

Afternoon session.

Date - June 2

Course - python

Topic - Scrape real estate
property data from web.

Repository - Rashmi k

Name - Rashmi

USN - 4AL18ECDU3

Sem - IV Sem.

Report:

Scrape real

- web scraping
- with python performing data
- python will load data and keep in given table.
- These attributes are automatic.

Request headers:

r = requests / real estate / rock springs.

r = requests.

headers = {'User-agent': 'mozilla / 5.0 (X11; Ubuntu;
Linux x86_64; rv: 61.0)

Gecko / 20100101 Firefox / 61.0'}

The rest of the code stays the same.

So we just changing the domain name from python to pyclass & were adding a header allows the script to impersonate a web browser.

loading a web page in python:

import replate

from bs4 import BeautifulSoup

Extracting "div" tags:

Extracting Adresses & Properties -

python find - will grab the proper collapse & find address & ignore.

Extracting elements without Unique -

produces the script. price, adress & some attributes features

source code - the actual value to extract

extracting elements & saved the extracted data:

- python dictionary & list will give lead to save the extracted data.
- Give the adres value.

crawling through webpage.