

## ИНСТРУКЦИЯ

### Практическое занятие 2.6

Откройте среду разработки Pycharm (или VScode) и создайте новый проект

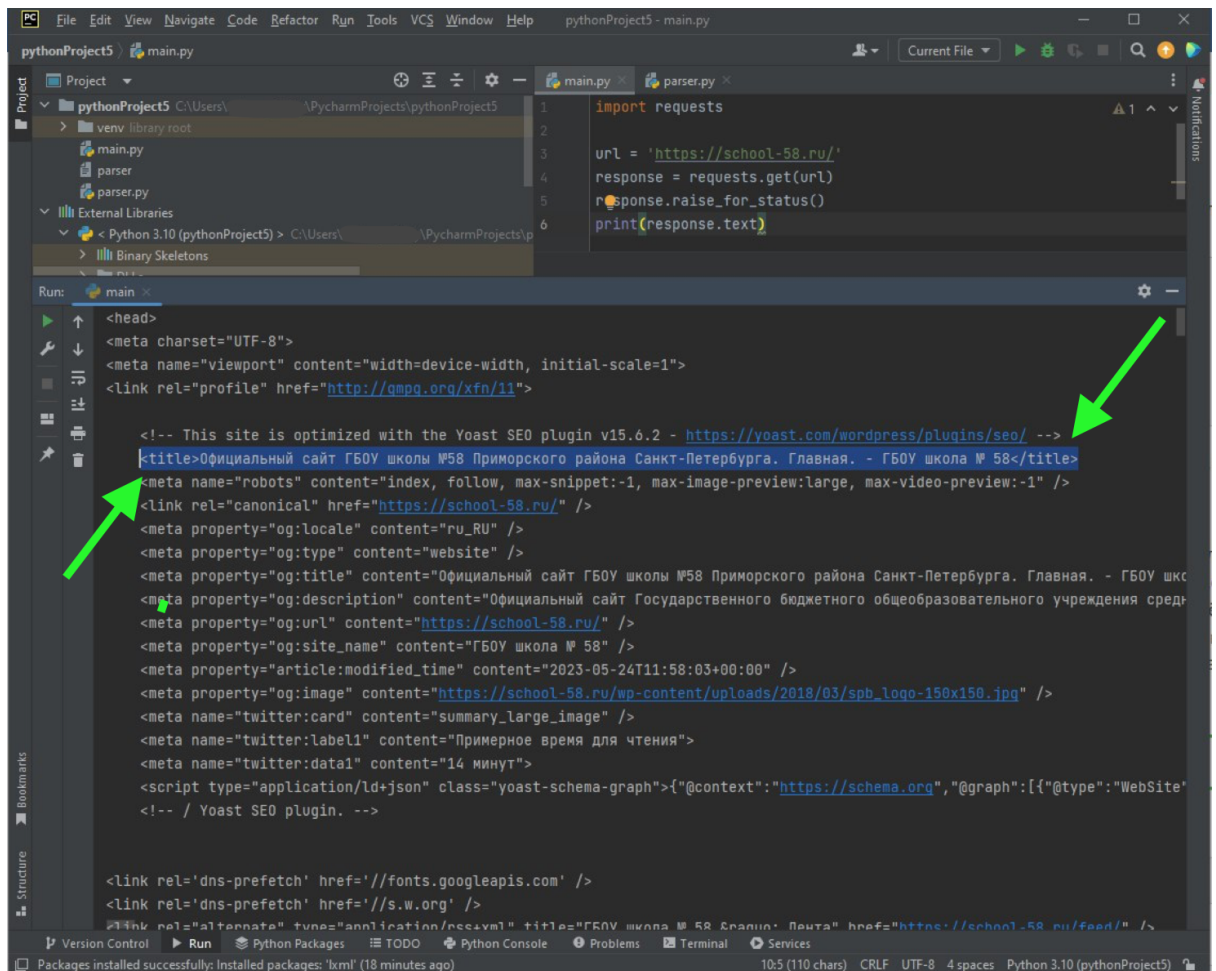
1. Для того, чтобы получить программный код веб-страницы, напишем следующий скрипт:

```
import requests /// импорт библиотеки

url = 'https://сайт_школы.ru' /// переменная адреса сайта
response = requests.get(url) /// направление GET запроса
на сайт
response.raise_for_status() /// проверка статуса запроса

print(response.text) /// вывод программного кода
веб-страницы
```

2. Скрипт выведет весь программный код html страницы. Чаще всего это довольно большой объем текста кода. Но заголовок <title>, который нам нужен обычно находится в начале. Поэтому нужно пролистать вывод программы наверх и найти строку (или строки), где находится этот заголовок. Скопировать его, как указано в примере:



The screenshot shows an IDE window with two tabs: 'main.py' and 'parser.py'. The 'main.py' tab is active, displaying a Python script that uses the 'requests' library to fetch data from 'https://school-58.ru/'. The 'parser.py' tab is also visible, showing a similar script. Below the code editor, the 'Run' window displays the HTML output of the script. The HTML is a meta-data page from a website, featuring various meta tags and a title. Two green arrows point to specific parts of the HTML: one points to the title tag, and the other points to a comment about the Yoast SEO plugin.

```
pythonProject5 - main.py
pythonProject5 \main.py
venv library root
main.py
parser.py
External Libraries
Python 3.10 (pythonProject5)
Binary Skeletons

Run: main
<head>
<meta charset="UTF-8">
<meta name="viewport" content="width=device-width, initial-scale=1">
<link rel="profile" href="http://gmpg.org/xfn/11">

<!-- This site is optimized with the Yoast SEO plugin v15.6.2 - https://yoast.com/wordpress/plugins/seo/ -->
<title>Официальный сайт ГБОУ школы №58 Приморского района Санкт-Петербурга. Главная. - ГБОУ школа № 58</title>
<meta name="robots" content="index, follow, max-snippet:-1, max-image-preview:large, max-video-preview:-1" />
<link rel="canonical" href="https://school-58.ru/" />
<meta property="og:locale" content="ru_RU" />
<meta property="og:type" content="website" />
<meta property="og:title" content="Официальный сайт ГБОУ школы №58 Приморского района Санкт-Петербурга. Главная. - ГБОУ школа № 58" />
<meta property="og:description" content="Официальный сайт Государственного бюджетного общеобразовательного учреждения средней общеобразовательной школы №58 Приморского района Санкт-Петербурга" />
<meta property="og:url" content="https://school-58.ru/" />
<meta property="og:site_name" content="ГБОУ школа № 58" />
<meta property="article:modified_time" content="2023-05-24T11:58:03+00:00" />
<meta property="og:image" content="https://school-58.ru/wp-content/uploads/2018/03/spb_logo-150x150.jpg" />
<meta name="twitter:card" content="summary_large_image" />
<meta name="twitter:label1" content="Примерное время для чтения" />
<meta name="twitter:data1" content="14 минут" />
<script type="application/ld+json" class="yoast-schema-graph">{"@context":"https://schema.org", "@graph":[{"@type":"WebSite"}]}
<!-- / Yoast SEO plugin. -->

<link rel="dns-prefetch" href="//fonts.googleapis.com" />
<link rel="dns-prefetch" href="//s.w.org" />
<link rel="alternate" type="application/rss+xml" title="ГБОУ школа № 58 Справочник: Печата" href="https://school-58.ru/feed/" />

Version Control Run Python Packages TODO Python Console Problems Terminal Services
Packages installed successfully: Installed packages: 'lxml' (18 minutes ago) 10:5 (110 chars) CRLF UTF-8 4 spaces Python 3.10 (pythonProject5)
```

\*Примечание. Если полученный объем текста веб-страницы слишком большой, его можно скопировать в документ и там выполнить поиск по слову title.