

# Data X Berkeley

## Plaksha Web scraping assignment

---

### Assignment problem

In this week's assignment you should extract live weather data in Mohali for the next 15 days:

<https://weather.com/en-IN/weather/tenday/l/d253814420804f28d49f7a75130e71b797c3c93f9ded842eaf6769e3784da352>

- Task scrape
  - day | period (as Wed 11 | Day, Wed 11 | Night etc.)
  - the temperature for the period
  - the short description (e.g. Mostly Clear, Sunny etc.)
  - the long weather description (e.g. Showers late at night. Low 12°C. Winds SE at 10 to 15 km/h. Chance of rain 40%.)

Store the scraped data strings in a Pandas DataFrame

**Hint:** The weather information is found in a div tag with `class_='DailyForecast--DisclosureList--nosQS'`

Your DataFrame should be similar to the below screenshot (with the same columns):

### Assignment solution

```
In [1]: # Necessary Libraries
import requests
import bs4 as bs
import pandas as pd
import numpy as np

In [2]: source = requests.get('https://weather.com/en-IN/weather/tenday/l/d253814420804f28d49f7a75130e71b797c3c93f9ded842eaf6769e3784da352')

In [3]: source

Out[3]: <Response [200]>
```

```
In [4]: soup = bs.BeautifulSoup(source.content, features = 'html.parser')
```

## Approach:

"DaypartDetails--DayPartDetail--2XOOV Disclosure--themeList--1Dz21"

This is a class which gives the consolidated weather data for each and every day. There are a total of 15 such classes.

'DaypartDetails--Content--2Yg3\_ DaypartDetails--contentGrid--2\_szQ'

This is the subclass which gives the entire weather data for a particular day.

'DailyContent--DailyContent--1yRkH'.

This class gives the top part of the weather info of a day. It includes the info about Date, Temperature, SHort description and Detailed description. There are 2 such classes for each day(one for day and other for night).

```
In [5]: desc_list = []          # This list stores the weather description for all dates
temp_list = []                # This list stores temperature values for all days
name_list = []                # This list stores name of day (Like this, Sat 21 | Night)
short_desc_list = []          # This list stores short weather description of the day
d = soup.find_all(class_ = "DaypartDetails--DayPartDetail--2XOOV Disclosure--themeList")
for element in d:
    tag = element.find(class_ = 'DaypartDetails--Content--2Yg3_ DaypartDetails--contentGrid--2_szQ')
    day_night = tag.find_all(class_ = 'DailyContent--DailyContent--1yRkH')
    for time in day_night:
        desc = time.find(class_ = 'DailyContent--narrative--3Ti6_')
        desc_list.append(desc.text)
        temp = time.find(class_ = 'DailyContent--temp--1s3a7')
        temp_list.append(temp.text)
        name = time.find(class_ = 'DailyContent--daypartName--3emSU')
        name_list.append(name.text)
        short_desc = time.find(class_ = 'DailyContent--Condition--1zRBJ')
        short_desc_list.append(short_desc.text)
```

```
In [6]: # Converting all temperatures, descriptions, short descriptions and dates into dataframe
df = pd.DataFrame(list(zip(name_list, temp_list, short_desc_list, desc_list)), columns = ['name', 'temp', 'short_desc', 'desc'])
```

```
In [7]: df
```

Out[7]:

	Day	Temp	Short Desc	Desc
0	Sat 21   Night	7°	Clear Night	Generally clear. Low 7°C. Winds N and variable.
1	Sun 22   Day	20°	Sunny	Generally clear. High 20°C. Winds NNW and vari...
2	Sun 22   Night	9°	Clear Night	Generally clear. Low 9°C. Winds N and variable.
3	Mon 23   Day	21°	Sunny	Generally clear. High 21°C. Winds NW and varia...
4	Mon 23   Night	11°	Scattered Showers Night	Showers late at night. Low 11°C. Winds NNE and...
5	Tue 24   Day	18°	Rain	Rain. High 18°C. Winds SE at 10 to 15 km/h. Ch...
6	Tue 24   Night	12°	Rain	Rain. Low 12°C. Winds SE and variable. Chance ...
7	Wed 25   Day	18°	Thunderstorm	Thunderstorms. High 18°C. Winds S at 10 to 15 ...
8	Wed 25   Night	11°	Thunderstorm	Thunderstorms. Hazy. Low 11°C. Winds N and var...
9	Thu 26   Day	19°	Isolated Thunderstorms	Partly cloudy, with a possible thunderstorm. H...
10	Thu 26   Night	10°	Thunderstorm	Late night thunderstorms. Low 10°C. Winds NNW ...
11	Fri 27   Day	20°	Mostly Sunny	Generally clear. High 20°C. Winds WNW and vari...
12	Fri 27   Night	11°	Partly Cloudy Night	Partly cloudy. Low 11°C. Winds NNW and variable.
13	Sat 28   Day	19°	Partly Cloudy	Partly cloudy. High 19°C. Winds WNW at 10 to 1...
14	Sat 28   Night	11°	Scattered Showers	Showers. Low 11°C. Winds NE and variable. Chan...
15	Sun 29   Day	16°	Scattered Showers	Showers. High 16°C. Winds SSE at 10 to 15 km/h...
16	Sun 29   Night	8°	Scattered Showers	Showers. Hazy. Low 8°C. Winds E and variable. ...
17	Mon 30   Day	17°	Scattered Showers	Morning showers. High 17°C. Winds WNW at 10 to...
18	Mon 30   Night	8°	Clear Night	Generally clear. Low 8°C. Winds NNW and variable.
19	Tue 31   Day	18°	Partly Cloudy	Partly cloudy. High 18°C. Winds NW at 10 to 15...
20	Tue 31   Night	8°	Clear Night	Generally clear. Low 8°C. Winds NNW and variable.
21	Wed 01   Day	18°	Sunny	Generally clear. High 18°C. Winds NW at 10 to ...
22	Wed 01   Night	8°	Mostly Clear Night	Generally clear. Low 8°C. Winds NNW and variable.
23	Thu 02   Day	18°	Partly Cloudy	Partly cloudy. High 18°C. Winds WNW at 10 to 1...
24	Thu 02   Night	8°	Mostly Clear Night	Generally clear. Low 8°C. Winds N and variable.
25	Fri 03   Day	19°	Sunny	Generally clear. High 19°C. Winds WNW at 10 to...
26	Fri 03   Night	9°	Clear Night	Generally clear. Low 9°C. Winds NNE and variable.
27	Sat 04   Day	20°	Sunny	Generally clear. High 20°C. Winds WNW at 10 to...
28	Sat 04   Night	9°	Mostly Clear Night	Generally clear. Low 9°C. Winds NNE and variable.

In [ ]: