

Problem 4

```
import scrapy

class KamernetSpider(Scrapy.Spider):
    name = 'Kamernet_Rooms'

    init_url = ['https://kamernet.nl/huren/kamers-nederland']

    custom_settings = { 'FEED_URI': "Kamernet_%(time)s.json",
                        'FEED_FORMAT': 'json' }

    def parse(self, response):

        # Extracting data using css selectors
        room_link = response.css("div.tile-img a::attr(href)").extract()

        city_name = response.css("div.tile-city::text").extract()

        rent = response.css("div.tile-rent::text").extract()

        room_type = response.css("div.tile-room-type::text").extract()

        surface = response.css("div.tile-surface::text").extract()

        furnished = response.css("div.tile-furnished::text").extract()

        fa_calendar = response.css("div.left::text").extract()

        row_data = zip(city_name,room_link,rent,room_type,
                       surface,furnished,fa_calendar)

        # Making extracted data row wise
        for item in row_data:
            # create a dictionary to store the scraped info
            scraped_info = {
                'city_name' : item[0],
                'room_link' : item[1],
                'rent' : item[2],
                'room_type' : item[3],
                'surface' : item[4],
                'furnished' : item[5],
                'fa_calendar' : item[6]
            }

            # yield or give the scraped info to Scrapy
            yield scraped_info
```

```

NEXT_PAGE_SELECTOR = '.pagination + a::attr(href)'

# you first extracted the link of the next page using
next_page = response.css(NEXT_PAGE_SELECTOR).extract_first()

# if the variable next_page gets a link and is not empty, it will
# enter the if body.
if next_page:
    yield scrapy.Request(
        # The parse() method will use this method to build a new
        # url and provide a new request, which will be sent later
        # to the callback.

        response.urljoin(next_page),
        callback = self.parse
    )

# After receiving the new URL, it will scrape that link executing
# the for body and again look for the next page.
# This will continue until it doesn't get a next page link.

```

Now you can run your spider

Scrapy crawl Kamernet_Rooms

Exporting data

You will need data to be presented as a CSV or JSON so that you can further use the data for analysis. This section of the tutorial will take you through how you can save CSV and JSON file for this data.

To save a CSV file, open settings.py from the project directory and add the following lines:

```

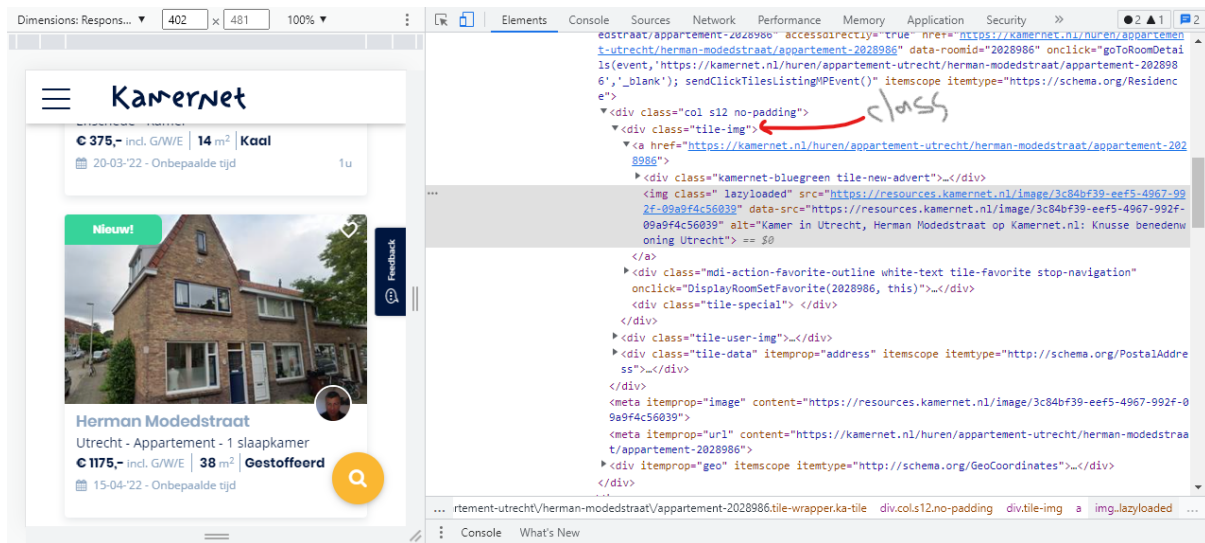
FEED_FORMAT="csv"
FEED_URI="kamernet.csv"

```

After saving the **settings.py**, rerun the **scrapy crawl kamernet_rooms** in your project directory.

For instance – Links to the rooms

As you can see below, name of the class that holds the links is **tile-img**.



`response.css("div.tile-img a::attr(href)").extract()`

`['https://kamernet.nl/huren/kamer-leeuwarden/nieuwesteeg/kamer-2029029',
'https://kamernet.nl/huren/kamer-delft/prof.-evertslaen/kamer-2029031',
'https://kamernet.nl/huren/studio-delft/jan-de-oudeweg/studio-2029033',
'https://kamernet.nl/huren/appartement-amsterdam/wingerdweg/appartement-2029023',
'https://kamernet.nl/huren/kamer-hengelo/sparweg/kamer-2029021',
'https://kamernet.nl/huren/kamer-enschede/kuipersdijk/kamer-2029019',
'https://kamernet.nl/huren/appartement-amsterdam/tweede-marnixplantsoen/appartement-2029017',
'https://kamernet.nl/huren/kamer-amsterdam/martin-ennalsplein/kamer-2028924',
'https://kamernet.nl/huren/studio-hilversum/langestraat/studio-2028946',
'https://kamernet.nl/huren/kamer-amsterdam/ceintuurbaan/kamer-2028760',
'https://kamernet.nl/huren/studio-amsterdam/alhambraan/studio-2028038',
'https://kamernet.nl/huren/kamer-amsterdam/nieuwe-nieuwstraat/kamer-2027894',
'https://kamernet.nl/huren/kamer-enschede/noorderhagen/kamer-1981767',
'https://kamernet.nl/huren/kamer-almere/presidentstraat/kamer-1989403',
'https://kamernet.nl/huren/kamer-amsterdam/jisperveldstraat/kamer-1909443',
'https://kamernet.nl/huren/kamer-amsterdam/badhuiskade/kamer-1899434',
'https://kamernet.nl/huren/kamer-arnhem/rosendaalsestraat/kamer-1893144',
'https://kamernet.nl/huren/kamer-amsterdam/valkhof/kamer-1899247',
'https://kamernet.nl/huren/kamer-utrecht/m.p.-lindostraat/kamer-2029025',
'https://kamernet.nl/huren/studio-groningen/gelkingestraat/studio-2029012',
'https://kamernet.nl/huren/kamer-delft/hugo-de-grootstraat/kamer-2028950',
'https://kamernet.nl/huren/kamer-den-haag/hendrik-zwaardcroonstraat/kamer-2029016',
'https://kamernet.nl/huren/kamer-delft/bagijnhof/kamer-2029015',
'https://kamernet.nl/huren/studio-groningen/lissabonstraat/studio-2029020',
'https://kamernet.nl/huren/kamer-den-haag/copernicuslaan/kamer-2029018',`

'https://kamernet.nl/huren/kamer-alkmaar/bergerweg/kamer-2029014',
'https://kamernet.nl/huren/kamer-groningen/petrus-campersingel/kamer-2029009',
'https://kamernet.nl/huren/kamer-delft/oude-delft/kamer-2029005',
'https://kamernet.nl/huren/kamer-enschede/oldenzaalsestraat/kamer-2029007',
'https://kamernet.nl/huren/kamer-utrecht/brederopplein/kamer-2029006']

.
.
.

