

Московский государственный технический университет им. Н. Э. Баумана  
(МГТУ им. Н. Э. Баумана)

Кафедра «Системы обработки информации и управления» (ИУ5)

Домашнее задание

По дисциплине: «Парадигмы и конструкции языков программирования»

«Анализ цен на съем однокомнатных квартир в Москве»

Выполнила: Холодова К. А.,  
студентка группы ИУ5Ц-52Б

Проверил: Гапанюк Ю. Е.

г. Москва 2023 г.

**Задание:** создать программу для парсинга страницы сайта cian.ru с информацией о съеме однокомнатных квартир в Москве и построения графика цен на них

## Код программы

### Файл *dz.py*

```
import requests
from bs4 import BeautifulSoup as bs
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import re

result_list = {'href': [], 'title': [], 'price': []}

for page_number in range(1, 6):
    url =
    'https://www.cian.ru/cat.php?deal_type=rent&engine_version=2&offer_type=flat&p=
    {}&region=1&room1=1&type=4'.format(str(page_number))
    r = requests.get(url)
    soup = bs(r.text, 'html.parser')
    links = soup.find_all(class_ = '_93444fe79c--container--kZeLu _93444fe79c--
    link--DqDOy')
    titles = soup.find_all(class_=' _93444fe79c--link--VtWj6')
    price = soup.find_all('span', {'data-mark': 'MainPrice'})

    for link in links:
        result_list['href'].append(link.a['href'])

    for title in titles:
        result_list['title'].append(title.text)

    for p in price:
        result_list['price'].append(int((re.search(r'\d+', p.text)[0])+ '000'))

df = pd.DataFrame(result_list)
df.sort_values(by=['price'])
sns.displot(df['price'], kind='kde')
sns.displot(df['price'], kind = 'hist')
plt.show()
```

Результат выполнения:

