Cento Universitário UNA

Sistemas de Informação

Recuperação de Informação

Práticas de Laboratório

Wesley Dias Maciel

2019/01

Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01

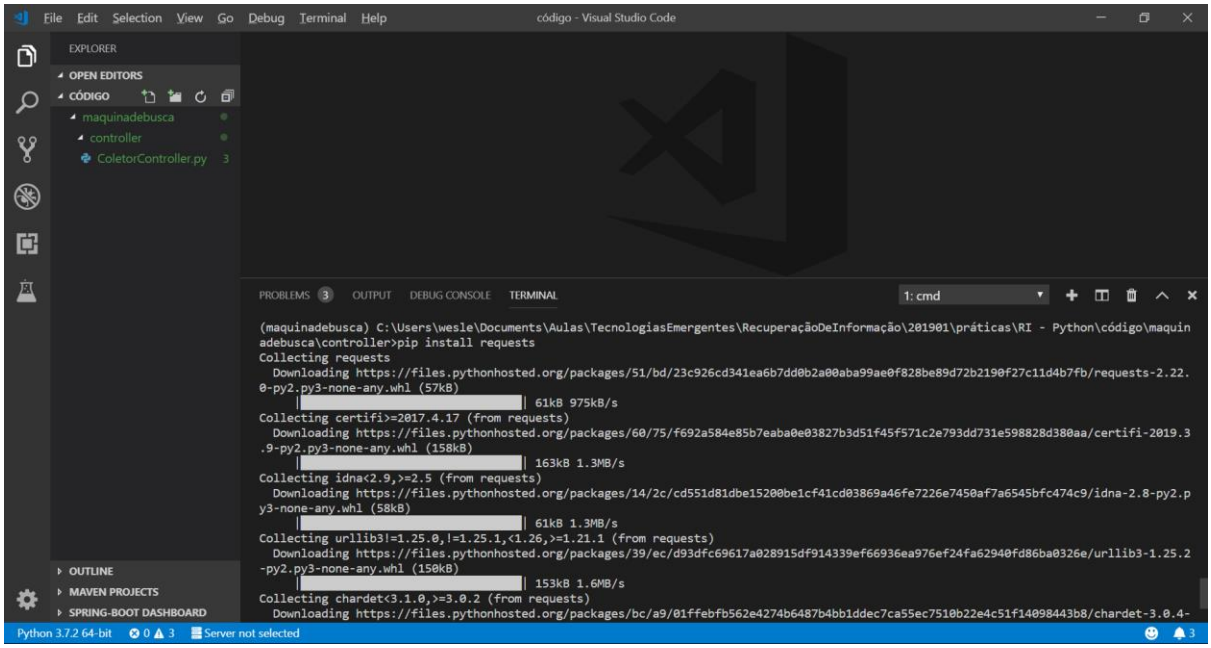
Python

Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01

Prática 07

requests

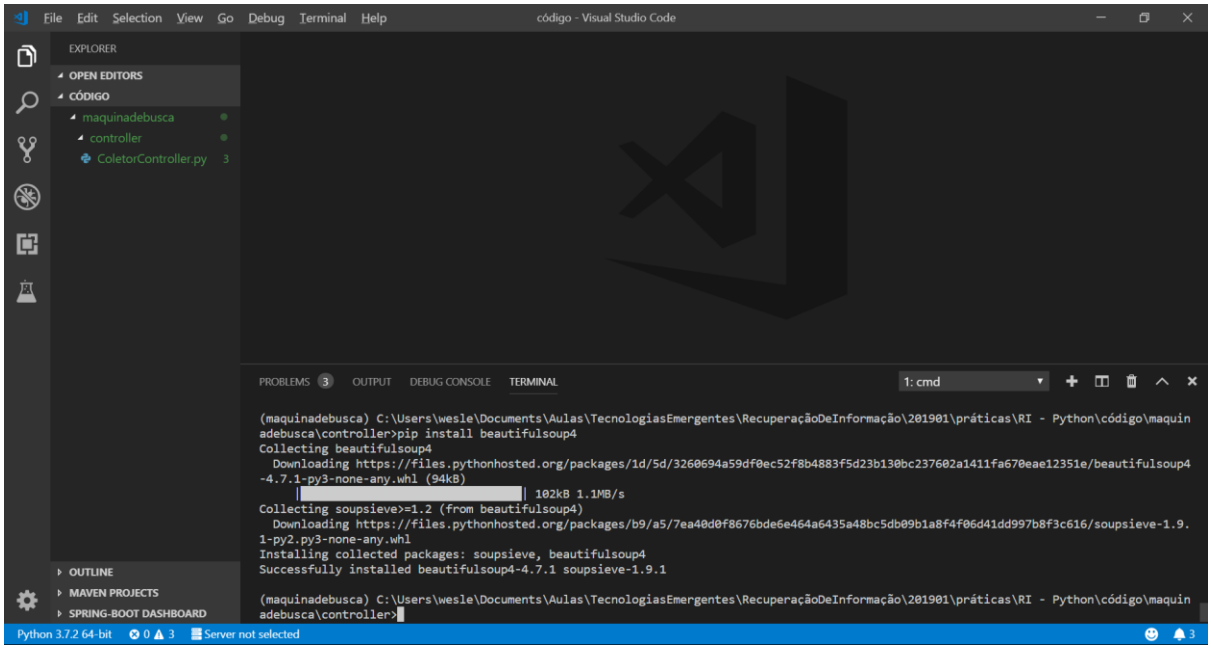
O módulo “requests” permite realizar requisições HTTP. Instale o módulo requests, executando o comando “pip install requests”.



BeautifulSoup

O múdulo “BeautifulSoup” permite tratar arquivos no formato HTML e XML. Instale o módulo BeautifulSoup, executando o comando “pip install beautifulsoup4”.

Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01

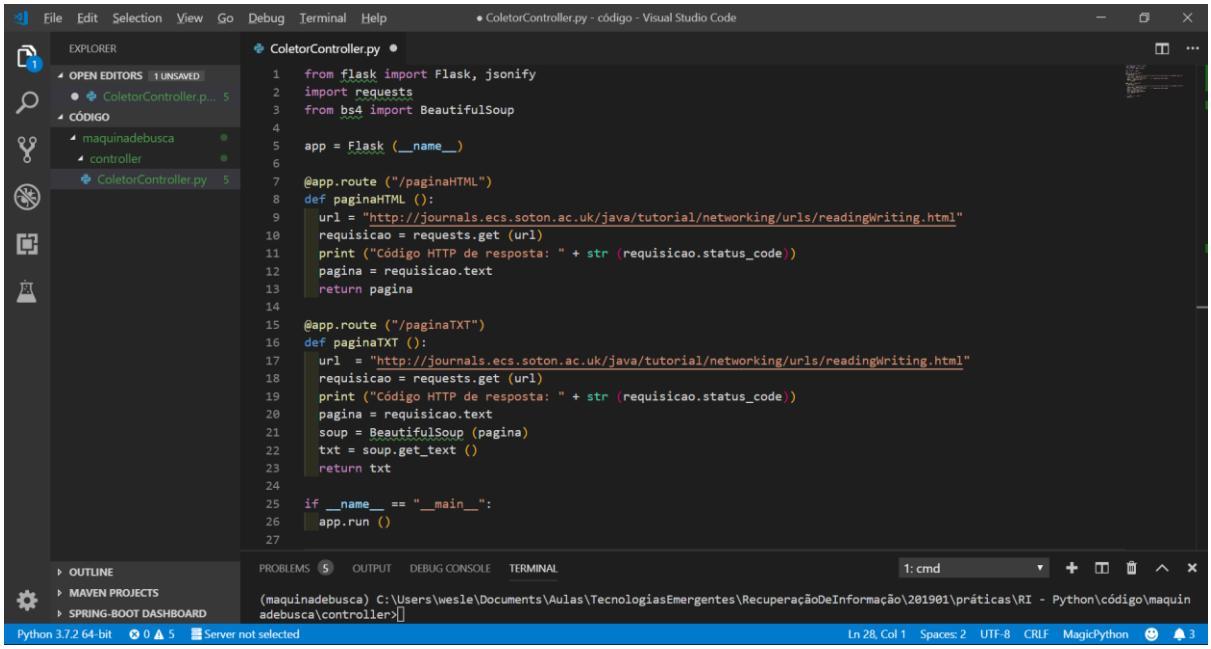


A partir da prática 06, altere o código do coletor, como apresentado abaixo:

```
from flask import Flask, jsonify
import requests
from bs4 import BeautifulSoup

app = Flask (__name__)

@app.route ("/paginaHTML")
def paginaHTML ():
  url =
"http://journals.ecs.soton.ac.uk/java/tutorial/networking/urls/readingWriting.
html"
  requisicao = requests.get (url)
  print ("Código HTTP de resposta: " + str (requisicao.status_code))
  pagina = requisicao.text
  return pagina

@app.route ("/paginaTXT")
def paginaTXT ():
  url  =
"http://journals.ecs.soton.ac.uk/java/tutorial/networking/urls/readingWriting.
html"
  requisicao = requests.get (url)
  print ("Código HTTP de resposta: " + str (requisicao.status_code))
  pagina = requisicao.text
```

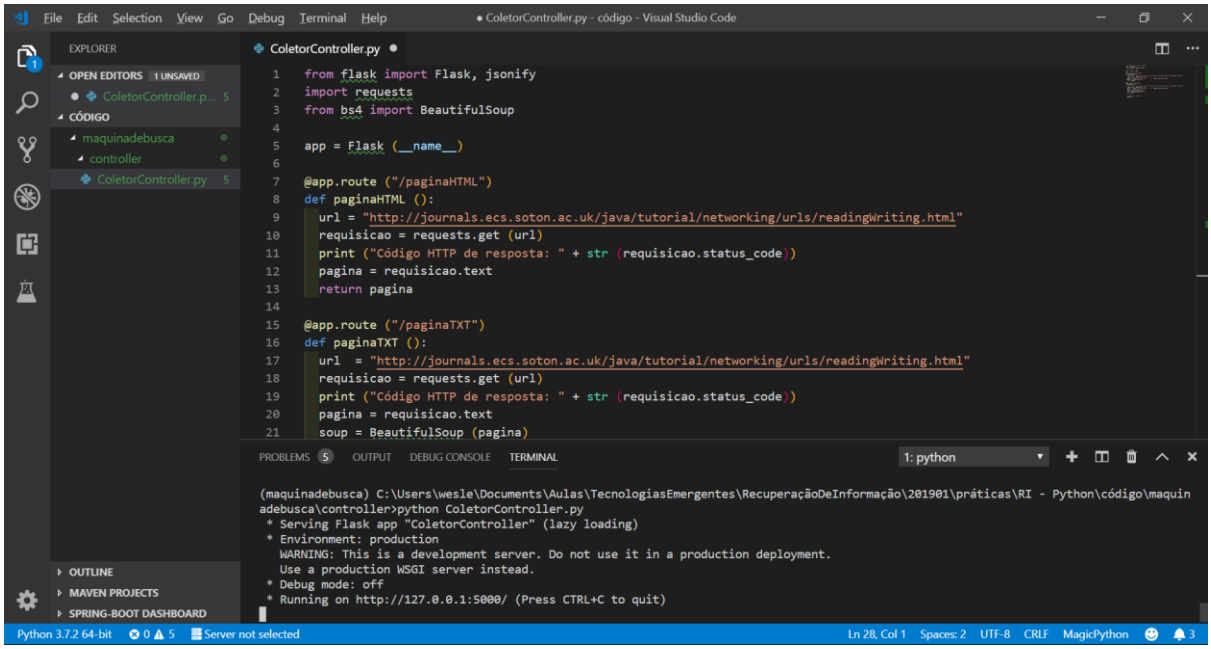
Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01

```
  soup = BeautifulSoup (pagina)
  txt = soup.get_text ()
  return txt

if __name__ == "__main__":
  app.run ()
```

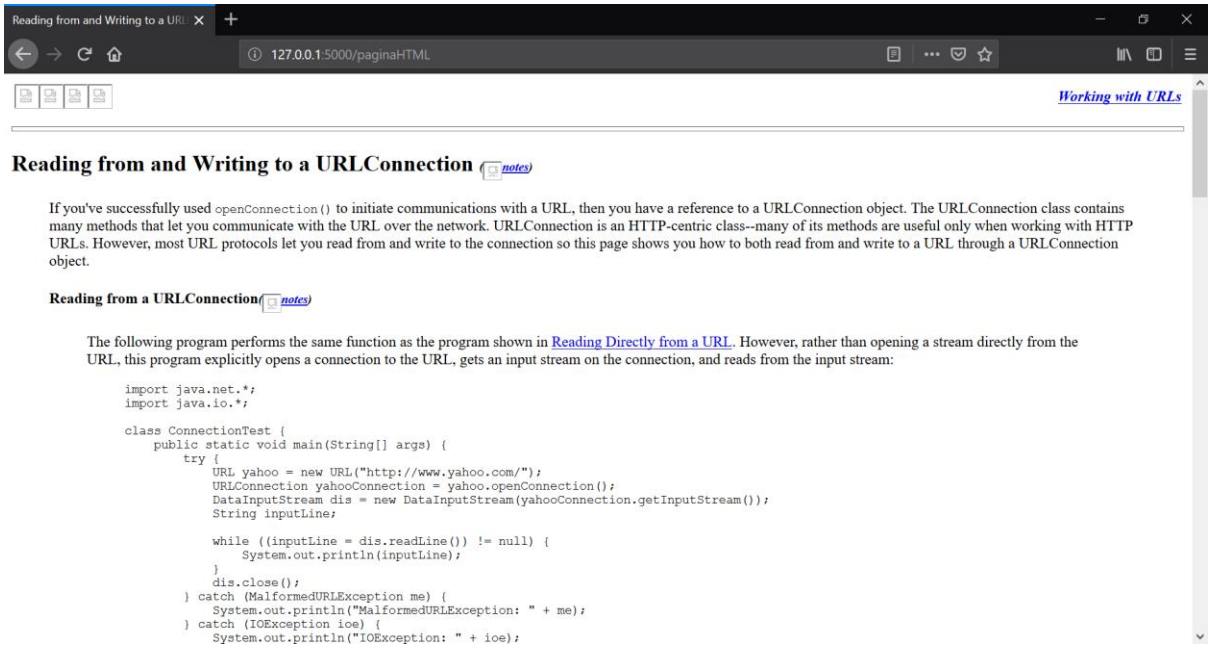


Inicie o servidor, executando o comando “python ColetorController.py”.

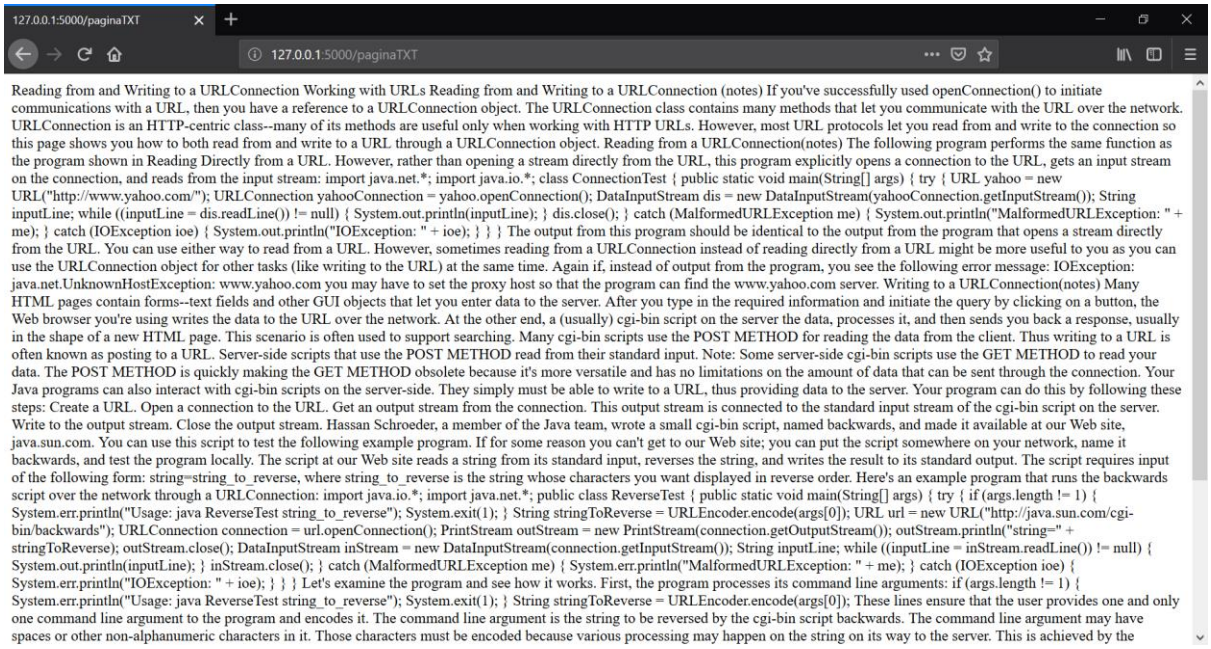
Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01



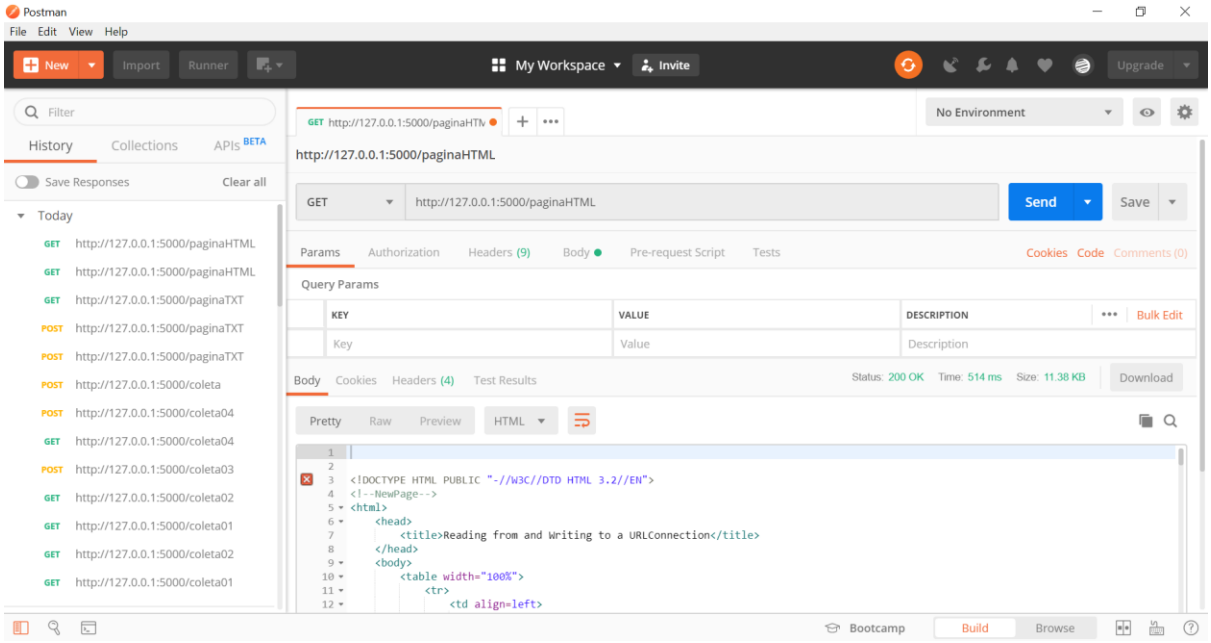
Teste a API nos navegadores Chrome ou FireFox.



Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01



Teste a API também no Postman.



Centro Universitário UNA
Sistemas de Informação
Recuperação de Informação
Prática de Laboratório
Wesley Dias Maciel
2019/01