**Урок 15: Работа с библиотекой BeautifulSoup**

**Повторение прошлого материала (10 минут):**

“Прежде чем мы начнем, давайте вспомним, что мы изучали на прошлом уроке. Кто может объяснить, что такое библиотека pygame и для чего она используется?”

“Правильно, pygame — это библиотека для создания игр и мультимедийных приложений на языке Python, которая позволяет работать с графикой, звуком и управлением.”

“Кто может рассказать, как мы создавали графические приложения и рисовали фигуры на экране?”

“Верно, мы создавали окно приложения, очищали экран и рисовали различные фигуры, такие как прямоугольники и круги.”

**Познание нового (25 минут):**

Основные понятия:

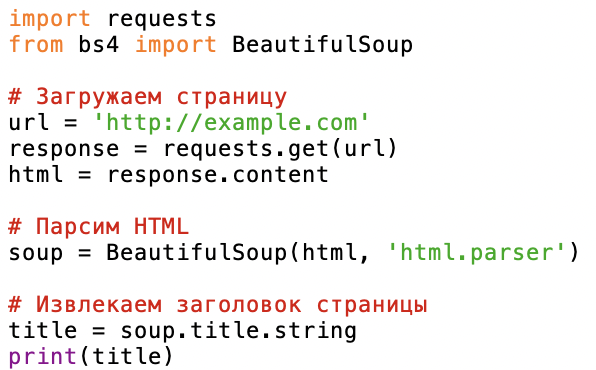
1. Что такое библиотека BeautifulSoup:

“Библиотека BeautifulSoup — это библиотека для парсинга HTML и XML документов. С ее помощью можно извлекать данные из веб-страниц, что полезно для анализа и обработки информации с сайтов.”

2. Установка библиотеки BeautifulSoup:

“Для начала нам нужно установить библиотеку BeautifulSoup4 и requests. Откройте терминал и выполните команду pip install beautifulsoup4 requests.”

3.Основы работы с BeautifulSoup:



“Мы используем библиотеку requests для загрузки HTML страницы, затем парсим ее с помощью BeautifulSoup. После этого мы можем извлекать элементы страницы, такие как заголовок.”

4. Извлечение данных из HTML:



“Мы используем метод find\_all для поиска всех тегов <a> на странице и выводим их атрибуты href, которые содержат ссылки.”

**Некомпьютерная активность (10 минут):**

Обсуждение:

Примеры использования библиотеки BeautifulSoup:

“Библиотеку BeautifulSoup можно использовать для различных задач, таких как сбор данных для анализа, создание веб-скраперов и автоматизация задач.”

Вопросы для размышления:

“Почему важно уметь работать с HTML и парсингом данных? Какие задачи можно решать с помощью библиотеки BeautifulSoup?”

Активность:

Ученики делятся на небольшие группы:

Учитель делит класс на группы по 3-4 человека и дает задание:

“Обсудите в группах, какие данные вы могли бы извлекать из веб-страниц с использованием библиотеки BeautifulSoup. У вас есть 5 минут.”

Каждая группа представляет свои идеи классу:

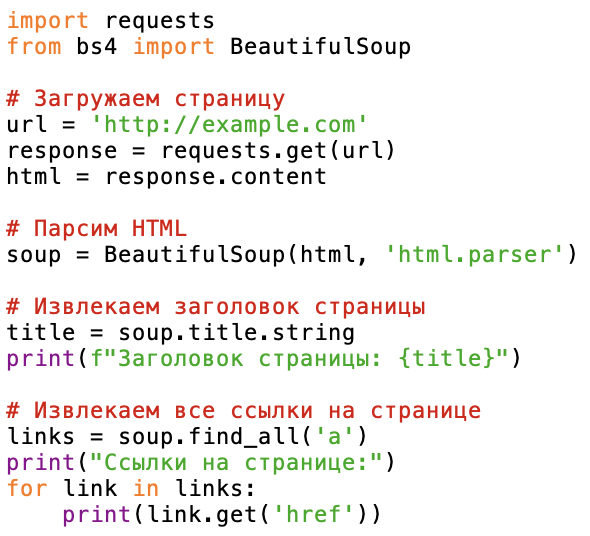
2. Учитель вызывает каждую группу по очереди для представления своих идей. После каждого выступления учитель обсуждает идеи с классом и приводит дополнительные примеры.

**Работа над проектом (25 минут):**

Задание: Создание веб-скрапера

“Давайте создадим веб-скрапер для извлечения данных с веб-страницы. Откройте IDLE, выберите ‘File’ -> ‘New File’.”

Написание кода для загрузки и парсинга страницы:



“Сохраните файл как web\_scraper.py, затем откройте терминал и выполните команду python web\_scraper.py. Вы должны увидеть заголовок страницы и список ссылок.”

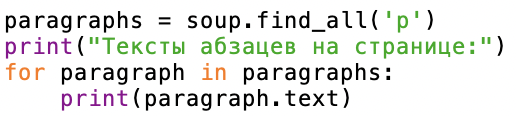
Структура программы:

“Мы создали веб-скрапер, который загружает HTML страницу, парсит ее и извлекает заголовок и ссылки. Это основа для создания более сложных скрапер-приложений.”

**Дополнительное задание:**

Задание: Изменить программу для извлечения других данных

“Теперь давайте изменим нашу программу, чтобы она извлекала текст всех абзацев на странице. Попробуйте добавить код для поиска тегов <p>:”

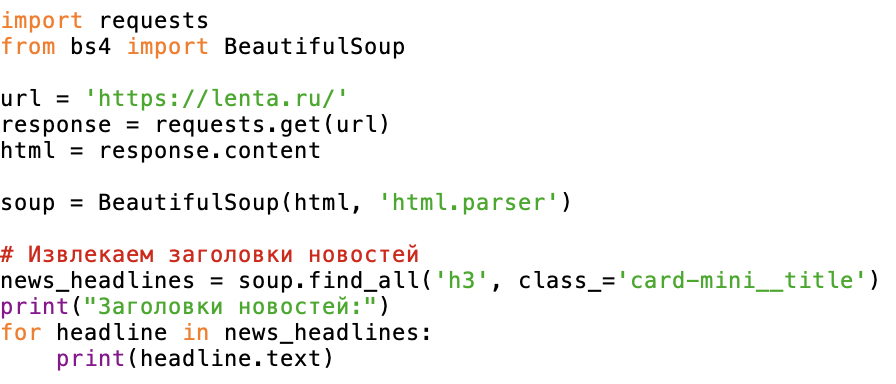


**Проблемная задача (10 минут):**

Задание: Написание программы для извлечения данных

Пример выполнения:

“Теперь давайте создадим программу, которая будет извлекать заголовки новостей с сайта новостей.”



“Запустите программу и посмотрите, как она работает. Какие заголовки новостей она извлекает?”

**Рефлексия (10 минут):**

“Сегодня мы узнали о библиотеке BeautifulSoup и как парсить HTML документы в Python. Кто может объяснить, как загрузить страницу и извлекать данные с помощью BeautifulSoup?”

“Верно, мы использовали requests для загрузки страницы, BeautifulSoup для парсинга HTML и извлечения различных элементов, таких как заголовки, ссылки и тексты абзацев.”