lab7.py - C:/Users/kisho/OneDrive/Desktop/Data Science/Python/lab7.py (3.11.3)

File Edit Format Run Options Window Help

```
"""Write a program that reads a text containing names in the format "First
Last" and uses regular expressions to extract and print all the names
found in the text."""
import re
def extract_names(text):
    name_pattern = r'\b[A-Z][a-zA-Z]*\s[A-Z][a-zA-Z]*\b'
    names = re.findall(name_pattern, text)
    return names
text = "Kishor Patil,Yash Patil, and Arvi are attending the meeting."
names = extract_names(text)
print("Names found:", names)
print()

"""Write a program that reads an HTML program
para  and removes all the HTML tags using regular expressions, leaving
only the plain text content."""
import re
def remove_html_tags(html):
    tag_pattern = r'<.*?>'
    plain_text = re.sub(tag_pattern, '', html)
    return plain_text
html_content = "<html><body><h1>Title</h1><p>I Kishor...!</p></body></html>"
plain_text = remove_html_tags(html_content)
print("Plain text:", plain_text)

print()
"""Write a program that reads a
paragraph of text and uses regular expressions to find and display all
the duplicate words in the paragraph."""
import re
def find_duplicate_words(paragraph):
    word_pattern = r'\b(\w+)\b(?=.*\b\1\b)'
    duplicates = re.findall(word_pattern, paragraph, re.IGNORECASE)
    return set(duplicates)

paragraph = "I am Kishor. I am 1MDS student Kishor."
duplicates = find_duplicate_words(paragraph)
print("Duplicate words found:", duplicates)
print()
```

```python
"""Write a program that reads a
text containing IP addresses and uses regular expressions to extract
and print all the IP addresses found in the text."""
import re
def extract_ip_addresses(text):
    ip_pattern = r'\b(?:\d{1,3}\.){3}\d{1,3}\b'
    ip_addresses = re.findall(ip_pattern, text)
    return ip_addresses
text = "My New Laptop IP address is 192.168.0.1 and the old laptop IP address is 10.0.0.2."
ip_addresses = extract_ip_addresses(text)
print("IP addresses found:", ip_addresses)
print()
""" Write a program
that reads a social media post and uses regular expressions to extract
and print all the hashtags (words starting with a #) found in the post."""

import re
def extract_hashtags(post):
    hashtag_pattern = r'#\w+'
    hashtags = re.findall(hashtag_pattern, post)
    return hashtags
post = "I am #Kishor. I am #1MDS #Student"
hashtags = extract_hashtags(post)
print("Hashtags found:", hashtags)
```

File Edit Shell Debug Options Window Help

```
Python 3.11.3 (tags/v3.11.3:f3909b8, Apr  4 2023, 23:49:59) [MSC v.1934 64 bit (
AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: C:/Users/kisho/OneDrive/Desktop/Data Science/Python/lab7.py =====
Names found: ['Kishor Patil', 'Yash Patil']

Plain text: TitleI Kishor...!

Duplicate words found: {'Kishor', 'am', 'I'}

IP addresses found: ['192.168.0.1', '10.0.0.2']

Hashtags found: ['#Kishor', '#1MDS', '#Student']
>>>
```