v Ireland, 24/01/2021
3	3	4 This player has moved down in the ran...	Chris Woakes	ENG	700	711 v Sri Lanka, 04/07/2021
4	4	5 This player has moved down in the ran...	Mehedi Hasan	BAN	692	725 v Sri Lanka, 25/05/2021
...	...	...	...	...	...	...
95	95	96 This player has moved up in the rank...	Sharafuddin Ashraf	AFG	347	354 v West Indies, 25/03/2018
96	96	This player has moved down in the rankings sin...	Roston Chase	NaN	347	NaN
97	97	98	Bilal Khan	OMA	344	344 v United States, 11/02/2020
98	98	99	Saurabh Netravalkar	USA	343	352 v Oman, 11/02/2020
99	99	100	Norman Vanua	PNG	342	353 v Namibia, 22/09/2019

100 rows × 6 columns

In [114]:
```
df=pd.DataFrame(data=df,columns=headers)
df.head(10)
```

Out[114]:

	Pos	Player	Team	Rating	Career Best Rating
0	1	Trent Boult	NZ	737	770 v West Indies, 22/06/2019
1	2 This player has moved up in the ranki...	Josh Hazlewood	AUS	709	733 v England, 26/01/2018
2	3 This player has moved down in the ran...	Mujeeb Ur Rahman	AFG	708	712 v Ireland, 24/01/2021
3	4 This player has moved down in the ran...	Chris Woakes	ENG	700	711 v Sri Lanka, 04/07/2021
4	5 This player has moved down in the ran...	Mehedi Hasan	BAN	692	725 v Sri Lanka, 25/05/2021
5	6 This player has moved down in the ran...	Matt Henry	NZ	691	691 v Bangladesh, 26/03/2021
6	7	Jasprit Bumrah	IND	679	841 v West Indies, 01/11/2018
7	8	Mitchell Starc	AUS	652	783 v New Zealand, 29/03/2015
8	9	Shakib Al Hasan	BAN	650	717 v Zimbabwe, 05/11/2009
9	10	Kagiso Rabada	SA	648	724 v England, 29/05/2017

6. Write a python program to scrape cricket rankings from ‘www.icc-cricket.com’ (http://www.icc-cricket.com%E2%80%99). You have toscrape:

i) Top 10 ODI teams in women’s cricket along with the records for matches, points and rating.

In [115]:
```
page=requests.get("https://www.icc-cricket.com/rankings/womens/team-rankings/odi")
page
```

Out[115]:
```
<Response [200]>
```

In [116]:

```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[116]:

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="ICC Ranking for ODI teams International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official International Cricket Council rankings for test match cricket teams. Discover latest ICC ran
kings table, predict upcoming matches, see points and ratings for all teams." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official International Cricket Council rankings for test match cricket teams. Discover latest ICC ran
kings table, predict upcoming matches, see points and ratings for all teams." name="twitter:description"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="ICC Ranking for ODI teams International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>ICC Ranking for ODI teams International Cricket Council</title>
<meta content="Official International Cricket Council rankings for test match cricket teams. Discover latest ICC ran
kings table, predict upcoming matches, see points and ratings for all teams." property="og:description"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
<script>
```

In [117]:
```
table=soup.find("table",class_="table")
table
```

Out[117]:
```
<table class="table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell rankings-table__team">
<span class="u-hide-mobile">Team</span>
<span class="u-show-mobile">T</span>
</th>
<th class="table-head__cell u-center-text">
<span class="u-hide-mobile">Matches</span>
<span class="u-show-mobile">M</span>
</th>
<th class="table-head__cell u-center-text">
<span class="u-hide-mobile">Points</span>
<span class="u-show-mobile">P</span>
</th>
<th class="table-head__cell u-text-right rating">
<span class="u-hide-mobile">Rating</span>
<span class="u-show-mobile">R</span>
```

In [118]:
```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n"," ").strip()
    headers.append(title)
```

In [119]:
```
df=pd.DataFrame(columns=headers)
df
```

Out[119]:

Pos	Team T	Matches M	Points P	Rating R

In [120]:

```
for row in table.find_all('tr')[1:]:
    data=row.find_all("td")
    row_data=[td.text.replace("\n","").strip() for td in data]
    length=len(df)
    df.loc[length]=row_data
    df.to_csv("top_teams_womens_cricket")
```

In [121]:

```
df=pd.read_csv("top_teams_womens_cricket")
df
```

Out[121]:

	Unnamed: 0	Pos	Team T	Matches M	Points P	Rating R
0	0	1	AustraliaAUS	18	2,955	164
1	1	2	EnglandENG	20	2,370	119
2	2	3	South AfricaSA	24	2,828	118
3	3	4	IndiaIND	23	2,535	110
4	4	5	New ZealandNZ	21	1,947	93
5	5	6	West IndiesWI	17	1,427	84
6	6	7	PakistanPAK	20	1,496	75
7	7	8	BangladeshBAN	5	306	61
8	8	9	Sri LankaSL	11	519	47
9	9	10	IrelandIRE	2	25	13

In [122]:
```
df=pd.DataFrame(data=df,columns=headers)
df
```

Out[122]:

	Pos	Team T	Matches M	Points P	Rating R
0	1	AustraliaAUS	18	2,955	164
1	2	EnglandENG	20	2,370	119
2	3	South AfricaSA	24	2,828	118
3	4	IndiaIND	23	2,535	110
4	5	New ZealandNZ	21	1,947	93
5	6	West IndiesWI	17	1,427	84
6	7	PakistanPAK	20	1,496	75
7	8	BangladeshBAN	5	306	61
8	9	Sri LankaSL	11	519	47
9	10	IrelandIRE	2	25	13

ii) Top 10 women’s ODI players along with the records of their team and rating.

In [123]:
```
page=requests.get("https://www.icc-cricket.com/rankings/womens/player-rankings/odi/batting")
page
```

Out[123]:
```
<Response [200]>
```

In [124]:

```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[124]:

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="Live Cricket Scores &amp; News International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="twitter:description"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="Live Cricket Scores &amp; News International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>Live Cricket Scores &amp; News International Cricket Council</title>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." property="og:description"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
<script>
```

In [125]:
```
table=soup.find("table",class_="table rankings-table")
table
```

Out[125]:
```
<table class="table rankings-table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell">Player</th>
<th class="table-head__cell">Team</th>
<th class="table-head__cell table-head__cell--rating">Rating</th>
<th class="table-head__cell u-text-right u-hide-phablet">Career Best Rating</th>
</tr>
</thead>
<tr class="rankings-block__banner">
<td class="rankings-block__position">1</td>
<td class="rankings-block__top-player-container">
<div class="u-flex-center u-text-left">
<div class="rankings-block__player-image-container rankings-block__player-image-container--large">
<a href="/rankings/womens/player-rankings/558">
<img alt="Player Image" class="rankings-block__player-image image-missing" data-error="true" data-player="558" data-
size="130x150" data-widget="player-image" src="//resources.pulse.icc-cricket.com/players/130x150/Photo-Missing.png"/
>
```

In [126]:
```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n"," ").strip()
    headers.append(title)
```

In [127]:
```
df=pd.DataFrame(columns=headers)
df
```

Out[127]:

Pos	Player	Team	Rating	Career Best Rating

```
In [128]: for row in table.find_all('tr')[1:]:
              data=row.find_all("td")
              row_data=[td.text.replace("\n","").strip() for td in data]
              length=len(df)
              df.loc[length]=row_data
              #df.to_csv("top_players_womens_teams")
```

```
In [129]: df=pd.read_csv("top_players_womens_teams")
          df
```

```
Out[129]:
          Unnamed: 0                                          Pos            Player  Team  Rating        Career Best Rating
      0            0                                            1       Mithali Raj   IND     762   839 v Australia, 24/12/2004
      1            1                                            2        Lizelle Lee    SA     758       773 v India, 14/03/2021
      2            2                                            3       Alyssa Healy   AUS     756  756 v New Zealand, 10/04/2021
      3            3                                            4    Tammy Beaumont   ENG     754       791 v India, 27/06/2021
      4            4                                            5    Stafanie Taylor    WI     736    766 v Pakistan, 07/07/2021
    ...          ...                                          ...               ...   ...     ...                           ...
     95           95     96 This player has moved up in the rank...    Chinelle Henry    WI     180   196 v Australia, 11/09/2019
     96           96    97 This player has moved down in the ra...          Jess Kerr    NZ     179   179 v Australia, 10/04/2021
     97           97  This player has moved down in the rankings sin...   Hannah Rowe    NZ     179   191 v Australia, 04/04/2021
     98           98                                          NaN     Sushma Verma   IND     179   261 v Australia, 18/03/2018
     99           99                                          100    Shabnim Ismail    SA     178     290 v Ireland, 14/01/2014

    100 rows × 6 columns
```

In [130]:

```
df=pd.DataFrame(data=df,columns=headers)
df.head(10)
```

Out[130]:

	Pos	Player	Team	Rating	Career Best Rating
0	1	Mithali Raj	IND	762	839 v Australia, 24/12/2004
1	2	Lizelle Lee	SA	758	773 v India, 14/03/2021
2	3	Alyssa Healy	AUS	756	756 v New Zealand, 10/04/2021
3	4	Tammy Beaumont	ENG	754	791 v India, 27/06/2021
4	5	Stafanie Taylor	WI	736	766 v Pakistan, 07/07/2021
5	6	Meg Lanning	AUS	723	834 v New Zealand, 24/02/2016
6	7	Amy Satterthwaite	NZ	715	756 v Australia, 02/03/2017
7	8	Natalie Sciver	ENG	706	712 v India, 25/02/2019
8	9	Smriti Mandhana	IND	701	797 v England, 28/02/2019
9	10	Laura Wolvaardt	SA	683	725 v India, 07/03/2021

iii) Top 10 women’s ODI all-rounder along with the records of their team and rating.

In [131]:

```
page=requests.get("https://www.icc-cricket.com/rankings/womens/player-rankings/odi/all-rounder")
page
```

Out[131]:

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

In [132]:

```
soup=BeautifulSoup(page.text,"lxml")
soup
```

Out[132]:

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta content="Live Cricket Scores &amp; News International Cricket Council" name="twitter:title"/>
<meta content="website" property="og:type"/>
<meta content="summary_large_image" property="twitter:card"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="description"/>
<meta content="@icc" property="twitter:site"/>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." name="twitter:description"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" name="twitter:image"/>
<meta content="Live Cricket Scores &amp; News International Cricket Council" property="og:title"/>
<meta content="https://www.icc-cricket.com/resources/ver/i/elements/default-thumbnail.jpg" property="og:image"/>
<title>Live Cricket Scores &amp; News International Cricket Council</title>
<meta content="Official ICC Cricket website - live matches, scores, news, highlights, commentary, rankings, videos a
nd fixtures from the International Cricket Council." property="og:description"/>
<!-- Designed and built by Pulselive - www.pulselive.com -->
<script>
```

In [133]:

```
table=soup.find("table",class_="table rankings-table")
table
```

Out[133]:

```
<table class="table rankings-table">
<thead>
<tr class="table-head">
<th class="table-head__cell u-text-right">Pos</th>
<th class="table-head__cell">Player</th>
<th class="table-head__cell">Team</th>
<th class="table-head__cell table-head__cell--rating">Rating</th>
<th class="table-head__cell u-text-right u-hide-phablet">Career Best Rating</th>
</tr>
</thead>
<tr class="rankings-block__banner">
<td class="rankings-block__position">1</td>
<td class="rankings-block__top-player-container">
<div class="u-flex-center u-text-left">
<div class="rankings-block__player-image-container rankings-block__player-image-container--large">
<a href="/rankings/womens/player-rankings/424">
<img alt="Player Image" class="rankings-block__player-image image-missing" data-error="true" data-player="424" data-
size="130x150" data-widget="player-image" src="//resources.pulse.icc-cricket.com/players/130x150/Photo-Missing.png"/
>
```

In [134]:

```
headers=[]
for i in table.find_all("th"):
    title=i.text.replace("\n"," ").strip()
    headers.append(title)
```

In [135]:

```
df=pd.DataFrame(columns=headers)
df
```

Out[135]:

Pos	Player	Team	Rating	Career Best Rating

In [136]:

```
for row in table.find_all('tr')[1:]:
    data=row.find_all("td")
    row_data=[td.text.replace("\n","").strip() for td in data]
    length=len(df)
    df.loc[length]=row_data
    #df.to_csv("top_odi_allrounders_womens")
```

In [137]:
```
df=pd.read_csv("top_odi_allrounders_womens")
df
```

Out[137]:

	Unnamed: 0	Pos	Player	Team	Rating	Career Best Rating
0	0	1.0	Marizanne Kapp	SA	418	418 v India, 17/03/2021
1	1	NaN	Ellyse Perry	AUS	418	548 v West Indies, 11/09/2019
2	2	3.0	Stafanie Taylor	WI	394	559 v New Zealand, 10/10/2013
3	3	4.0	Natalie Sciver	ENG	365	365 v India, 03/07/2021
4	4	5.0	Deepti Sharma	IND	331	397 v South Africa, 09/10/2019
5	5	6.0	Jess Jonassen	AUS	307	308 v West Indies, 11/09/2019
6	6	7.0	Ashleigh Gardner	AUS	252	256 v New Zealand, 04/04/2021
7	7	8.0	Dane van Niekerk	SA	243	421 v Sri Lanka, 11/02/2019
8	8	9.0	Sophie Devine	NZ	242	305 v Australia, 05/10/2020
9	9	10.0	Katherine Brunt	ENG	239	270 v Sri Lanka, 16/03/2019
10	10	11.0	Amelia Kerr	NZ	236	247 v Australia, 07/04/2021
11	11	12.0	Hayley Matthews	WI	215	227 v Pakistan, 12/07/2021
12	12	NaN	Jhulan Goswami	IND	215	308 v Australia, 02/02/2016
13	13	14.0	Shikha Pandey	IND	195	262 v South Africa, 14/10/2019
14	14	15.0	Heather Knight	ENG	181	249 v West Indies, 13/06/2019
15	15	16.0	Rumana Ahmed	BAN	179	180 v Pakistan, 04/11/2019
16	16	17.0	Nida Dar	PAK	176	178 v West Indies, 15/07/2021
17	17	18.0	Chamari Athapaththu	SL	164	187 v South Africa, 17/02/2019
18	18	19.0	Sune Luus	SA	153	217 v India, 05/02/2018
19	19	20.0	Harmanpreet Kaur	IND	149	174 v South Africa, 14/03/2021

In [138]:
```
df=pd.DataFrame(data=df,columns=headers)
df.head(10)
```

Out[138]:

	Pos	Player	Team	Rating	Career Best Rating
0	1.0	Marizanne Kapp	SA	418	418 v India, 17/03/2021
1	NaN	Ellyse Perry	AUS	418	548 v West Indies, 11/09/2019
2	3.0	Stafanie Taylor	WI	394	559 v New Zealand, 10/10/2013
3	4.0	Natalie Sciver	ENG	365	365 v India, 03/07/2021
4	5.0	Deepti Sharma	IND	331	397 v South Africa, 09/10/2019
5	6.0	Jess Jonassen	AUS	307	308 v West Indies, 11/09/2019
6	7.0	Ashleigh Gardner	AUS	252	256 v New Zealand, 04/04/2021
7	8.0	Dane van Niekerk	SA	243	421 v Sri Lanka, 11/02/2019
8	9.0	Sophie Devine	NZ	242	305 v Australia, 05/10/2020
9	10.0	Katherine Brunt	ENG	239	270 v Sri Lanka, 16/03/2019

7. Write a python program to scrape details of all the mobile phones under Rs. 20,000 listed on Amazon.in. The scraped data should include Product Name, Price, Image URL and Average Rating.

In [20]:
```
page=requests.get("https://www.amazon.in/Mobile-Phone-Under-20000-Rupees/s?k=Mobile+Phone+Under+20000+Rupees")
page
```

Out[20]:
```
<Response [200]>
```

In [21]:
```
soup=BeautifulSoup(page.content,"lxml")
soup
```

Out[21]:
```
<!DOCTYPE html>
<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-in"><!-- sp:feature:head-start -->
<head><script>var aPageStart = (new Date()).getTime();</script><meta charset="utf-8"/>
<!-- sp:feature:cs-optimization -->
<meta content="on" http-equiv="x-dns-prefetch-control"/>
<link href="https://images-eu.ssl-images-amazon.com" rel="dns-prefetch"/>
<link href="https://m.media-amazon.com" rel="dns-prefetch"/>
<link href="https://completion.amazon.com" rel="dns-prefetch"/>
<!-- sp:feature:aui-assets -->
<link href="https://images-eu.ssl-images-amazon.com/images/I/11EIQ5IGqaL._RC|01ZTHTZObnL.css,41SIz69qHYL.css,31qGOnS
AToL.css,013z33uKh2L.css,017DsKjNQJL.css,0131vqwP5UL.css,41EWOOlBJ9L.css,11TIuySqr6L.css,01ElnPiDxWL.css,11bGSgD5pD
L.css,01Dm5eKVxwL.css,01IdKcBuAdL.css,01y-XAlI+2L.css,21N4kUH7pxL.css,01oDR3IULNL.css,41-PwE7+H0L.css,21j0IlW7xKL.cs
s,01XPHJk60-L.css,014OeDQisGL.css,21aPhFy+riL.css,11gneA3MtJL.css,21fecG8pUzL.css,01RddH8vm-L.css,01CFUgsA-YL.css,31
C80IiXalL.css,11qour3ND0L.css,11tRp6+0HHL.css,11061HxnEvL.css,11oHt2HYxnL.css,013RDhw9hoL.css,11JQtnL-6eL.css,11RKoG
Sb-gL.css,11jtXRmppwL.css,01QrWuRrZ-L.css,21pIv-yKhaL.css,11QyqG8yiqL.css,11kwKGWmBfL.css,11F2+OBzLyL.css,11Y05DTEL6
L.css,01cbS3UK11L.css,21F85am0yFL.css,01giMEP+djL.css_.css?AUIClients/AmazonUI#page_type-Search.not-trident" rel="st
ylesheet"/>
<script>
(function(d,f,Q,D){function E(a){u&&u.tag&&u.tag(p(":","aui",a))}function r(a,b){u&&u.count&&u.count("aui:"+a,0===b?
```

In [22]:

```
names=soup.find_all("h2",class_="a-size-mini a-spacing-none a-color-base s-line-clamp-2")
names
```

Out[22]:

```
[<h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Re
dmi-9A-2GB-32GB-Storage/dp/B08696XB4B?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-text-norma
l">Redmi 9A (Nature Green, 2GB RAM, 32GB Storage) | 2GHz Octa-core Helio G25 Processor | 5000 mAh Battery</span> </a
> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/re
alme-Storage-Additional-Exchange-Offers/dp/B096RWS54N?dchild=1" target="_blank"><span class="a-size-medium a-color-b
ase a-text-normal">realme C11 (2021) (Cool Blue, 2GB RAM, 32GB Storage) with No Cost EMI/Additional Exchange Offers
</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Op
po-Fantasy-Storage-Additional-Exchange/dp/B08444SXZ6?dchild=1" target="_blank"><span class="a-size-medium a-color-ba
se a-text-normal">Oppo A31 (Fantasy White, 6GB RAM, 128GB Storage) with No Cost EMI/Additional Exchange Offers</span
> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Sa
msung-Galaxy-Ocean-128GB-Storage/dp/B07HGGYWL6?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-t
ext-normal">Samsung Galaxy M31 (Ocean Blue, 6GB RAM, 128GB Storage)</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Te
st-Exclusive_2020_1112-Multi-3GB-Storage/dp/B089MS7D8F?dchild=1" target="_blank"><span class="a-size-medium a-color-
base a-text-normal">Redmi 9 Power (Mighty Black 4GB RAM 64GB Storage) - 6000mAh Battery |FHD+ Screen | 48MP Quad Cam
era | Alexa Hands-Free Capable</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Te
st-Exclusive_2020_1138-Multi-3GB-Storage/dp/B089MTR9JW?dchild=1" target="_blank"><span class="a-size-medium a-color-
base a-text-normal">Redmi 9 Power (Mighty Black, 6GB RAM, 128GB Storage) - 6000mAh Battery |FHD+ Screen| 48MP Quad C
amera | Snapdragon 662 Processor | Alexa Hands-Free Capable</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Sa
msung-Galaxy-Space-Black-Storage/dp/B07HGN617M?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-t
ext-normal">Samsung Galaxy M31 (Space Black, 6GB RAM, 128GB Storage)</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/re
alme-Storage-Additional-Exchange-Offers/dp/B096RT57F9?dchild=1" target="_blank"><span class="a-size-medium a-color-b
ase a-text-normal">realme C11 (2021) (Cool Grey, 2GB RAM, 32GB Storage) with No Cost EMI/Additional Exchange Offers
</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Sa
msung-Galaxy-Storage-Triple-Camera/dp/B08RSXPPFY?dchild=1" target="_blank"><span class="a-size-medium a-color-base a
-text-normal">Samsung Galaxy M02s (Black,4GB RAM, 64GB Storage) | 5000 mAh | Triple Camera</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Te
st-Exclusive_2020_1140-Multi-3GB-Storage/dp/B089MT36T8?dchild=1" target="_blank"><span class="a-size-medium a-color-
base a-text-normal">Redmi Note 10 (Aqua Green, 4GB RAM, 64GB Storage) -Amoled Dot Display | 48MP Sony Sensor IMX582
| Snapdragon 678 Processor</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Re
dmi-9A-Midnight-2GB-32GB/dp/B08697N43G?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-text-norm
al">Redmi 9A (Midnight Black, 2GB RAM, 32GB Storage) | 2GHz Octa-Core Helio G25 Processor</span> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Re
dmi-Carbon-Black-64GB-Storage/dp/B086985T6R?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-text
-normal">Redmi 9 (Carbon Black, 4GB RAM, 64GB Storage) | 2.3GHz Mediatek Helio G35 Octa core Processor</span> </a>
</h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Re
dmi-Note-Pebble-Grey-Storage/dp/B086977TR6?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-text-
normal">Redmi Note 9 (Pebble Grey, 4GB RAM 64GB Storage) - 48MP Quad Camera &amp; Full HD+ Display</span> </a> </h2
>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Fa
ntasy-Storage-Additional-Exchange-Offers/dp/B08444XVNP?dchild=1" target="_blank"><span class="a-size-medium a-color-
base a-text-normal">OPPO A31 (Fantasy White, 4GB RAM, 64GB Storage) with No Cost EMI/Additional Exchange Offers</spa
n> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Sa
msung-Galaxy-Ocean-128GB-Storage/dp/B07HG8S7KP?dchild=1" target="_blank"><span class="a-size-medium a-color-base a-t
ext-normal">Samsung Galaxy M31 (Ocean Blue, 8GB RAM, 128GB Storage) 6 Months Free Screen Replacement for Prime</span
> </a> </h2>,
 <h2 class="a-size-mini a-spacing-none a-color-base s-line-clamp-2"><a class="a-link-normal a-text-normal" href="/Op
po-Mystery-Storage-Additional-Exchange/dp/B08444S68L?dchild=1" target="_blank"><span class="a-size-medium a-color-ba
se a-text-normal">Oppo A31 (Mystery Black, 6GB RAM, 128GB Storage) with No Cost EMI/Additional Exchange Offers</span
> </a> </h2>]
```

In [23]:
```
mobile_name=[]
for i in range(0,len(names)):
    mobile_name.append(names[i].get_text())
mobile_name
```

Out[23]:
```
['Redmi 9A (Nature Green, 2GB RAM, 32GB Storage) | 2GHz Octa-core Helio G25 Processor | 5000 mAh Battery  ',
 'realme C11 (2021) (Cool Blue, 2GB RAM, 32GB Storage) with No Cost EMI/Additional Exchange Offers  ',
 'Oppo A31 (Fantasy White, 6GB RAM, 128GB Storage) with No Cost EMI/Additional Exchange Offers  ',
 'Samsung Galaxy M31 (Ocean Blue, 6GB RAM, 128GB Storage)  ',
 'Redmi 9 Power (Mighty Black 4GB RAM 64GB Storage) - 6000mAh Battery |FHD+ Screen | 48MP Quad Camera | Alexa Hands-Fre
e Capable  ',
 'Redmi 9 Power (Mighty Black, 6GB RAM, 128GB Storage) - 6000mAh Battery |FHD+ Screen| 48MP Quad Camera | Snapdragon 66
2 Processor | Alexa Hands-Free Capable  ',
 'Samsung Galaxy M31 (Space Black, 6GB RAM, 128GB Storage)  ',
 'realme C11 (2021) (Cool Grey, 2GB RAM, 32GB Storage) with No Cost EMI/Additional Exchange Offers  ',
 'Samsung Galaxy M02s (Black,4GB RAM, 64GB Storage) | 5000 mAh | Triple Camera  ',
 'Redmi Note 10 (Aqua Green, 4GB RAM, 64GB Storage) -Amoled Dot Display | 48MP Sony Sensor IMX582 | Snapdragon 678 Proc
essor  ',
 'Redmi 9A (Midnight Black, 2GB RAM, 32GB Storage) | 2GHz Octa-Core Helio G25 Processor  ',
 'Redmi 9 (Carbon Black, 4GB RAM, 64GB Storage) | 2.3GHz Mediatek Helio G35 Octa core Processor  ',
 'Redmi Note 9 (Pebble Grey, 4GB RAM 64GB Storage) - 48MP Quad Camera & Full HD+ Display  ',
 'OPPO A31 (Fantasy White, 4GB RAM, 64GB Storage) with No Cost EMI/Additional Exchange Offers  ',
 'Samsung Galaxy M31 (Ocean Blue, 8GB RAM, 128GB Storage) 6 Months Free Screen Replacement for Prime  ',
 'Oppo A31 (Mystery Black, 6GB RAM, 128GB Storage) with No Cost EMI/Additional Exchange Offers  ']
```

In [24]:
```
len("mobile_name")
```

Out[24]:
```
11
```

In [ ]:

In [25]:

```
prices=soup.find_all("span",class_="a-offscreen")
prices
```

Out[25]:

```
[<span class="a-offscreen">₹6,999</span>,
 <span class="a-offscreen">₹8,499</span>,
 <span class="a-offscreen">₹6,999</span>,
 <span class="a-offscreen">₹7,999</span>,
 <span class="a-offscreen">₹12,490</span>,
 <span class="a-offscreen">₹15,990</span>,
 <span class="a-offscreen">₹14,999</span>,
 <span class="a-offscreen">₹19,999</span>,
 <span class="a-offscreen">₹10,999</span>,
 <span class="a-offscreen">₹13,999</span>,
 <span class="a-offscreen">₹13,499</span>,
 <span class="a-offscreen">₹16,999</span>,
 <span class="a-offscreen">₹14,999</span>,
 <span class="a-offscreen">₹19,999</span>,
 <span class="a-offscreen">₹6,999</span>,
 <span class="a-offscreen">₹7,999</span>,
 <span class="a-offscreen">₹10,499</span>,
 <span class="a-offscreen">₹11,499</span>,
 <span class="a-offscreen">₹13,499</span>,
 <span class="a-offscreen">₹15,999</span>,
 <span class="a-offscreen">₹6,999</span>,
 <span class="a-offscreen">₹8,499</span>,
 <span class="a-offscreen">₹11,999</span>,
 <span class="a-offscreen">₹14,999</span>,
 <span class="a-offscreen">₹10,990</span>,
 <span class="a-offscreen">₹12,990</span>,
 <span class="a-offscreen">₹16,999</span>,
 <span class="a-offscreen">₹21,999</span>,
 <span class="a-offscreen">₹12,490</span>,
 <span class="a-offscreen">₹15,990</span>]
```

In [ ]:

```
```

In [26]:

```
mobile_price=[]
for i in range(0,len(prices)):
    mobile_price.append(prices[i].get_text())
mobile_price
```

Out[26]:

```
['₹6,999',
 '₹8,499',
 '₹6,999',
 '₹7,999',
 '₹12,490',
 '₹15,990',
 '₹14,999',
 '₹19,999',
 '₹10,999',
 '₹13,999',
 '₹13,499',
 '₹16,999',
 '₹14,999',
 '₹19,999',
 '₹6,999',
 '₹7,999',
 '₹10,499',
 '₹11,499',
 '₹13,499',
 '₹15,999',
 '₹6,999',
 '₹8,499',
 '₹11,999',
 '₹14,999',
 '₹10,990',
 '₹12,990',
 '₹16,999',
 '₹21,999',
 '₹12,490',
 '₹15,990']
```

In [27]:
```
len(mobile_price)
```

Out[27]:
```
30
```

In [28]:
```
my_list=(mobile_price)
my_list=list(dict.fromkeys(my_list))
print(my_list)
```

```
['₹6,999', '₹8,499', '₹7,999', '₹12,490', '₹15,990', '₹14,999', '₹19,999', '₹10,999', '₹13,999', '₹13,499', '₹16,999',
'₹10,499', '₹11,499', '₹15,999', '₹11,999', '₹10,990', '₹12,990', '₹21,999']
```

In [ ]:
```

```

In [29]:

```
mobile_price=pd.DataFrame(my_list)
mobile_price
```

Out[29]:

```
          0
 0   ₹6,999
 1   ₹8,499
 2   ₹7,999
 3  ₹12,490
 4  ₹15,990
 5  ₹14,999
 6  ₹19,999
 7  ₹10,999
 8  ₹13,999
 9  ₹13,499
10  ₹16,999
11  ₹10,499
12  ₹11,499
13  ₹15,999
14  ₹11,999
15  ₹10,990
16  ₹12,990
17  ₹21,999
```

In [ ]:

```

```

In [30]:

```
mobile_url=[]
x=soup.find_all("img",class_="s-image")
for img in x:
    mobile_url.append(img["src"])
mobile_url
```

Out[30]:

```
['https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71FYSKYFupS._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71hEzQGO5qL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/61LHaUOhehL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71OxJeyywSL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/618UBhFmaQS._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71IkA3T7hIL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/810GQ7CWdDL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/716nHhG9SWL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71X5I1cVfbL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY218_.jpg',
 'https://m.media-amazon.com/images/I/71KCwNV6MuL._AC_UY218_.jpg']
```

In [31]:

```
len(mobile_url)
```

Out[31]:

```
16
```

In [ ]:

```

```

In [32]:

```
images=soup.find_all("img",class_="s-image")

for image in images:
    name=image["alt"]
    link=image["src"]
    print(link)
```

```
https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71FYSKYFupS._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71FYSKYFupS._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY218_.jpg (https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71hEzQGO5qL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71hEzQGO5qL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/61LHaUOhehL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/61LHaUOhehL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71OxJeyywSL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71OxJeyywSL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/618UBhFmaQS._AC_UY218_.jpg (https://m.media-amazon.com/images/I/618UBhFmaQS._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71IkA3T7hIL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71IkA3T7hIL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/810GQ7CWdDL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/810GQ7CWdDL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71sxlhYhKWL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/716nHhG9SWL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/716nHhG9SWL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71X5I1cVfbL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71X5I1cVfbL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY218_.jpg (https://m.media-amazon.com/images/I/61CnyJ-IbML._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71-Su4Wr0HL._AC_UY2
18_.jpg)
https://m.media-amazon.com/images/I/71KCwNV6MuL._AC_UY218_.jpg (https://m.media-amazon.com/images/I/71KCwNV6MuL._AC_UY2
18_.jpg)
```

In [ ]:

```
```

In [33]:

```
rating=soup.find_all("span",class_="a-icon-alt")
rating
```

Out[33]:

```
[<span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.1 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.3 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.3 out of 5 stars</span>,
 <span class="a-icon-alt">4.1 out of 5 stars</span>,
 <span class="a-icon-alt">4.1 out of 5 stars</span>,
 <span class="a-icon-alt">4.1 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.3 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4.3 out of 5 stars</span>,
 <span class="a-icon-alt">4.2 out of 5 stars</span>,
 <span class="a-icon-alt">4 Stars &amp; Up</span>,
 <span class="a-icon-alt">3 Stars &amp; Up</span>,
 <span class="a-icon-alt">2 Stars &amp; Up</span>,
 <span class="a-icon-alt">1 Star &amp; Up</span>]
```

In [34]:

```
mobile_rate=[]
for i in range(0,len(rating)):
    mobile_rate.append(rating[i].get_text())
mobile_rate
```

Out[34]:

```
['4.2 out of 5 stars',
 '4.1 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4 Stars & Up',
 '3 Stars & Up',
 '2 Stars & Up',
 '1 Star & Up']
```

In [35]:

```
mobile_rate.reverse()
```

In [36]:
```
mobile_rate
```

Out[36]:
```
['1 Star & Up',
 '2 Stars & Up',
 '3 Stars & Up',
 '4 Stars & Up',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.2 out of 5 stars',
 '4.1 out of 5 stars',
 '4.2 out of 5 stars']
```

In [37]:
```
mobile_rate.pop(0)
```

Out[37]:
```
'1 Star & Up'
```

In [38]:
```
mobile_rate.pop(0)
```

Out[38]:
```
'2 Stars & Up'
```

In [39]:
```
mobile_rate.pop(0)
```

Out[39]:
```
'3 Stars & Up'
```

In [40]:
```
mobile_rate.pop(0)
```

Out[40]:
```
'4 Stars & Up'
```

In [41]:
```
len(mobile_rate)
```

Out[41]:
```
16
```

In [42]:
```
mobile_rate.sort()
```

In [43]:
```
mobile_rate
```

Out[43]:
```
['4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.1 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.2 out of 5 stars',
 '4.3 out of 5 stars',
 '4.3 out of 5 stars',
 '4.3 out of 5 stars',
 '4.3 out of 5 stars']
```

In [ ]:

In [44]:
```
import pandas as pd
import csv
```

In [47]:
```
details_mobiles_phones=pd.DataFrame({})
details_mobiles_phones['mobile_name']=mobile_name
details_mobiles_phones['mobile_price']=mobile_price
details_mobiles_phones["mobile_url"]=mobile_url
details_mobiles_phones['mobile_rate']=mobile_rate
details_mobiles_phones
```

Out[47]:

	mobile_name	mobile_price	mobile_url	mobile_rate
0	Redmi 9A (Nature Green, 2GB RAM, 32GB Storage)...	₹6,999	https://m.media-amazon.com/images/I/71sxlhYhKW...	4.1 out of 5 stars
1	realme C11 (2021) (Cool Blue, 2GB RAM, 32GB St...	₹8,499	https://m.media-amazon.com/images/I/71FYSKYFup...	4.1 out of 5 stars
2	Oppo A31 (Fantasy White, 6GB RAM, 128GB Storag...	₹7,999	https://m.media-amazon.com/images/I/61CnyJ-IbM...	4.1 out of 5 stars
3	Samsung Galaxy M31 (Ocean Blue, 6GB RAM, 128GB...	₹12,490	https://m.media-amazon.com/images/I/71-Su4Wr0H...	4.1 out of 5 stars
4	Redmi 9 Power (Mighty Black 4GB RAM 64GB Stora...	₹15,990	https://m.media-amazon.com/images/I/71hEzQGO5q...	4.2 out of 5 stars
5	Redmi 9 Power (Mighty Black, 6GB RAM, 128GB St...	₹14,999	https://m.media-amazon.com/images/I/61LHaUOheh...	4.2 out of 5 stars
6	Samsung Galaxy M31 (Space Black, 6GB RAM, 128G...	₹19,999	https://m.media-amazon.com/images/I/71OxJeyywS...	4.2 out of 5 stars
7	realme C11 (2021) (Cool Grey, 2GB RAM, 32GB St...	₹10,999	https://m.media-amazon.com/images/I/618UBhFmaQ...	4.2 out of 5 stars
8	Samsung Galaxy M02s (Black,4GB RAM, 64GB Stora...	₹13,999	https://m.media-amazon.com/images/I/71IkA3T7hI...	4.2 out of 5 stars
9	Redmi Note 10 (Aqua Green, 4GB RAM, 64GB Stora...	₹13,499	https://m.media-amazon.com/images/I/810GQ7CWdD...	4.2 out of 5 stars
10	Redmi 9A (Midnight Black, 2GB RAM, 32GB Storag...	₹16,999	https://m.media-amazon.com/images/I/71sxlhYhKW...	4.2 out of 5 stars
11	Redmi 9 (Carbon Black, 4GB RAM, 64GB Storage) ...	₹10,499	https://m.media-amazon.com/images/I/716nHhG9SW...	4.2 out of 5 stars
12	Redmi Note 9 (Pebble Grey, 4GB RAM 64GB Storag...	₹11,499	https://m.media-amazon.com/images/I/71X5I1cVfb...	4.3 out of 5 stars
13	OPPO A31 (Fantasy White, 4GB RAM, 64GB Storage...	₹15,999	https://m.media-amazon.com/images/I/61CnyJ-IbM...	4.3 out of 5 stars
14	Samsung Galaxy M31 (Ocean Blue, 8GB RAM, 128GB...	₹11,999	https://m.media-amazon.com/images/I/71-Su4Wr0H...	4.3 out of 5 stars
15	Oppo A31 (Mystery Black, 6GB RAM, 128GB Storag...	₹10,990	https://m.media-amazon.com/images/I/71KCwNV6Mu...	4.3 out of 5 stars

In [49]:
```
df=pd.DataFrame(details_mobiles_phones)
df.to_csv("details_mobiles_phones")
```

In [50]:
```
df=pd.read_csv("details_mobiles_phones")
df
```

Out[50]:

	Unnamed: 0	mobile_name	mobile_price	mobile_url	mobile_rate
0	0	Redmi 9A (Nature Green, 2GB RAM, 32GB Storage)...	₹6,999	https://m.media-amazon.com/images/I/71sxlhYhKW...	4.1 out of 5 stars
1	1	realme C11 (2021) (Cool Blue, 2GB RAM, 32GB St...	₹8,499	https://m.media-amazon.com/images/I/71FYSKYFup...	4.1 out of 5 stars
2	2	Oppo A31 (Fantasy White, 6GB RAM, 128GB Storag...	₹7,999	https://m.media-amazon.com/images/I/61CnyJ-IbM...	4.1 out of 5 stars
3	3	Samsung Galaxy M31 (Ocean Blue, 6GB RAM, 128GB...	₹12,490	https://m.media-amazon.com/images/I/71-Su4Wr0H...	4.1 out of 5 stars
4	4	Redmi 9 Power (Mighty Black 4GB RAM 64GB Stora...	₹15,990	https://m.media-amazon.com/images/I/71hEzQGO5q...	4.2 out of 5 stars
5	5	Redmi 9 Power (Mighty Black, 6GB RAM, 128GB St...	₹14,999	https://m.media-amazon.com/images/I/61LHaUOheh...	4.2 out of 5 stars
6	6	Samsung Galaxy M31 (Space Black, 6GB RAM, 128G...	₹19,999	https://m.media-amazon.com/images/I/71OxJeyywS...	4.2 out of 5 stars
7	7	realme C11 (2021) (Cool Grey, 2GB RAM, 32GB St...	₹10,999	https://m.media-amazon.com/images/I/618UBhFmaQ...	4.2 out of 5 stars
8	8	Samsung Galaxy M02s (Black,4GB RAM, 64GB Stora...	₹13,999	https://m.media-amazon.com/images/I/71IkA3T7hI...	4.2 out of 5 stars
9	9	Redmi Note 10 (Aqua Green, 4GB RAM, 64GB Stora...	₹13,499	https://m.media-amazon.com/images/I/810GQ7CWdD...	4.2 out of 5 stars
10	10	Redmi 9A (Midnight Black, 2GB RAM, 32GB Storag...	₹16,999	https://m.media-amazon.com/images/I/71sxlhYhKW...	4.2 out of 5 stars
11	11	Redmi 9 (Carbon Black, 4GB RAM, 64GB Storage) ...	₹10,499	https://m.media-amazon.com/images/I/716nHhG9SW...	4.2 out of 5 stars
12	12	Redmi Note 9 (Pebble Grey, 4GB RAM 64GB Storag...	₹11,499	https://m.media-amazon.com/images/I/71X5I1cVfb...	4.3 out of 5 stars
13	13	OPPO A31 (Fantasy White, 4GB RAM, 64GB Storage...	₹15,999	https://m.media-amazon.com/images/I/61CnyJ-IbM...	4.3 out of 5 stars
14	14	Samsung Galaxy M31 (Ocean Blue, 8GB RAM, 128GB...	₹11,999	https://m.media-amazon.com/images/I/71-Su4Wr0H...	4.3 out of 5 stars
15	15	Oppo A31 (Mystery Black, 6GB RAM, 128GB Storag...	₹10,990	https://m.media-amazon.com/images/I/71KCwNV6Mu...	4.3 out of 5 stars

8. Write a python program to extract information about the local weather from the National Weather Service website of USA, https://www.weather.gov/ (https://www.weather.gov/) for the city, San Francisco. You need to extract data about 7 day extended forecast display for the city. The data should include period, short description, temperature and description.

In [120]:
```
page=requests.get("https://forecast.weather.gov/MapClick.php?lat=37.777120000000025&lon=-122.41963999999996#.YSeKfEvitPY
page
```

Out[120]:
```
<Response [200]>
```

In [122]:
```
soup=BeautifulSoup(page.content,"html.parser")
soup
```

Out[122]:
```
<!DOCTYPE html>

<html class="no-js">
<head>
<!-- Meta -->
<meta content="width=device-width" name="viewport"/>
<link href="http://purl.org/dc/elements/1.1/" rel="schema.DC"/><title>National Weather Service</title><meta content
="National Weather Service" name="DC.title"><meta content="NOAA National Weather Service National Weather Service" n
ame="DC.description"/><meta content="US Department of Commerce, NOAA, National Weather Service" name="DC.creator"/><
meta content="" name="DC.date.created" scheme="ISO8601"/><meta content="EN-US" name="DC.language" scheme="DCTERMS.RF
C1766"/><meta content="weather, National Weather Service" name="DC.keywords"/><meta content="NOAA's National Weather
Service" name="DC.publisher"/><meta content="National Weather Service" name="DC.contributor"/><meta content="http://
www.weather.gov/disclaimer.php" name="DC.rights"/><meta content="General" name="rating"/><meta content="index,follo
w" name="robots"/>
<!-- Icons -->
<link href="./images/favicon.ico" rel="shortcut icon" type="image/x-icon"/>
<!-- CSS -->
<link href="css/bootstrap-3.2.0.min.css" rel="stylesheet"/>
<link href="css/bootstrap-theme-3.2.0.min.css" rel="stylesheet"/>
```

In [123]:

```
period=soup.find_all("p",class_="period-name")
period
```

Out[123]:

```
[<p class="period-name">Today<br/><br/></p>,
 <p class="period-name">Tonight<br/><br/></p>,
 <p class="period-name">Friday<br/><br/></p>,
 <p class="period-name">Friday<br/>Night</p>,
 <p class="period-name">Saturday<br/><br/></p>,
 <p class="period-name">Saturday<br/>Night</p>,
 <p class="period-name">Sunday<br/><br/></p>,
 <p class="period-name">Sunday<br/>Night</p>,
 <p class="period-name">Monday<br/><br/></p>]
```

In [124]:

```
period_name=[]
for i in range(0,len(period)):
    period_name.append(period[i].get_text().replace("\n",""))
period_name
```

Out[124]:

```
['Today',
 'Tonight',
 'Friday',
 'FridayNight',
 'Saturday',
 'SaturdayNight',
 'Sunday',
 'SundayNight',
 'Monday']
```

In [125]:
```
description=soup.find_all("p",class_="short-desc")
description
```

Out[125]:
```
[<p class="short-desc">Becoming<br/>Sunny</p>,
 <p class="short-desc">Increasing<br/>Clouds</p>,
 <p class="short-desc">Sunny then<br/>Haze</p>,
 <p class="short-desc">Haze then<br/>Mostly Clear</p>,
 <p class="short-desc">Sunny</p>,
 <p class="short-desc">Mostly Clear<br/>and Breezy<br/>then Mostly<br/>Clear</p>,
 <p class="short-desc">Sunny</p>,
 <p class="short-desc">Partly Cloudy</p>,
 <p class="short-desc">Mostly Sunny</p>]
```

In [126]:
```
short_description=[]
for i in range(0,len(description)):
    short_description.append(description[i].get_text().replace("\n",""))
short_description
```

Out[126]:
```
['BecomingSunny',
 'IncreasingClouds',
 'Sunny thenHaze',
 'Haze thenMostly Clear',
 'Sunny',
 'Mostly Clearand Breezythen MostlyClear',
 'Sunny',
 'Partly Cloudy',
 'Mostly Sunny']
```

In [159]:
```
temp_high=soup.find_all("p",class_="temp temp-high")
temp_high
```

Out[159]:
```
[<p class="temp temp-high">High: 72 °F</p>,
 <p class="temp temp-high">High: 74 °F</p>,
 <p class="temp temp-high">High: 75 °F</p>,
 <p class="temp temp-high">High: 73 °F</p>,
 <p class="temp temp-high">High: 67 °F</p>]
```

In [160]:

```
temperature_high=[]
for i in range(0,len(temp_high)):
    temperature_high.append(temp_high[i].get_text().replace("\n",""))
temperature_high
```

Out[160]:

```
['High: 72 °F', 'High: 74 °F', 'High: 75 °F', 'High: 73 °F', 'High: 67 °F']
```

In [161]:

```
temp_low=soup.find_all("p",class_="temp temp-low")
temp_low
```

Out[161]:

```
[<p class="temp temp-low">Low: 57 °F</p>,
 <p class="temp temp-low">Low: 57 °F</p>,
 <p class="temp temp-low">Low: 58 °F</p>,
 <p class="temp temp-low">Low: 56 °F</p>]
```

In [162]:

```
temperature_low=[]
for i in range(0,len(temp_low)):
    temperature_low.append(temp_low[i].get_text().replace("\n",""))
temperature_low
```

Out[162]:

```
['Low: 57 °F', 'Low: 57 °F', 'Low: 58 °F', 'Low: 56 °F']
```

In [163]:

```
temperature_high.insert(1,'Low: 57 °F')
```

In [164]:

```
temperature_high.insert(3,'Low: 57 °F')
```

In [165]:

```
temperature_high.insert(5,'Low: 58 °F')
```

In [166]:

```
temperature_high.insert(7,'Low: 56 °F')
```

In [205]:
```
temp_data=temperature_high
temp_data
```

Out[205]:
```
['High: 72 °F',
 'Low: 57 °F',
 'High: 74 °F',
 'Low: 57 °F',
 'High: 75 °F',
 'Low: 58 °F',
 'High: 73 °F',
 'Low: 56 °F',
 'High: 67 °F']
```

In [187]:
```
desc_odd=soup.find_all("div",class_="row row-odd row-forecast")
desc_odd
```

Out[187]:
```
[<div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Today</b></div><div class="col-sm-10 forecast-text">Partly sunny, then gradually becoming sunny, with a high near 72. Light and variable wind becoming west 15 to 20 mph in the afternoon. Winds could gust as high as 26 mph. </div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Friday</b></div><div class="col-sm-10 forecast-text">Widespread haze after noon. Sunny, with a high near 74. Light west wind increasing to 15 to 20 mph in the afternoon. Winds could gust as high as 25 mph. </div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Saturday</b></div><div class="col-sm-10 forecast-text">Sunny, with a high near 75. Light west southwest wind increasing to 13 to 18 mph in the afternoon. Winds could gust as high as 24 mph. </div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Sunday</b></div><div class="col-sm-10 forecast-text">Sunny, with a high near 73.</div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Monday</b></div><div class="col-sm-10 forecast-text">Mostly sunny, with a high near 67.</div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Tuesday</b></div><div class="col-sm-10 forecast-text">Sunny, with a high near 66.</div></div>,
 <div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>Wednesday</b></div><div class="col-sm-10 forecast-text">Sunny, with a high near 65.</div></div>]
```

In [189]:

```
description_odd=[]
for i in range(0,len(desc_odd)):
    description_odd.append(desc_odd[i].get_text().replace("\n"," "))
description_odd
```

Out[189]:

```
['TodayPartly sunny, then gradually becoming sunny, with a high near 72. Light and variable wind becoming west 15 to 20
mph in the afternoon. Winds could gust as high as 26 mph. ',
 'FridayWidespread haze after noon. Sunny, with a high near 74. Light west wind increasing to 15 to 20 mph in the after
noon. Winds could gust as high as 25 mph. ',
 'SaturdaySunny, with a high near 75. Light west southwest wind increasing to 13 to 18 mph in the afternoon. Winds coul
d gust as high as 24 mph. ',
 'SundaySunny, with a high near 73.',
 'MondayMostly sunny, with a high near 67.',
 'TuesdaySunny, with a high near 66.',
 'WednesdaySunny, with a high near 65.']
```

In [190]:

```
desc_even=soup.find_all("div",class_="row row-even row-forecast")
desc_even
```

Out[190]:

```
[<div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Tonight</b></div><div class="col-sm-10
forecast-text">Increasing clouds, with a low around 57. West wind 13 to 18 mph decreasing to 6 to 11 mph after midnigh
t. Winds could gust as high as 23 mph. </div></div>,
 <div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Friday Night</b></div><div class="col-s
m-10 forecast-text">Widespread haze before 9pm. Mostly clear, with a low around 57. Southwest wind 16 to 21 mph decreas
ing to 6 to 11 mph after midnight. Winds could gust as high as 26 mph. </div></div>,
 <div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Saturday Night</b></div><div class="col
-sm-10 forecast-text">Mostly clear, with a low around 58. Breezy. </div></div>,
 <div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Sunday Night</b></div><div class="col-s
m-10 forecast-text">Partly cloudy, with a low around 56.</div></div>,
 <div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Monday Night</b></div><div class="col-s
m-10 forecast-text">Mostly clear, with a low around 54.</div></div>,
 <div class="row row-even row-forecast"><div class="col-sm-2 forecast-label"><b>Tuesday Night</b></div><div class="col-
sm-10 forecast-text">Mostly clear, with a low around 54.</div></div>]
```

In [191]:

```
description_even=[]
for i in range(0,len(desc_even)):
    description_even.append(desc_even[i].get_text().replace("\n"," "))
description_even
```

Out[191]:

```
['TonightIncreasing clouds, with a low around 57. West wind 13 to 18 mph decreasing to 6 to 11 mph after midnight. Wind
s could gust as high as 23 mph. ',
 'Friday NightWidespread haze before 9pm. Mostly clear, with a low around 57. Southwest wind 16 to 21 mph decreasing to
6 to 11 mph after midnight. Winds could gust as high as 26 mph. ',
 'Saturday NightMostly clear, with a low around 58. Breezy. ',
 'Sunday NightPartly cloudy, with a low around 56.',
 'Monday NightMostly clear, with a low around 54.',
 'Tuesday NightMostly clear, with a low around 54.']
```

In [192]:

```
description_odd.insert(1,'TonightIncreasing clouds, with a low around 57. West wind 13 to 18 mph decreasing to 6 to 11 mp
)
```

In [193]:

```
description_odd.insert(3, 'Friday NightWidespread haze before 9pm. Mostly clear, with a low around 57. Southwest wind 16
)
```

In [194]:

```
description_odd.insert(5, 'Saturday NightMostly clear, with a low around 58. Breezy.'
)
```

In [195]:

```
description_odd.insert(7, 'Sunday NightPartly cloudy, with a low around 56.'
)
```

In [196]:

```
description_odd.insert(9, 'Monday NightMostly clear, with a low around 54.'
)
```

In [197]:

```
description_odd.insert(11, 'Tuesday NightMostly clear, with a low around 54.'
)
```

In [206]:

```
desc_data=description_odd
desc_data
```

Out[206]:

```
['TodayPartly sunny, then gradually becoming sunny, with a high near 72. Light and variable wind becoming west 15 to 20
mph in the afternoon. Winds could gust as high as 26 mph. ',
 'TonightIncreasing clouds, with a low around 57. West wind 13 to 18 mph decreasing to 6 to 11 mph after midnight. Wind
s could gust as high as 23 mph.',
 'FridayWidespread haze after noon. Sunny, with a high near 74. Light west wind increasing to 15 to 20 mph in the after
noon. Winds could gust as high as 25 mph. ',
 'Friday NightWidespread haze before 9pm. Mostly clear, with a low around 57. Southwest wind 16 to 21 mph decreasing to
6 to 11 mph after midnight. Winds could gust as high as 26 mph.',
 'SaturdaySunny, with a high near 75. Light west southwest wind increasing to 13 to 18 mph in the afternoon. Winds coul
d gust as high as 24 mph. ',
 'Saturday NightMostly clear, with a low around 58. Breezy.',
 'SundaySunny, with a high near 73.',
 'Sunday NightPartly cloudy, with a low around 56.',
 'MondayMostly sunny, with a high near 67.',
 'Monday NightMostly clear, with a low around 54.',
 'TuesdaySunny, with a high near 66.',
 'Tuesday NightMostly clear, with a low around 54.',
 'WednesdaySunny, with a high near 65.']
```

In [210]:

```
df=pd.DataFrame(description_odd)
df
```

Out[210]:

	0
0	TodayPartly sunny, then gradually becoming sun...
1	TonightIncreasing clouds, with a low around 57...
2	FridayWidespread haze after noon. Sunny, with ...
3	Friday NightWidespread haze before 9pm. Mostly...
4	SaturdaySunny, with a high near 75. Light west...
5	Saturday NightMostly clear, with a low around ...
6	SundaySunny, with a high near 73.
7	Sunday NightPartly cloudy, with a low around 56.
8	MondayMostly sunny, with a high near 67.
9	Monday NightMostly clear, with a low around 54.
10	TuesdaySunny, with a high near 66.
11	Tuesday NightMostly clear, with a low around 54.
12	WednesdaySunny, with a high near 65.

In [211]:

```
detailed_forecast=pd.DataFrame({})
detailed_forecast['period_name']=period_name
detailed_forecast['short_description']=short_description
detailed_forecast['temp_data']=temp_data
detailed_forecast['desc_data']=desc_data[:9]
detailed_forecast
```

Out[211]:

	period_name	short_description	temp_data	desc_data
0	Today	BecomingSunny	High: 72 °F	TodayPartly sunny, then gradually becoming sun...
1	Tonight	IncreasingClouds	Low: 57 °F	TonightIncreasing clouds, with a low around 57...
2	Friday	Sunny thenHaze	High: 74 °F	FridayWidespread haze after noon. Sunny, with ...
3	FridayNight	Haze thenMostly Clear	Low: 57 °F	Friday NightWidespread haze before 9pm. Mostly...
4	Saturday	Sunny	High: 75 °F	SaturdaySunny, with a high near 75. Light west...
5	SaturdayNight	Mostly Clearand Breezythen MostlyClear	Low: 58 °F	Saturday NightMostly clear, with a low around ...
6	Sunday	Sunny	High: 73 °F	SundaySunny, with a high near 73.
7	SundayNight	Partly Cloudy	Low: 56 °F	Sunday NightPartly cloudy, with a low around 56.
8	Monday	Mostly Sunny	High: 67 °F	MondayMostly sunny, with a high near 67.

In [ ]:

```

```

9. Write a python program to scrape fresher job listings from ‘https://internshala.com/’ (https://internshala.com/%E2%80%99). It should include job title, company name, CTC, and apply date.

In [299]:

```
page=requests.get("https://internshala.com/fresher-jobs")
page
```

Out[299]:

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

In [300]:
```
soup=BeautifulSoup(page.content,"lxml")
soup
```

Out[300]:
```
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:fb="https://www.facebook.com/2008/fbml" xmlns:og="http://ogp.me/ns
#">
<head>
<meta content="IE=9" http-equiv="X-UA-Compatible"/>
<meta charset="utf-8"/>
<meta content="width=device-width, initial-scale=1.0 user-scalable=0" name="viewport"/>
<meta content="272234782795210" property="fb:app_id"/>
<meta content="article" property="og:type"/>
<meta content="1200" property="og:image:width"/>
<meta content="630" property="og:image:height"/>
<meta content="@Internshala" name="twitter:site"/>
<meta content="summary_large_image" name="twitter:card"/>
<meta content="@internshala" name="twitter:creator"/>
<meta content="#ffffff" name="theme-color"/>
<meta content="#ffffff" name="msapplication-navbutton-color"/>
<meta content="telephone=no" name="format-detection"/>
<link as="font" crossorigin="" href="/static/fonts/Inter.woff2?v=3.11" rel="preload" type="font/woff2"/>
<link as="font" crossorigin="" href="/static/fonts/internshala-icons.woff2?8zxqsu" rel="preload" type="font/woff2"/>
```

In [301]:
```
title=soup.find_all("div",class_="heading_4_5 profile")
title
```

Out[301]:
```
[<div class="heading_4_5 profile">
 <a href="/fresher-job/detail/omni-sport-leader-fresher-jobs-in-hyderabad-at-decathlon-sport-india-private-limited16
29355813">OMNI SPORT LEADER</a> </div>,
 <div class="heading_4_5 profile">
 <a href="/fresher-job/detail/executive-senior-executive-partnerships-fresher-jobs-in-multiple-locations-at-freechar
ge-payments-technology-private-limited1628770904">Executive/Senior Executive - Partnerships</a> </div>,
 <div class="heading_4_5 profile">
 <a href="/fresher-job/detail/executive-sales-fresher-jobs-in-multiple-locations-at-freecharge-payments-technology-p
rivate-limited1628770715">Executive - Sales</a> </div>,
 <div class="heading_4_5 profile">
 <a href="/fresher-job/detail/junior-operations-executive-fresher-jobs-in-gurgaon-at-freecharge-payments-technology-
private-limited1628075825">Junior Operations Executive</a> </div>,
 <div class="heading_4_5 profile">
 <a href="/fresher-job/detail/remote-student-success-teacher-fresher-jobs-at-code-a-block-kedtech-llc1629488510">Stu
dent Success - Teacher</a> </div>,
 <div class="heading_4_5 profile">
 <a href="/fresher-job/detail/supply-chain-executive-fresher-jobs-in-mumbai-at-serturner1629870500">Supply Chain Exe
cutive</a> </div>,
 <div class="heading_4_5 profile">
```

In [302]:

```
job_title=[]
for i in range(0,len(title)):
    job_title.append(title[i].get_text().replace("\n"," "))
job_title
```

Out[302]:

```
[' OMNI SPORT LEADER ',
 ' Executive/Senior Executive - Partnerships ',
 ' Executive - Sales ',
 ' Junior Operations Executive ',
 ' Student Success - Teacher ',
 ' Supply Chain Executive ',
 ' Game Developer ',
 ' Associate - Digital Products ',
 ' Program Impact In-charge (Science And Mathematics) ',
 ' Software Engineer Trainee ',
 ' Sales Executive ',
 ' Digital Marketing Executive ',
 ' Corporate Sales Manager ',
 ' Software Developer ',
 ' Executive Assistant To Director ',
 ' C++ Embedded/Semiconductor Expert (SoC Modeling) ',
 ' Mobile App Developer ',
 ' Software Developer ',
 ' Coding Teacher (Remote) ',
 ' Validation Expert ',
 ' Full Stack Developer ',
 ' Business Development Associate ',
 ' Content & E-commerce Management Executive ',
 ' Business Development Specialist ',
 ' Assistant Manager - Outreach & Events ',
 ' Investor Pitch Deck - Content Writer ',
 ' Full Stack Engineer (MERN) ',
 ' Business Development Executive ',
 ' Junior Graphic Designer ',
 ' Sales Development Representative ',
 ' Junior Sales Executive ',
 ' Online Coding Educator ',
 ' Executive - Statistical Modeling ',
 ' Customer Experience And Sales Advisor ',
 ' Firmware Engineer ',
 ' Junior Software Developer ',
 ' Full Stack Engineer ',
 ' Backend Developer (Java/Spring Boot/Hibernate) ',
 ' Associate Content Writer ',
 ' Social Media Manager ']
```

In [303]:

```
import pandas as pd
df=pd.DataFrame(job_title)
df
```

Out[303]:

	0
0	OMNI SPORT LEADER
1	Executive/Senior Executive - Partnerships
2	Executive - Sales
3	Junior Operations Executive
4	Student Success - Teacher
5	Supply Chain Executive
6	Game Developer
7	Associate - Digital Products
8	Program Impact In-charge (Science And Mathema...
9	Software Engineer Trainee
10	Sales Executive
11	Digital Marketing Executive
12	Corporate Sales Manager
13	Software Developer
14	Executive Assistant To Director
15	C++ Embedded/Semiconductor Expert (SoC Modeli...
16	Mobile App Developer
17	Software Developer
18	Coding Teacher (Remote)
19	Validation Expert
20	Full Stack Developer
21	Business Development Associate
22	Content & E-commerce Management Executive

	0
23	Business Development Specialist
24	Assistant Manager - Outreach & Events
25	Investor Pitch Deck - Content Writer
26	Full Stack Engineer (MERN)
27	Business Development Executive
28	Junior Graphic Designer
29	Sales Development Representative
30	Junior Sales Executive
31	Online Coding Educator
32	Executive - Statistical Modeling
33	Customer Experience And Sales Advisor
34	Firmware Engineer
35	Junior Software Developer
36	Full Stack Engineer
37	Backend Developer (Java/Spring Boot/Hibernate)
38	Associate Content Writer
39	Social Media Manager

```
In [ ]:
```

In [304]:
```
name=soup.find_all("div",class_="heading_6 company_name")
name
```

Out[304]:
```
[<div class="heading_6 company_name">
 <a class="link_display_like_text" href="/fresher-jobs/fresher-job-at-Decathlon%20Sport%20India%20Private%20Limite
d">
                         Decathlon Sport India Private Limited                    </a>
 </div>,
 <div class="heading_6 company_name">
 <a class="link_display_like_text" href="/fresher-jobs/fresher-job-at-Freecharge%20Payments%20Technology%20Private%2
0Limited">
                         Freecharge Payments Technology Private Limited                    </a>
 </div>,
 <div class="heading_6 company_name">
 <a class="link_display_like_text" href="/fresher-jobs/fresher-job-at-Freecharge%20Payments%20Technology%20Private%2
0Limited">
                         Freecharge Payments Technology Private Limited                    </a>
 </div>,
 <div class="heading_6 company_name">
 <a class="link_display_like_text" href="/fresher-jobs/fresher-job-at-Freecharge%20Payments%20Technology%20Private%2
0Limited">
                         Freecharge Payments Technology Private Limited                    </a>
 </div>,
```

In [305]:

```
company_name=[]
for i in range(0,len(name)):
    company_name.append(name[i].get_text().replace("\n",""))
company_name
```

Out[305]:

```
['                        Decathlon Sport India Private Limited                    ',
 '                        Freecharge Payments Technology Private Limited                    ',
 '                        Freecharge Payments Technology Private Limited                    ',
 '                        Freecharge Payments Technology Private Limited                    ',
 '                        Code A Block (KEDTech LLC)                    ',
 '                        Serturner                    ',
 '                        Cogent Web Services                    ',
 '                        RocSearch                    ',
 '                        Open Door Education                    ',
 '                        Markytics                    ',
 '                        1st In Class                    ',
 '                        TradeVinder                    ',
 '                        FinDestination                    ',
 '                        Innovantes IT Solutions LLP                    ',
 '                        Best Roadways Limited                    ',
 '                        Circuitsutra Technologies Private Limited                    ',
 '                        Cogent Web Services                    ',
 '                        Swabhav Techlabs                    ',
 '                        Codingal Education Private Limited                    ',
 '                        Apna                    ',
 '                        ImbueDesk                    ',
 '                        Tutedude                    ',
 '                        Anmol Baby                    ',
 '                        The SmartWare                    ',
 '                        Formskart                    ',
 '                        PitchyDeck                    ',
 '                        Softway Solutions Private Limited                    ',
 '                        Volopay                    ',
 '                        Habitate Technologies Private Limited                    ',
 '                        Ziguar                    ',
 '                        Picocrew                    ',
 '                        BrightChamps Tech Private Limited                    ',
 "                        D'Well Research                    ",
 '                        Tilfi Banaras                    ',
 '                        NBase2 Systems Private Limited                    ',
 '                        Finnovus Technologies Private Limited (CreditSiddhi)                    ',
 '                        Getart                    ',
 '                        Krenai                    ',
 '                        MiM-Essay                    ',
 '                        Finvision Financial Services                    ']
```

```
In [306]: import pandas as pd
          df=pd.DataFrame(company_name)
          df
```

```
Out[306]:
                                0
     0        Decathlon Sport India ...
     1     Freecharge Payments Te...
     2     Freecharge Payments Te...
     3     Freecharge Payments Te...
     4     Code A Block (KEDTech ...
     5                    Serturner ...
     6        Cogent Web Services ...
     7                    RocSearch ...
     8       Open Door Education ...
     9                    Markytics ...
    10                   1st In Class ...
    11                  TradeVinder ...
    12                FinDestination ...
    13       Innovantes IT Solution...
    14      Best Roadways Limited ...
    15       Circuitsutra Technolog...
    16        Cogent Web Services ...
    17           Swabhav Techlabs ...
    18       Codingal Education Pri...
    19                            Apna
    20                   ImbueDesk ...
    21                    Tutedude ...
    22                 Anmol Baby ...
```

	0
23	The SmartWare ...
24	Formskart ...
25	PitchyDeck ...
26	Softway Solutions Priv...
27	Volopay ...
28	Habitate Technologies ...
29	Ziguar ...
30	Picocrew ...
31	BrightChamps Tech Priv...
32	D'Well Research ...
33	Tilfi Banaras ...
34	NBase2 Systems Private...
35	Finnovus Technologies ...
36	Getart ...
37	Krenai ...
38	MiM-Essay ...
39	Finvision Financial Se...

In [329]:

```
ctc=soup.find_all("div",class_="item_body")
ctc
```

Out[329]:

```
[<div class="item_body" id="start-date-first">
 <span class="start_immediately_mobile">Starts </span>Immediately
                         </div>,
 <div class="item_body">
 <i class="ic-16-currency-inr"></i>  3 - 4 LPA                                                    </div>,
 <div class="item_body">18 Sep' 21</div>,
 <div class="item_body" id="start-date-first">
 <span class="start_immediately_mobile">Starts </span>Immediately
                         </div>,
 <div class="item_body">
 <i class="ic-16-currency-inr"></i>  3 - 4.2 LPA                                                    </div>,
 <div class="item_body">11 Sep' 21</div>,
 <div class="item_body" id="start-date-first">
 <span class="start_immediately_mobile">Starts </span>Immediately
                         </div>,
 <div class="item_body">
 <i class="ic-16-currency-inr"></i>  3 - 3.5 LPA                                                    </div>,
 <div class="item_body">11 Sep' 21</div>,
 <div class="item_body" id="start-date-first">
 <span class="start_immediately_mobile">Starts </span>Immediately
```

In [324]:

```
company_ctc=[]
for i in range(0,len(ctc)):
        company_ctc.append(ctc[i].get_text().replace("\n",""))
company_ctc
```

Out[324]:

```
['Starts\xa0Immediately                          ',
 '  3 - 4 LPA                                                     ',
 "18 Sep' 21",
 'Starts\xa0Immediately                         ',
 '  3 - 4.2 LPA                                                     ',
 "11 Sep' 21",
 'Starts\xa0Immediately                         ',
 '  3 - 3.5 LPA                                                     ',
 "11 Sep' 21",
 'Starts\xa0Immediately                         ',
 '  3 - 4 LPA                                                     ',
 "4 Sep' 21",
 'Starts\xa0Immediately                         ',
 '  4 - 7.2 LPA                                                     ',
 "25 Sep' 21",
 'Starts\xa0Immediately                         ',
 '  3 LPA                                                     ',
 "25 Sep' 21",
 'Starts\xa0Immediately                         ',
```

In [325]:

```
import pandas as pd
```

In [332]:

```
df= pd.DataFrame(company_ctc)
df
```

Out[332]:

```
                      0
  0   Starts Immediately
  1           3 - 4 LPA ...
  2          18 Sep' 21
  3   Starts Immediately
  4         3 - 4.2 LPA ...
...                   ...
115         4 - 4.5 LPA ...
116          22 Sep' 21
117   Starts Immediately
118         3 - 3.5 LPA ...
119          22 Sep' 21

120 rows × 1 columns
```

In [313]:

```
import csv
company_details=pd.DataFrame({})
```

In [314]:

```
company_details['job_title']=job_title
company_details['company_name']=company_name
company_details
```

Out[314]:

	job_title	company_name
0	OMNI SPORT LEADER	Decathlon Sport India ...
1	Executive/Senior Executive - Partnerships	Freecharge Payments Te...
2	Executive - Sales	Freecharge Payments Te...
3	Junior Operations Executive	Freecharge Payments Te...
4	Student Success - Teacher	Code A Block (KEDTech ...
5	Supply Chain Executive	Serturner ...
6	Game Developer	Cogent Web Services ...
7	Associate - Digital Products	RocSearch ...
8	Program Impact In-charge (Science And Mathema...	Open Door Education ...
9	Software Engineer Trainee	Markytics ...
10	Sales Executive	1st In Class ...
11	Digital Marketing Executive	TradeVinder ...
12	Corporate Sales Manager	FinDestination ...
13	Software Developer	Innovantes IT Solution...
14	Executive Assistant To Director	Best Roadways Limited ...
15	C++ Embedded/Semiconductor Expert (SoC Modeli...	Circuitsutra Technolog...
16	Mobile App Developer	Cogent Web Services ...
17	Software Developer	Swabhav Techlabs ...
18	Coding Teacher (Remote)	Codingal Education Pri...
19	Validation Expert	Apna
20	Full Stack Developer	ImbueDesk ...
21	Business Development Associate	Tutedude ...
22	Content & E-commerce Management Executive	Anmol Baby ...

	job_title	company_name
23	Business Development Specialist	The SmartWare ...
24	Assistant Manager - Outreach & Events	Formskart ...
25	Investor Pitch Deck - Content Writer	PitchyDeck ...
26	Full Stack Engineer (MERN)	Softway Solutions Priv...
27	Business Development Executive	Volopay ...
28	Junior Graphic Designer	Habitate Technologies ...
29	Sales Development Representative	Ziguar ...
30	Junior Sales Executive	Picocrew ...
31	Online Coding Educator	BrightChamps Tech Priv...
32	Executive - Statistical Modeling	D'Well Research ...
33	Customer Experience And Sales Advisor	Tilfi Banaras ...
34	Firmware Engineer	NBase2 Systems Private...
35	Junior Software Developer	Finnovus Technologies ...
36	Full Stack Engineer	Getart ...
37	Backend Developer (Java/Spring Boot/Hibernate)	Krenai ...
38	Associate Content Writer	MiM-Essay ...
39	Social Media Manager	Finvision Financial Se...

10. Write a python program to scrape house details from https://www.nobroker.in/ (https://www.nobroker.in/) for any location. It should include house title, location, area, emi and price

In [96]:

```
from csv import writer
```

In [58]:
```
page=requests.get("https://www.nobroker.in/property/sale/pune/multiple?searchParam=W3sibGF0IjoxOC41MzYyMDg0LCJsb24iOjczL
page
```

Out[58]:
```
<Response [200]>
```

In [59]:
```
soup=BeautifulSoup(page.content,"html.parser")
soup
```

Out[59]:
```
<!DOCTYPE html>
<html lang="en"><head>
<meta content="794951570520699" property="fb:pages"/>
<link href="https://www.nobroker.in" rel="canonical"/>
<link href="//assets.nobroker.in/static/img/favicon.png" id="favicon" rel="shortcut icon"/>
<link href="https://images.nobroker.in/static/img/fav64.png" rel="apple-touch-icon"/>
<meta charset="utf-8"/><meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<meta content="app-id=com.nobroker.app&amp;referrer=utm_source%3Dnobroker%26utm_medium%3DmobileWeb" name="google-pla
y-app"/>
<meta content="app-id=1200507100, app-argument=nobrokerapp://" name="apple-itunes-app"/>
<meta content="#fd3752" name="theme-color"/>
<meta content="BHK2,BHK3,BHK4,BHK4PLUS  flats for sale in Koregaon Park,Viman Nagar, BHK2,BHK3,BHK4,BHK4PLUS  apartm
ents for sale in Koregaon Park,Viman Nagar,  houses for sale in Koregaon Park,Viman Nagar" name="keywords">
<meta content="BHK2,BHK3,BHK4,BHK4PLUS  Flats, Apartments for Sale in Koregaon Park,Viman Nagar, Pune | BHK2,BHK3,BH
K4,BHK4PLUS Flats in  Koregaon Park,Viman Nagar -   NoBroker" property="og:title">
<meta content="Searching BHK2,BHK3,BHK4,BHK4PLUS   Flats, Apartments for Sale in Koregaon Park,Viman Nagar, Pune ? G
et verified list of flats for sale WITHOUT BROKER and SAVE YOUR BROKERAGE at NoBroker." property="og:description">
<meta content="https://www.nobroker.in" property="og:url"/>
<meta content="https://assets.nobroker.in/nb-new/public/List-Page/ogImage.png" property="og:image"/>
```

In [60]:
```
title=soup.find_all("h2",class_="heading-6 font-semi-bold nb__1AShY")
title
```

Out[60]:
```
[<h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Flat  For Sale  In Reverie Apartments In Va
dgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path
d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h
-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Flat  For Sale  In Reverie Apartments In Va
dgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path
d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h
-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Apartment  For Sale  In Gurukrupa Astter In
Wadgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><pa
th d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v
-7h-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Apartment  For Sale  In Gurukrupa Astter In
Wadgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><pa
th d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v
-7h-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Apartment  For Sale  In Gurukrupa Astter In
Wadgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><pa
th d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v
-7h-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Flat  For Sale  In Spring Dale Society In K
haradi</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path d="M
0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h-2v7
zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Apartment  For Sale  In Kalp Homes In Vadga
on Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path d
="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h-
2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Flat  For Sale  In Nav Deviya Home In Wadga
on Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path d
="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h-
2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">2 BHK Apartment  For Sale  In Prestige Kalphomes
In Vadgaon Sheri</span> <svg aria-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24">
<path d="M0 0h24v24H0z" fill="none"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2
-2v-7h-2v7zM14 3v2h3.59l-9.83 9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>,
 <h2 class="heading-6 font-semi-bold nb__1AShY"><span itemprop="name">4+ BHK Flat  For Sale  In Kharadi,</span> <svg ar
ia-hidden="true" class="jss1" focusable="false" role="presentation" viewbox="0 0 24 24"><path d="M0 0h24v24H0z" fill="n
one"></path><path d="M19 19H5V5h7V3H5c-1.11 0-2 .9-2 2v14c0 1.1.89 2 2 2h14c1.1 0 2-.9 2-2v-7h-2v7zM14 3v2h3.59l-9.83
9.83 1.41 1.41L19 6.41V10h2V3h-7z"></path></svg></h2>]
```

In [61]:

```
house_title=[]
for i in range(0,len(title)):
    house_title.append(title[i].get_text().replace("\n",""))
house_title
```

Out[61]:

```
['2 BHK Flat  For Sale  In Reverie Apartments In Vadgaon Sheri ',
 '2 BHK Flat  For Sale  In Reverie Apartments In Vadgaon Sheri ',
 '2 BHK Apartment  For Sale  In Gurukrupa Astter In Wadgaon Sheri ',
 '2 BHK Apartment  For Sale  In Gurukrupa Astter In Wadgaon Sheri ',
 '2 BHK Apartment  For Sale  In Gurukrupa Astter In Wadgaon Sheri ',
 '2 BHK Flat  For Sale  In Spring Dale Society In Kharadi ',
 '2 BHK Apartment  For Sale  In Kalp Homes In Vadgaon Sheri ',
 '2 BHK Flat  For Sale  In Nav Deviya Home In Wadgaon Sheri ',
 '2 BHK Apartment  For Sale  In Prestige Kalphomes In Vadgaon Sheri ',
 '4+ BHK Flat  For Sale  In Kharadi, ']
```

In [62]:

```
location=soup.find_all("div",class_="nb__2CMjv")
location
```

Out[62]:

```
[<div class="nb__2CMjv">Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtra 411014, India</div>,
 <div class="nb__2CMjv">Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtra 411014, India</div>,
 <div class="nb__2CMjv"><a class="nb__17tcx" href="/gurukrupa-astter-wadgaon-sheri_pune-prjt-5ba3f027714b56aa0857151d"
target="_blank">Gurukrupa Astter</a>  <!-- -->Behind Bollywood Theatre, Wadgaon Sheri, Pune, Maharashtra, INDIA.</div>,
 <div class="nb__2CMjv"><a class="nb__17tcx" href="/gurukrupa-astter-wadgaon-sheri_pune-prjt-5ba3f027714b56aa0857151d"
target="_blank">Gurukrupa Astter</a>  <!-- -->Wadgaon Sheri, Pune, Maharashtra, INDIA.</div>,
 <div class="nb__2CMjv"><a class="nb__17tcx" href="/gurukrupa-astter-wadgaon-sheri_pune-prjt-5ba3f027714b56aa0857151d"
target="_blank">Gurukrupa Astter</a>  <!-- -->Behind Bollywood Theatre, Wadgaon Sheri, Pune, Maharashtra, INDIA.</div>,
 <div class="nb__2CMjv">Old Kharadi Mundhwa Road Near Bollywood Multiplex</div>,
 <div class="nb__2CMjv"><a class="nb__17tcx" href="/kalpak-homes-katraj_pune-prjt-5ba3f027714b56aa0857146e" target="_bl
ank">Kalp Homes</a>  <!-- -->Old Mundhwa Rd, Near Bollywood Multiplex</div>,
 <div class="nb__2CMjv">Mahadev Nagar, Near Bollywood E-Square</div>,
 <div class="nb__2CMjv"><a class="nb__17tcx" href="/prestige-kalphomes-vadgaon-sheri_pune-prjt-5ba3f027714b56aa0857158
0" target="_blank">Prestige Kalphomes</a>  <!-- -->Sr. No 52/2, Near Bollywood Multiplex, Vadgaon Sheri, Pune, Maharash
tra, INDIA.</div>,
 <div class="nb__2CMjv">Standalone Building, Old Mundhwa Rd, near Pavan Hotel</div>]
```

In [63]:

```
house_location=[]
for i in range(0,len(location)):
    house_location.append(location[i].get_text().replace("\n",""))
house_location
```

Out[63]:

```
['Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtra 411014, India',
 'Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtra 411014, India',
 'Gurukrupa Astter\xa0 Behind Bollywood Theatre, Wadgaon Sheri, Pune, Maharashtra, INDIA.',
 'Gurukrupa Astter\xa0 Wadgaon Sheri, Pune, Maharashtra, INDIA.',
 'Gurukrupa Astter\xa0 Behind Bollywood Theatre, Wadgaon Sheri, Pune, Maharashtra, INDIA.',
 'Old Kharadi Mundhwa Road Near Bollywood Multiplex',
 'Kalp Homes\xa0 Old Mundhwa Rd, Near Bollywood Multiplex',
 'Mahadev Nagar, Near Bollywood E-Square',
 'Prestige Kalphomes\xa0 Sr. No 52/2, Near Bollywood Multiplex, Vadgaon Sheri, Pune, Maharashtra, INDIA.',
 'Standalone Building, Old Mundhwa Rd, near Pavan Hotel']
```

In [64]:

```
area=soup.find_all("div",class_="nb__3oNyC")
area
```

Out[64]:

```
[<div class="nb__3oNyC">803 sqft</div>,
 <div class="nb__3oNyC">750 sqft</div>,
 <div class="nb__3oNyC">865 sqft</div>,
 <div class="nb__3oNyC">865 sqft</div>,
 <div class="nb__3oNyC">910 sqft</div>,
 <div class="nb__3oNyC">756 sqft</div>,
 <div class="nb__3oNyC">880 sqft</div>,
 <div class="nb__3oNyC">903 sqft</div>,
 <div class="nb__3oNyC">940 sqft</div>,
 <div class="nb__3oNyC">3,000 sqft</div>]
```

In [65]:

```
house_area=[]
for i in range(0,len(area)):
    house_area.append(area[i].get_text().replace("\n"," "))
house_area
```

Out[65]:

```
['803 sqft',
 '750 sqft',
 '865 sqft',
 '865 sqft',
 '910 sqft',
 '756 sqft',
 '880 sqft',
 '903 sqft',
 '940 sqft',
 '3,000 sqft']
```

In [66]:

```
emi=soup.find_all("div",class_="font-semi-bold heading-6",id="roomType")
emi
```

Out[66]:

```
[<div class="font-semi-bold heading-6" id="roomType"><span>₹</span>31,522/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>34,388/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>36,681/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>37,254/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>33,815/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>31,522/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>37,254/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>37,254/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>37,254/Month</div>,
 <div class="font-semi-bold heading-6" id="roomType"><span>₹</span>85,971/Month</div>]
```

In [67]:

```
house_emi=[]
for i in range(0,len(emi)):
    house_emi.append(emi[i].get_text().replace("\n",""))
house_emi
```

Out[67]:

```
['₹31,522/Month',
 '₹34,388/Month',
 '₹36,681/Month',
 '₹37,254/Month',
 '₹33,815/Month',
 '₹31,522/Month',
 '₹37,254/Month',
 '₹37,254/Month',
 '₹37,254/Month',
 '₹85,971/Month']
```

In [76]:
```
price=soup.find_all("div",class_="nb__2NPHR",id="minDeposit")
price
```

Out[76]:
```
[<div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->55 Lacs</span></div><div cl
ass="heading-7">₹6,849 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->60 Lacs</span></div><div cl
ass="heading-7">₹8,000 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->64 Lacs</span></div><div cl
ass="heading-7">₹7,399 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->65 Lacs</span></div><div cl
ass="heading-7">₹7,514 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->59 Lacs</span></div><div cl
ass="heading-7">₹6,484 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->55 Lacs</span></div><div cl
ass="heading-7">₹7,275 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->65 Lacs</span></div><div cl
ass="heading-7">₹7,386 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->65 Lacs</span></div><div cl
ass="heading-7">₹7,198 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->65 Lacs</span></div><div cl
ass="heading-7">₹6,915 per sq.ft.</div></div>,
 <div class="nb__2NPHR" id="minDeposit" itemprop="valueReference" itemtype="http://schema.org/PropertyValue"><div class
="font-semi-bold heading-6"><meta content="Minimum Deposit" itemprop="name"/><span>₹<!-- -->1.5 Crores</span></div><div
class="heading-7">₹5,000 per sq.ft.</div></div>]
```

In [79]:

```
house_price=[]
for i in range(0,len(price)):
    house_price.append(price[i].get_text().replace("\n"," "))
house_price
```

Out[79]:

```
['₹55 Lacs₹6,849 per sq.ft.',
 '₹60 Lacs₹8,000 per sq.ft.',
 '₹64 Lacs₹7,399 per sq.ft.',
 '₹65 Lacs₹7,514 per sq.ft.',
 '₹59 Lacs₹6,484 per sq.ft.',
 '₹55 Lacs₹7,275 per sq.ft.',
 '₹65 Lacs₹7,386 per sq.ft.',
 '₹65 Lacs₹7,198 per sq.ft.',
 '₹65 Lacs₹6,915 per sq.ft.',
 '₹1.5 Crores₹5,000 per sq.ft.']
```

In [ ]:

In [80]:

```
import pandas as pd
house_details=pd.DataFrame({})
house_details['house_title']=house_title
house_details['house_location']=house_location
house_details['house_area']=house_area
house_details['house_emi']=house_emi
house_details['house_price']=house_price
house_details
```

Out[80]:

	house_title	house_location	house_area	house_emi	house_price
0	2 BHK Flat For Sale In Reverie Apartments In...	Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtr...	803 sqft	₹31,522/Month	₹55 Lacs₹6,849 per sq.ft.
1	2 BHK Flat For Sale In Reverie Apartments In...	Mahadev Nagar, Vadgaon Sheri, Pune, Maharashtr...	750 sqft	₹34,388/Month	₹60 Lacs₹8,000 per sq.ft.
2	2 BHK Apartment For Sale In Gurukrupa Astter...	Gurukrupa Astter Behind Bollywood Theatre, Wa...	865 sqft	₹36,681/Month	₹64 Lacs₹7,399 per sq.ft.
3	2 BHK Apartment For Sale In Gurukrupa Astter...	Gurukrupa Astter Wadgaon Sheri, Pune, Maharas...	865 sqft	₹37,254/Month	₹65 Lacs₹7,514 per sq.ft.
4	2 BHK Apartment For Sale In Gurukrupa Astter...	Gurukrupa Astter Behind Bollywood Theatre, Wa...	910 sqft	₹33,815/Month	₹59 Lacs₹6,484 per sq.ft.
5	2 BHK Flat For Sale In Spring Dale Society I...	Old Kharadi Mundhwa Road Near Bollywood Multiplex	756 sqft	₹31,522/Month	₹55 Lacs₹7,275 per sq.ft.
6	2 BHK Apartment For Sale In Kalp Homes In Va...	Kalp Homes Old Mundhwa Rd, Near Bollywood Mul...	880 sqft	₹37,254/Month	₹65 Lacs₹7,386 per sq.ft.
7	2 BHK Flat For Sale In Nav Deviya Home In Wa...	Mahadev Nagar, Near Bollywood E-Square	903 sqft	₹37,254/Month	₹65 Lacs₹7,198 per sq.ft.
8	2 BHK Apartment For Sale In Prestige Kalphom...	Prestige Kalphomes Sr. No 52/2, Near Bollywoo...	940 sqft	₹37,254/Month	₹65 Lacs₹6,915 per sq.ft.
9	4+ BHK Flat For Sale In Kharadi,	Standalone Building, Old Mundhwa Rd, near Pava...	3,000 sqft	₹85,971/Month	₹1.5 Crores₹5,000 per sq.ft.

In [ ]: