Project_1_1

July 28, 2020

1 Project 1 Exercise 2. UOK website scraping

```python
[1]: from bs4 import BeautifulSoup
     import requests
```

1.1 Get website source

```python
[2]: # get home page of University of Kurdistan website
     home_page = requests.get('https://uok.ac.ir/fa.aspx')
     home_page_souped = BeautifulSoup(home_page.text, 'html.parser')
```

1.2 Get all news and set it's kinds

```python
[3]: # save news kind in a list (we can save it automaticaly but its' not necessary)
     main_titles = ['       ', '          ', '            ', '          ',          ',
     '              ', '        '[

     # gather all news from website
     all_news_stories = home_page_souped.findAll('div', attrs={'class':
      ↪'NewsStories-Inner'})
     print('news count: ',len(all_news_stories))
```

```
news count:  7
```

1.3 Show all news with their requested informations

```python
[5]: # for each kind of news in the website
     for main_count in range(7):

         # get the news kind name from main_title list
         news_kind = main_titles[main_count]
         print('-'*50)
```

```python
    # all news in that kind
    news_stories = all_news_stories[main_count].findAll('div', attrs={'class':
→'NewsStory'})

    # for each news in this kind
    for news_number in range(len(news_stories)):

        this_news = news_stories[news_number]

        # get the needed information
        news_title = this_news.findAll('h2')[0].text
        news_pub_date = this_news.findAll('div', attrs={'class':'PubDate'})[0].
→text
        news_text = this_news.findAll('p')[0].text
        news_link = 'http://uok.ac.ir' + this_news.findAll('a', attrs={'class':
→'More'})[0]['href']

        # print all gathered inforamtion
        print('\n* Title:', news_title)
        print('\n* Kind:', news_kind)
        print('\n* Pub date:', news_pub_date)
        print('\n* News: ', news_text)
        print('\n* Link:', news_link)

        print('-'*50)
```

----

* Title:

                .

* Kind:

* Pub date: 6   1399

* News:

        .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54132
----

* Title:

* Kind:

* Pub date: 5   1399

* News:
                                                   1398

  .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54128
----

* Title:

* Kind:

* Pub date: 2   1399

* News:
         1399/04/30   13:00   15:15
                     LMS   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54124
----

* Title:

* Kind:

* Pub date: 2   1399

* News:
                                    .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54123
----

* Title:

* Kind:

* Pub date: 1   1399

* News: .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54122
----

* Title:

* Kind:

* Pub date: 30   1399

* News:
                              1399/04/25   13:00   15:00
                  Google Meet   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54120
----

* Title:

* Kind:

* Pub date: 24   1399

* News:
                                            .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54115
----

* Title: 1399

* Kind:

* Pub date: 23   1399

* News:
                  1399   (   )   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54112
----

* Title: :

* Kind:

* Pub date: 21   1399

* News:
                              1399/04/18   12:30
13:50   Google Meet   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54106
----

* Title:

* Kind:

* Pub date: 21   1399

* News:

        17/04/1399   11:00   12:30
Google Meet   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54105

----

* Title:

* Kind:

* Pub date: 2   1399

* News:

                                     .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54063
----

* Title:

* Kind:

* Pub date: 31   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54057
----

* Title:
       1400-1399

* Kind:

* Pub date: 29   1399

* News:
       1399-1400

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54048
----

* Title:

* Kind:

* Pub date: 19   1399

* News:
                                99

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54020
----

* Title:

* Kind:

* Pub date: 18   1399

* News:)
      (99/3/20   99/4/04   .
                                        .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54014

----

* Title:

* Kind:

* Pub date: 2   1399

* News:
                                        .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54063
----

* Title:
      1400-1399

* Kind:

* Pub date: 29   1399

* News:
       1399-1400

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54048
----

* Title:

* Kind:

* Pub date: 21   1399

* News: 1399/1400 (   )

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54023
----

* Title:)
1398   (

* Kind:

* Pub date: 20   1399

* News:
                    99/03/20   99/04/20
                                 1398
                          .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54022
----

* Title:

* Kind:

* Pub date: 18   1399

* News:)
     (99/3/20   99/4/04   .
                          .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54014

----

* Title:)   (

* Kind:

* Pub date: 5   1399

* News:)   (
        .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54127
----

* Title: 99

* Kind:

* Pub date: 5   1399

* News: 99
   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54126
----

* Title:

* Kind:

* Pub date: 22   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54110
----

* Title:

* Kind:

* Pub date: 21   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54109
----

* Title: ( )   (9)   1399

* Kind:

* Pub date: 15   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54103

----

* Title:

* Kind:

* Pub date: 19   1399

* News:
                                            .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54021
----

* Title:

* Kind:

* Pub date: 21   1398

* News:

  .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53267
----

* Title:

* Kind:

* Pub date: 16   1398

* News: .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53239
----

* Title:

* Kind:

* Pub date: 27   1398

* News:

                                      .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=52888
----

* Title:

* Kind:

* Pub date: 12   1398

* News:
   .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=52863

----

* Title: 300

* Kind:

* Pub date: 4   1399

* News:
          300-251

     .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54071
----

* Title:

* Kind:

* Pub date: 30   1399

* News:
         .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54047
----

* Title: 250 201-

* Kind:

* Pub date: 15   1399

* News:
                                              250-201
          .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54008
----

* Title:
) (1399

* Kind:

* Pub date: 26   1399

* News:

                           .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53897
----

* Title:

* Kind:

* Pub date: 20   1399

* News:
                           .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53887
----

* Title: "   "

* Kind:

* Pub date: 3   1398

* News: "   "
          –
          .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53775
----

* Title:

* Kind:

* Pub date: 29   1398

* News: –

                    .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53771
----

* Title:

* Kind:

* Pub date: 10   1398

* News: 8   1398
                              –

                    .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53563
----

* Title:

* Kind:

* Pub date: 19   1398

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53441
----

* Title:

* Kind:

* Pub date: 17   1398

* News: 9   98

                                              _
                        .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53420

----

* Title: Benchmarking

* Kind:

* Pub date: 5   1399

* News: Benchmarking
                              .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54130
----

* Title:

* Kind:

* Pub date: 4   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54068
----

* Title:

* Kind:

* Pub date: 14   1399

* News:

                    .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=54005
----

* Title:

* Kind:

* Pub date: 3   1399

* News:

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53973
----

* Title:

* Kind:

* Pub date: 3   1399

* News: .

* Link: http://uok.ac.ir/fa/news/story.aspx?NewsID=53970