

WEB SCRAPING

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

Every answer to the question should be in form of a python function which should take URL as the parameter. 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'.
2. Write a python program to display IMDB's Top rated 100 movies' data (i.e. Name, IMDB rating, Year of release) and make data frame.
3. Write a python program to display IMDB's Top rated 100 Indian movies' data (i.e. Name, IMDB rating, Year of release) and make data frame.
4. 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 and rating.
 - ii) Top 10 ODI Batsmen in men along with the records of their team and rating.
 - iii) Top 10 ODI bowlers along with the records of their team and rating.
5. Write a python program to scrape cricket rankings from 'www.icc-cricket.com'. You have to scrape:
 - i) Top 10 ODI teams in women's cricket along with the records for matches, points and rating.
 - ii) Top 10 women's ODI players along with the records of their team and rating.
 - iii) Top 10 women's ODI all-rounder along with the records of their team and rating.
6. 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.
7. Write a python program to scrape house details from mentioned url. It should include house title, location, area, emi and price
<https://www.nobroker.in/property/sale/bangalore/Electronic%20City?type=BHK4&searchParam=W3sibGF0IjoxMi44NDUyMTQ1LCJsb24iOjc3LjY2MDE2OTUsInBsYWNIc3Q1OjE0dG91IGUWQ0cHNyanNSSGZkYXpnXzhYRW8iLCJwbGFjZU5hbWUiOiJFbGVjdHJvbmljIENpdHkifV0=&propertyAge=0&radius=2.0>
8. Write a python program to scrape mentioned details from 'https://www.dineout.co.in/delhi-restaurants/buffet-special' :
 - i) Restaurant name
 - ii) Cuisine
 - iii) Location
 - iv) Ratings
 - v) Image url
9. Write a python program to scrape weather details for last 24 hours from 'https://en.tutiempo.net/delhi.html?data=last-24-hours' :
 - i) Hour
 - ii) Temperature
 - iii) Wind
 - iv) Weather condition
 - v) Humidity
 - vi) Pressure
10. Write a python program to scrape monument name, monument description, image url about top 10 monuments from 'https://www.puredestinations.co.uk/top-10-famous-monuments-to-visit-in-india/'