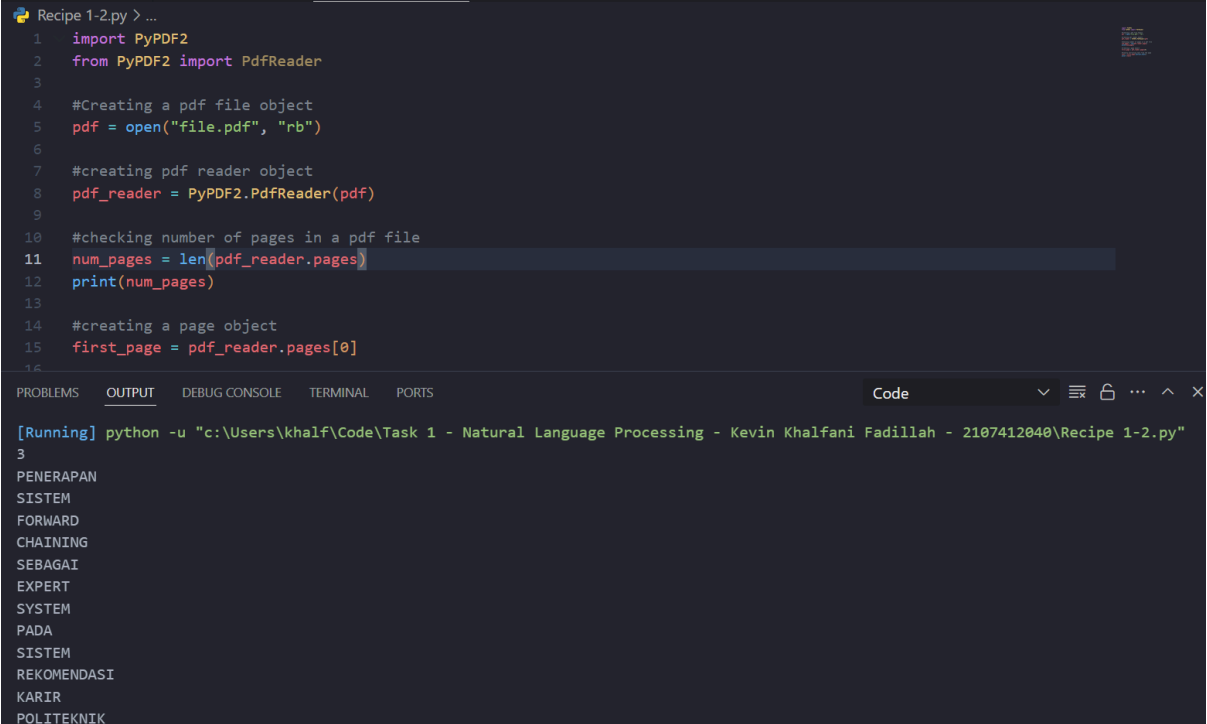


Jawaban

1. Bisa dilihat pada gambar berikut yang merupakan hasil keluaran dari Module Recipe 1-2 tentang pdf reader



The image shows a screenshot of a code editor with a Python script named 'Recipe 1-2.py'. The script uses the PyPDF2 library to open a PDF file, create a PdfReader object, check the number of pages, and create a page object. The output window shows the command 'python -u "c:\Users\khalif\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-2.py"' being executed, resulting in the output '3'. Below the output, there is a list of topics: PENERAPAN, SISTEM, FORWARD, CHAINING, SEBAGAI, EXPERT, SYSTEM, PADA, SISTEM, REKOMENDASI, KARIR, and POLITEKNIK.

```
Recipe 1-2.py > ...
1  import PyPDF2
2  from PyPDF2 import PdfReader
3
4  #Creating a pdf file object
5  pdf = open("file.pdf", "rb")
6
7  #creating pdf reader object
8  pdf_reader = PyPDF2.PdfReader(pdf)
9
10 #checking number of pages in a pdf file
11 num_pages = len(pdf_reader.pages)
12 print(num_pages)
13
14 #creating a page object
15 first_page = pdf_reader.pages[0]
16
```

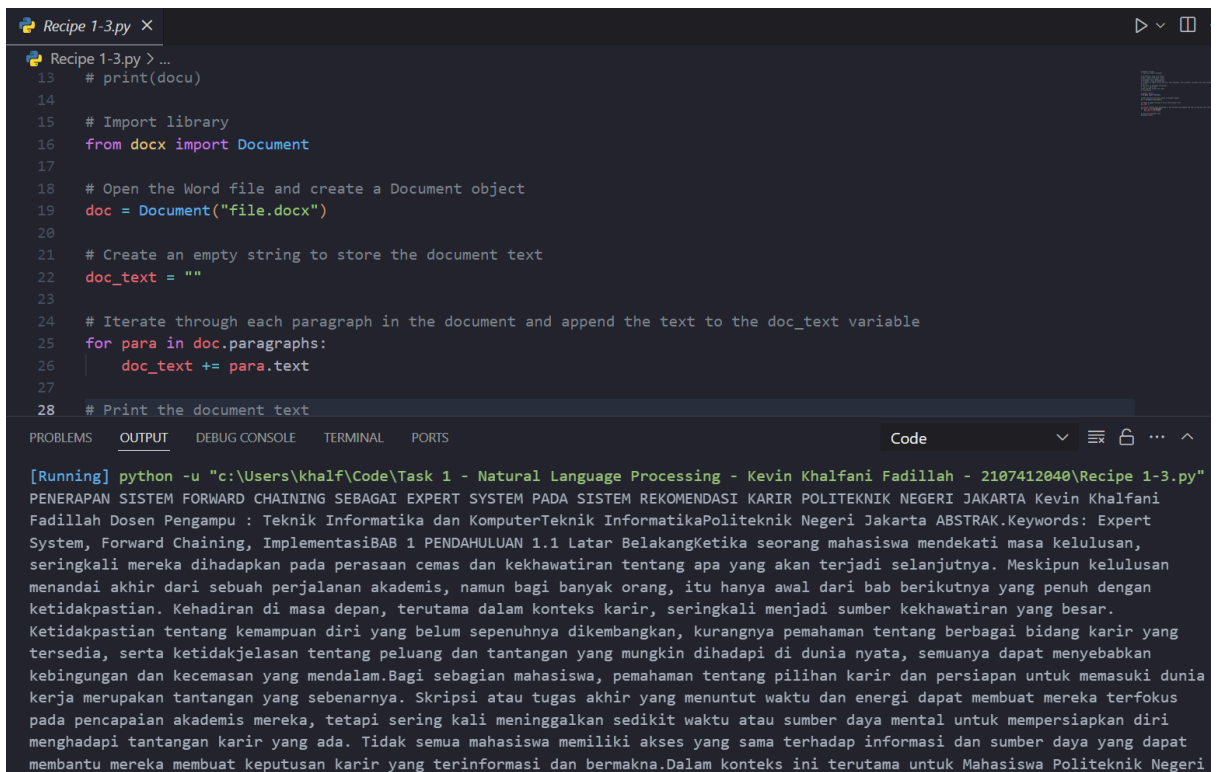
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Code

[Running] python -u "c:\Users\khalif\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-2.py"

3

PENERAPAN
SISTEM
FORWARD
CHAINING
SEBAGAI
EXPERT
SYSTEM
PADA
SISTEM
REKOMENDASI
KARIR
POLITEKNIK

2. Bisa dilihat pada gambar berikut yang merupakan hasil keluaran dari Module Recipe 1-3 Docx reader pada python



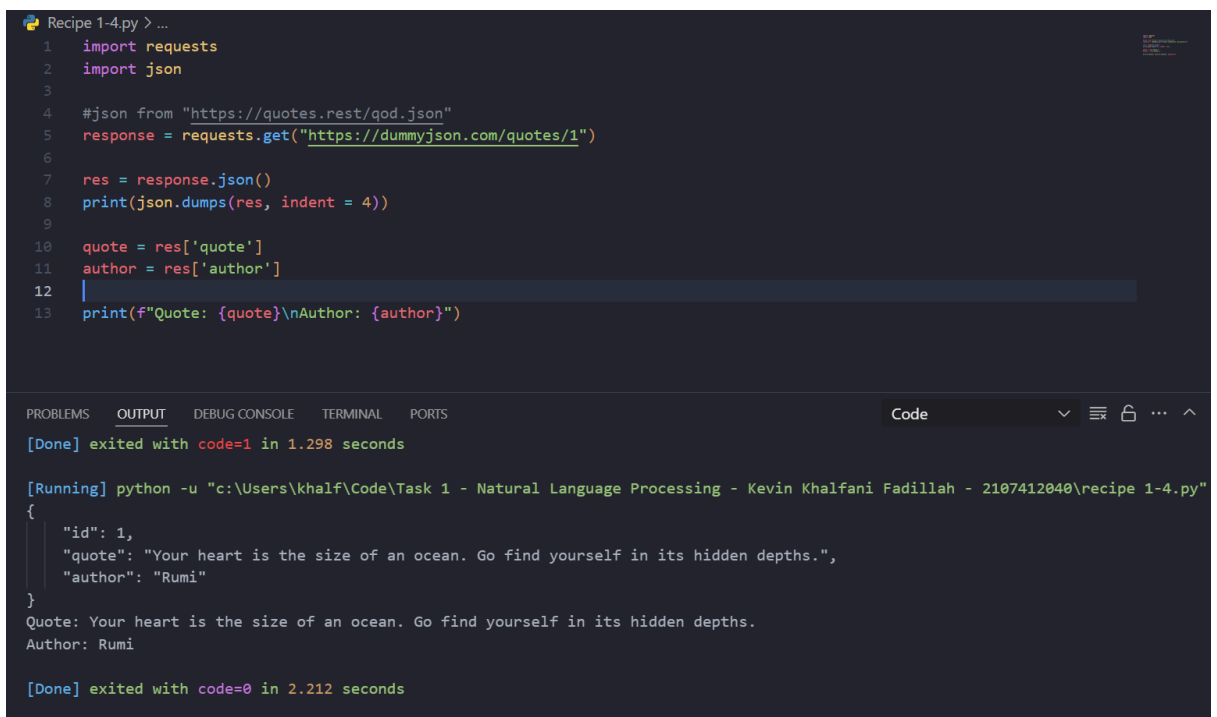
```
Recipe 1-3.py X
Recipe 1-3.py > ...
13 # print(docu)
14
15 # Import library
16 from docx import Document
17
18 # Open the Word file and create a Document object
19 doc = Document("file.docx")
20
21 # Create an empty string to store the document text
22 doc_text = ""
23
24 # Iterate through each paragraph in the document and append the text to the doc_text variable
25 for para in doc.paragraphs:
26     doc_text += para.text
27
28 # Print the document text
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Code

[Running] python -u "c:\Users\khalif\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-3.py"

PENERAPAN SISTEM FORWARD CHAINING SEBAGAI EXPERT SYSTEM PADA SISTEM REKOMENDASI KARIR POLITEKNIK NEGERI JAKARTA Kevin Khalfani Fadillah Dosen Pengampu : Teknik Informatika dan Komputer Teknik Informatika Politeknik Negeri Jakarta ABSTRAK. Keywords: Expert System, Forward Chaining, Implementasi BAB 1 PENDAHULUAN 1.1 Latar Belakang Ketika seorang mahasiswa mendekati masa kelulusan, seringkali mereka dihadapkan pada perasaan cemas dan kekhawatiran tentang apa yang akan terjadi selanjutnya. Meskipun kelulusan menandai akhir dari sebuah perjalanan akademis, namun bagi banyak orang, itu hanya awal dari bab berikutnya yang penuh dengan ketidakpastian. Kehadiran di masa depan, terutama dalam konteks karir, seringkali menjadi sumber kekhawatiran yang besar. Ketidakpastian tentang kemampuan diri yang belum sepenuhnya dikembangkan, kurangnya pemahaman tentang berbagai bidang karir yang tersedia, serta ketidakjelasan tentang peluang dan tantangan yang mungkin dihadapi di dunia nyata, semuanya dapat menyebabkan kebingungan dan kecemasan yang mendalam. Bagi sebagian mahasiswa, pemahaman tentang pilihan karir dan persiapan untuk memasuki dunia kerja merupakan tantangan yang sebenarnya. Skripsi atau tugas akhir yang menuntut waktu dan energi dapat membuat mereka terfokus pada pencapaian akademis mereka, tetapi sering kali meninggalkan sedikit waktu atau sumber daya mental untuk mempersiapkan diri menghadapi tantangan karir yang ada. Tidak semua mahasiswa memiliki akses yang sama terhadap informasi dan sumber daya yang dapat membantu mereka membuat keputusan karir yang terinformasi dan bermakna. Dalam konteks ini terutama untuk Mahasiswa Politeknik Negeri

3. Bisa dilihat pada gambar berikut yang merupakan hasil keluaran dari Module Recipe 1-4 yang mana merupakan sebuah cara untuk mengakses data JSON dan melihat outputnya



```
Recipe 1-4.py > ...
1 import requests
2 import json
3
4 #json from "https://quotes.rest/qod.json"
5 response = requests.get("https://dummyjson.com/quotes/1")
6
7 res = response.json()
8 print(json.dumps(res, indent = 4))
9
10 quote = res['quote']
11 author = res['author']
12
13 print(f"Quote: {quote}\nAuthor: {author}")
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Code

[Done] exited with code=1 in 1.298 seconds

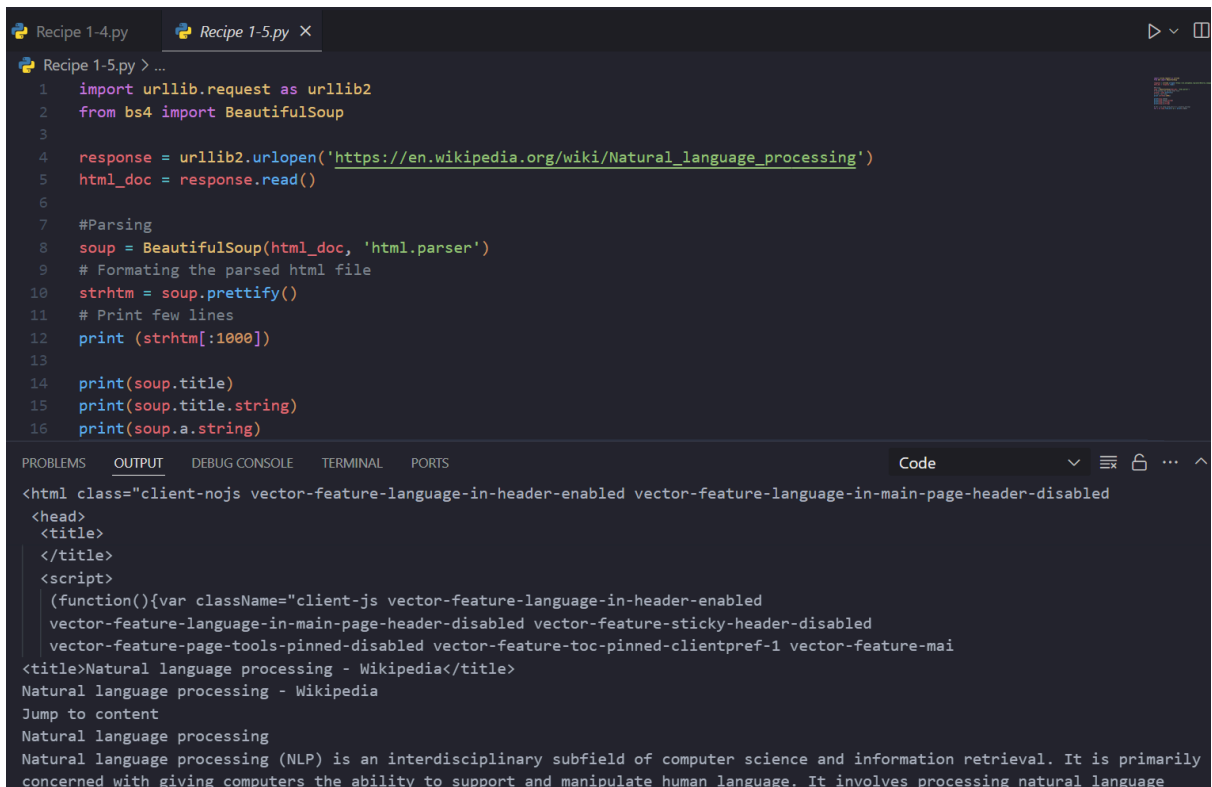
[Running] python -u "c:\Users\khalif\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\recipe 1-4.py"

```
{
  "id": 1,
  "quote": "Your heart is the size of an ocean. Go find yourself in its hidden depths.",
  "author": "Rumi"
}
```

Quote: Your heart is the size of an ocean. Go find yourself in its hidden depths.
Author: Rumi

[Done] exited with code=0 in 2.212 seconds

4. Berikut merupakan output dari Modul Recipe 1-5 yang merupakan cara mengekstrak sebuah tampilan website beserta text di dalamnya

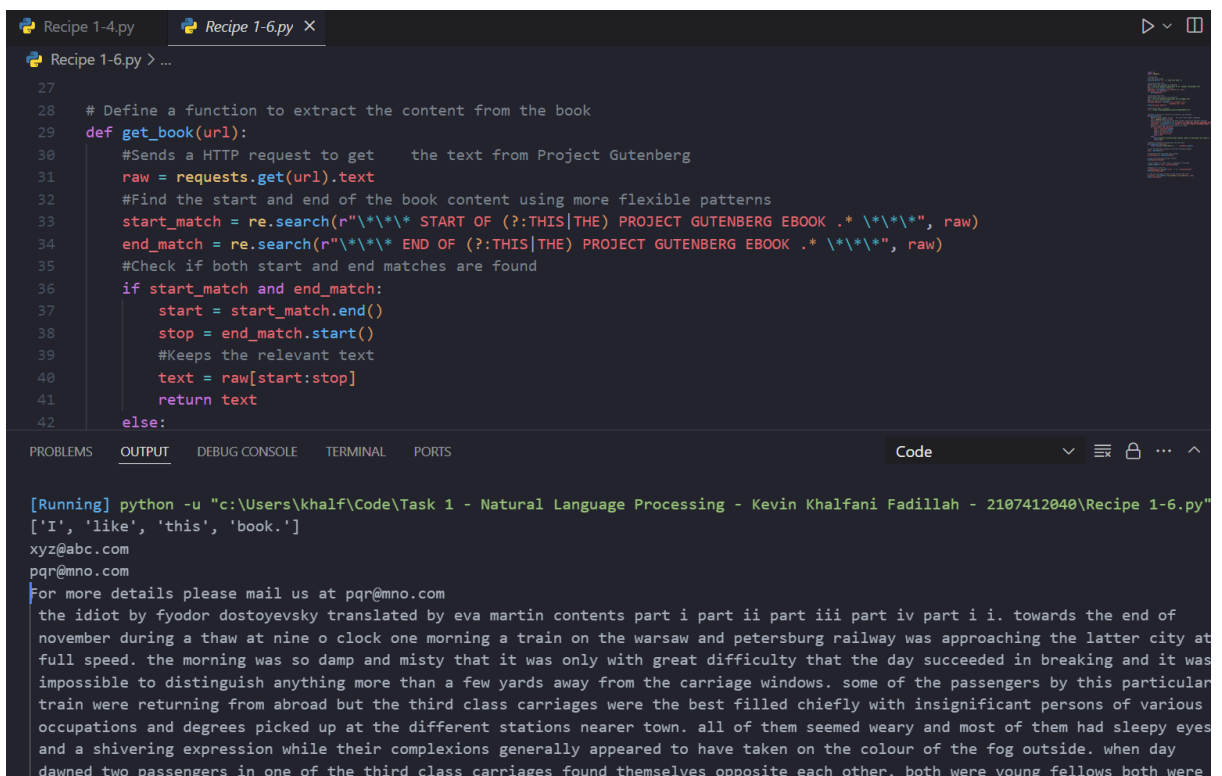


```
Recipe 1-5.py > ...
1  import urllib.request as urllib2
2  from bs4 import BeautifulSoup
3
4  response = urllib2.urlopen('https://en.wikipedia.org/wiki/Natural_language_processing')
5  html_doc = response.read()
6
7  #Parsing
8  soup = BeautifulSoup(html_doc, 'html.parser')
9  # Formatting the parsed html file
10 strhtml = soup.prettify()
11 # Print few lines
12 print (strhtml[:1000])
13
14 print(soup.title)
15 print(soup.title.string)
16 print(soup.a.string)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Code

```
<html class="client-nojs vector-feature-language-in-header-enabled vector-feature-language-in-main-page-header-disabled
<head>
<title>
</title>
<script>
(function(){var className="client-js vector-feature-language-in-header-enabled
vector-feature-language-in-main-page-header-disabled vector-feature-sticky-header-disabled
vector-feature-page-tools-pinned-disabled vector-feature-toc-pinned-clientpref-1 vector-feature-mai
<title>Natural language processing - Wikipedia</title>
Natural language processing - Wikipedia
Jump to content
Natural language processing
Natural language processing (NLP) is an interdisciplinary subfield of computer science and information retrieval. It is primarily
concerned with giving computers the ability to support and manipulate human language. It involves processing natural language
```

5. Berikut merupakan sebuah metode di mana kita bisa mengekstrak sebuah e-book

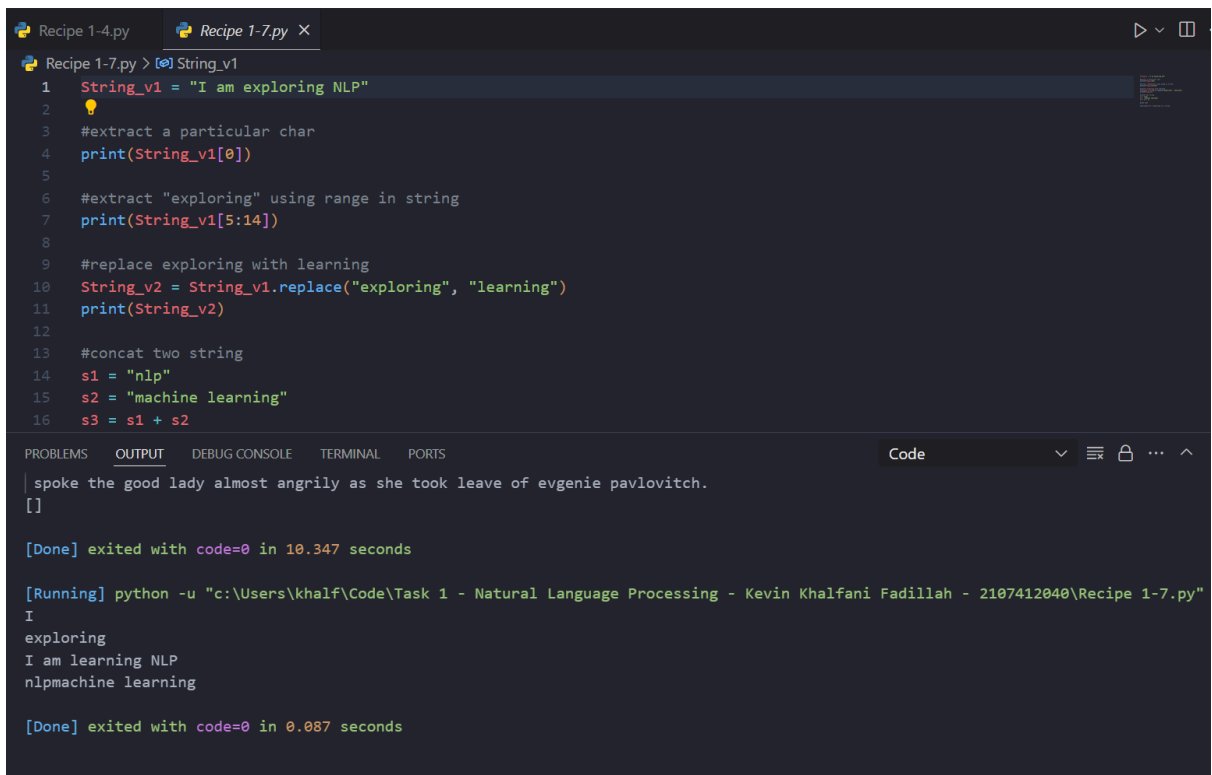


```
Recipe 1-6.py > ...
27
28 # Define a function to extract the content from the book
29 def get_book(url):
30     #Sends a HTTP request to get the text from Project Gutenberg
31     raw = requests.get(url).text
32     #Find the start and end of the book content using more flexible patterns
33     start_match = re.search(r"[*]* START OF (?:THIS|THE) PROJECT GUTENBERG EBOOK .* [*]*", raw)
34     end_match = re.search(r"[*]* END OF (?:THIS|THE) PROJECT GUTENBERG EBOOK .* [*]*", raw)
35     #Check if both start and end matches are found
36     if start_match and end_match:
37         start = start_match.end()
38         stop = end_match.start()
39         #Keeps the relevant text
40         text = raw[start:stop]
41         return text
42     else:
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Code

```
[Running] python -u "c:\Users\khalf\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-6.py"
['I', 'like', 'this', 'book.']
xyz@abc.com
pqr@mno.com
For more details please mail us at pqr@mno.com
the idiot by fyodor dostoyevsky translated by eva martin contents part i part ii part iii part iv part i i. towards the end of
november during a thaw at nine o clock one morning a train on the warsaw and petersburg railway was approaching the latter city at
full speed. the morning was so damp and misty that it was only with great difficulty that the day succeeded in breaking and it was
impossible to distinguish anything more than a few yards away from the carriage windows. some of the passengers by this particular
train were returning from abroad but the third class carriages were the best filled chiefly with insignificant persons of various
occupations and degrees picked up at the different stations nearer town. all of them seemed weary and most of them had sleepy eyes
and a shivering expression while their complexions generally appeared to have taken on the colour of the fog outside. when day
dawned two passengers in one of the third class carriages found themselves opposite each other. both were young fellows both were
```

6. Berikut merupakan sebuah cara untuk memanipulasi string dengan menggabungkan ataupun memisahnya ataupun juga bisa mengganti sebuah string yang sudah ada dengan input yang baru



The screenshot shows a Python IDE with two tabs: 'Recipe 1-4.py' and 'Recipe 1-7.py'. The active tab is 'Recipe 1-7.py', which contains the following Python code:

```
1 String_v1 = "I am exploring NLP"
2
3 #extract a particular char
4 print(String_v1[0])
5
6 #extract "exploring" using range in string
7 print(String_v1[5:14])
8
9 #replace exploring with learning
10 String_v2 = String_v1.replace("exploring", "learning")
11 print(String_v2)
12
13 #concat two string
14 s1 = "nlp"
15 s2 = "machine learning"
16 s3 = s1 + s2
```

The IDE's output pane shows the following output:

```
| spoke the good lady almost angrily as she took leave of evgenie pavlovitch.
[]

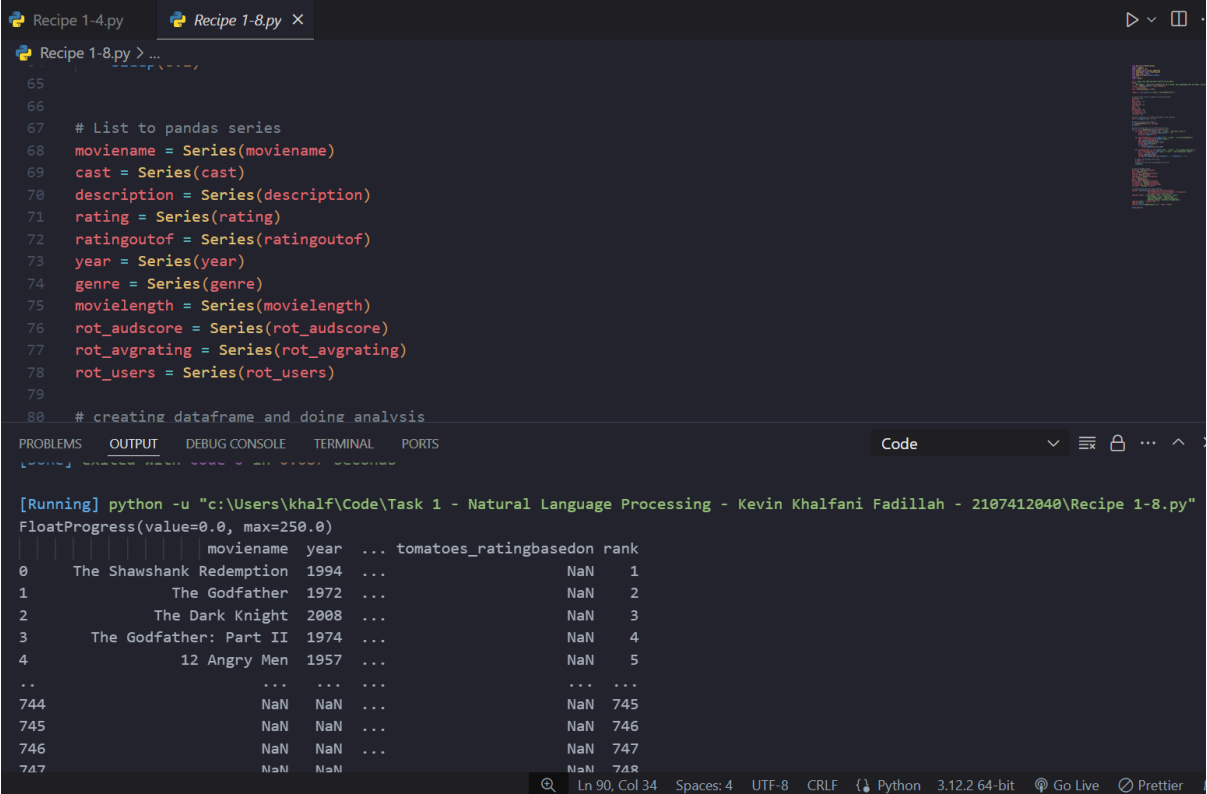
[Done] exited with code=0 in 10.347 seconds

[Running] python -u "c:\Users\khalif\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-7.py"
I
exploring
I am learning NLP
nlpmachine learning

[Done] exited with code=0 in 0.087 seconds
```

7. Pada recipe terakhir di bab 1 ini kita bisa belajar scrapping text dari sebuah web dengan menggunakan berbagai library yang ada di python dan mengubah outputnya menjadi

sebuah file csv



The image shows a VS Code editor with two tabs: 'Recipe 1-4.py' and 'Recipe 1-8.py'. The active tab is 'Recipe 1-8.py', which contains a Python script. The script defines several Series objects for movie data and then creates a DataFrame. The output window shows the execution of the script, displaying a DataFrame with columns: 'moviename', 'year', 'tomatoes_ratingbasedon', and 'rank'. The DataFrame contains 748 rows of data, with the first 5 rows being visible in the output window.

```
65
66
67 # List to pandas series
68 moviename = Series(moviename)
69 cast = Series(cast)
70 description = Series(description)
71 rating = Series(rating)
72 ratingoutof = Series(ratingoutof)
73 year = Series(year)
74 genre = Series(genre)
75 movielength = Series(movielength)
76 rot_audscore = Series(rot_audscore)
77 rot_avgrating = Series(rot_avgrating)
78 rot_users = Series(rot_users)
79
80 # creating dataframe and doing analysis
```

[Running] python -u "c:\Users\khalf\Code\Task 1 - Natural Language Processing - Kevin Khalfani Fadillah - 2107412040\Recipe 1-8.py"

FloatProgress(value=0.0, max=250.0)

	moviename	year	tomatoes_ratingbasedon	rank
0	The Shawshank Redemption	1994	...	NaN 1
1	The Godfather	1972	...	NaN 2
2	The Dark Knight	2008	...	NaN 3
3	The Godfather: Part II	1974	...	NaN 4
4	12 Angry Men	1957	...	NaN 5
...
744	NaN	NaN	...	NaN 745
745	NaN	NaN	...	NaN 746
746	NaN	NaN	...	NaN 747
747	NaN	NaN	...	NaN 748

Ln 90, Col 34 Spaces: 4 UTF-8 CRLF Python 3.12.2 64-bit Go Live Prettier