Лабораторная работа №5

РАБОТА С ФАЙЛАМИ. РАЗРАБОТКА СИНТАКСИЧЕСКОГО АНАЛИЗАТОРА. ВЫВОД ФОРМАТИРОВАННЫХ ДАННЫХ В ФОРМАТЕ JSON

**Цель работы**: изучить работу с файлами с помощью функций из стандартной библиотеки; рассмотреть понятие синтаксического анализа текста и ознакомиться с его применением на языке Python; изучить возможности взаимодействия Python с форматом хранения данных JSON.

**Задача 19** - Написать функцию, которая принимает путь к HTML файлу и html тег («p», «h1», «article» и др.) и возвращает количество повторений полученного тега в файле с учетом того, что требуется вернуть только количество тегов, который имеет открывающую и закрывающую часть.

**Описание программы**: В коде функция открывает файл и считывает его содержимое, для каждого тега создается отдельная переменная. Функция ищет открывающие и закрывающие теги, увеличивая счетчик, когда находит пару. Указываем отдельно tag - парный тег, его мы и будем выводить

Листинг кода:   
*def count\_complete\_tags(file\_path, tag):  
 with open(file\_path, 'r', encoding='utf-8') as file:  
 content = file.read()  
  
 opening\_tag = f'<{tag}>'  
 closing\_tag = f'</{tag}>'  
  
 count = 0  
 pos = 0  
  
 while True:  
 pos = content.find(opening\_tag, pos)  
 if pos == -1:  
 break  
  
 closing\_pos = content.find(closing\_tag, pos)  
 if closing\_pos == -1:  
 break  
  
 count += 1  
 pos = closing\_pos + len(closing\_tag)  
  
 return count  
  
tag = 'p'  
count = count\_complete\_tags('E:/lython/test/text.html', tag)  
print(f'Количество полных тегов <{tag}>: {count}')*

Результаты выполнения кода:

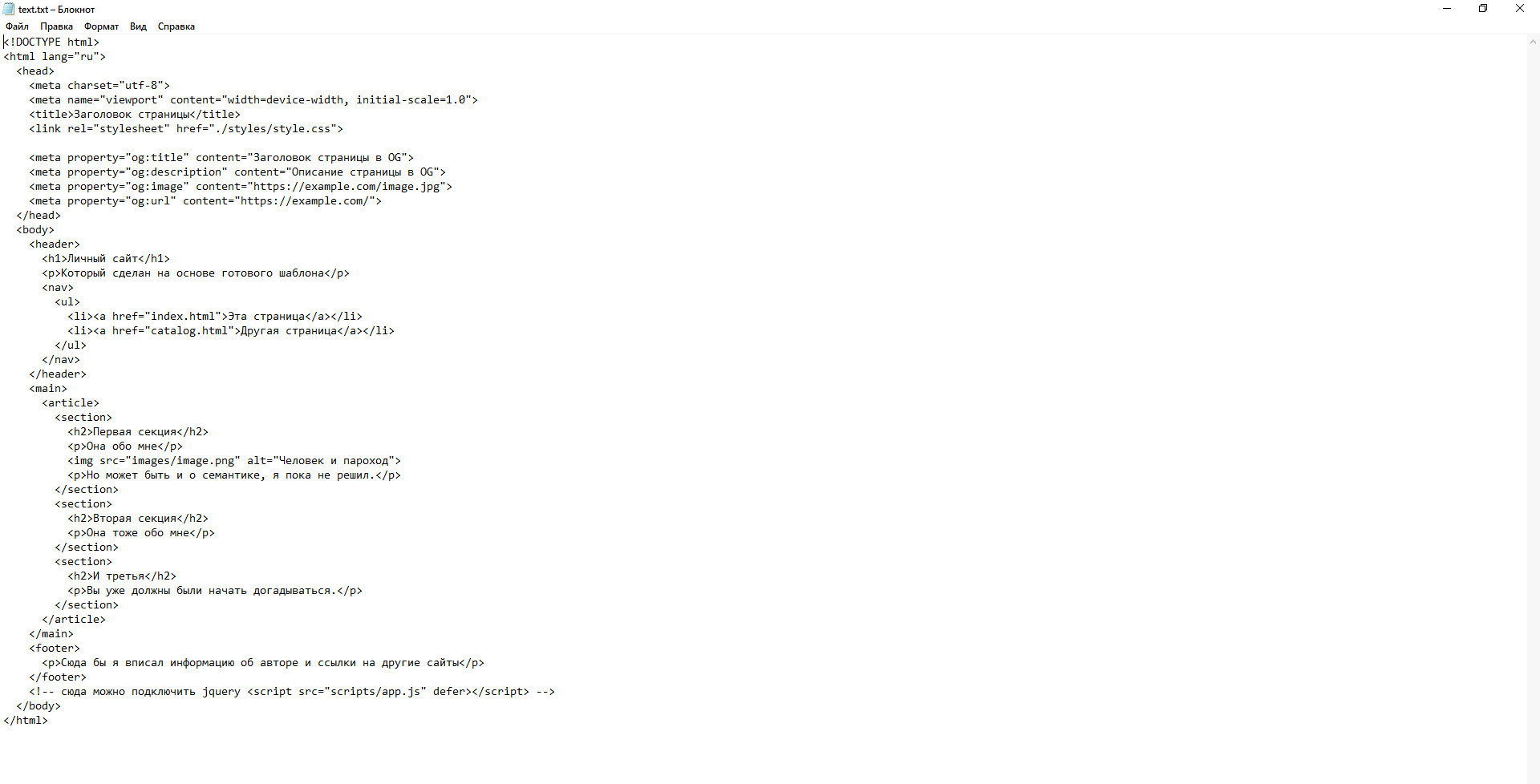
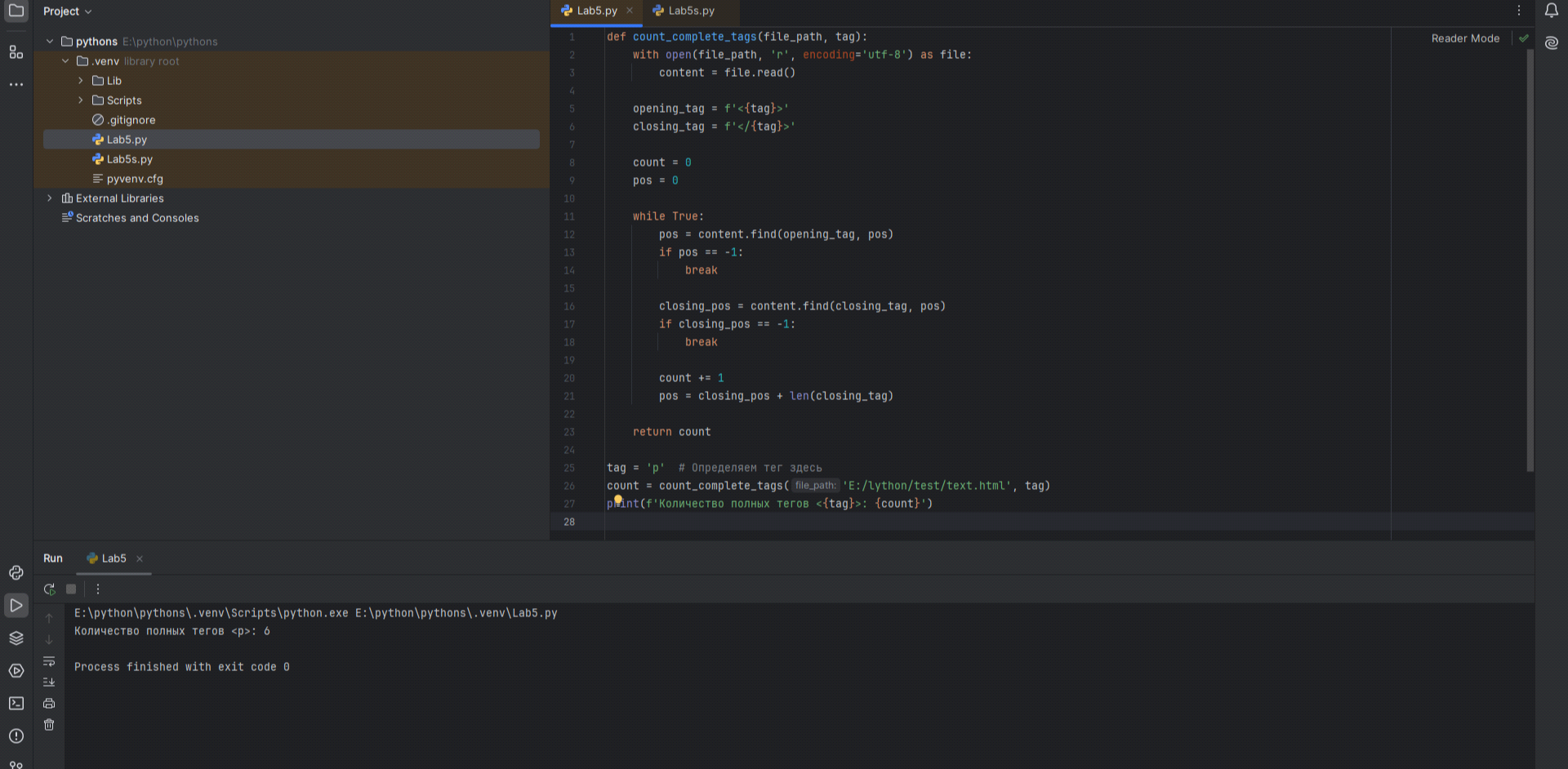


Рис. 1 - HTML файл

  
Рис. 2 - Результат работы