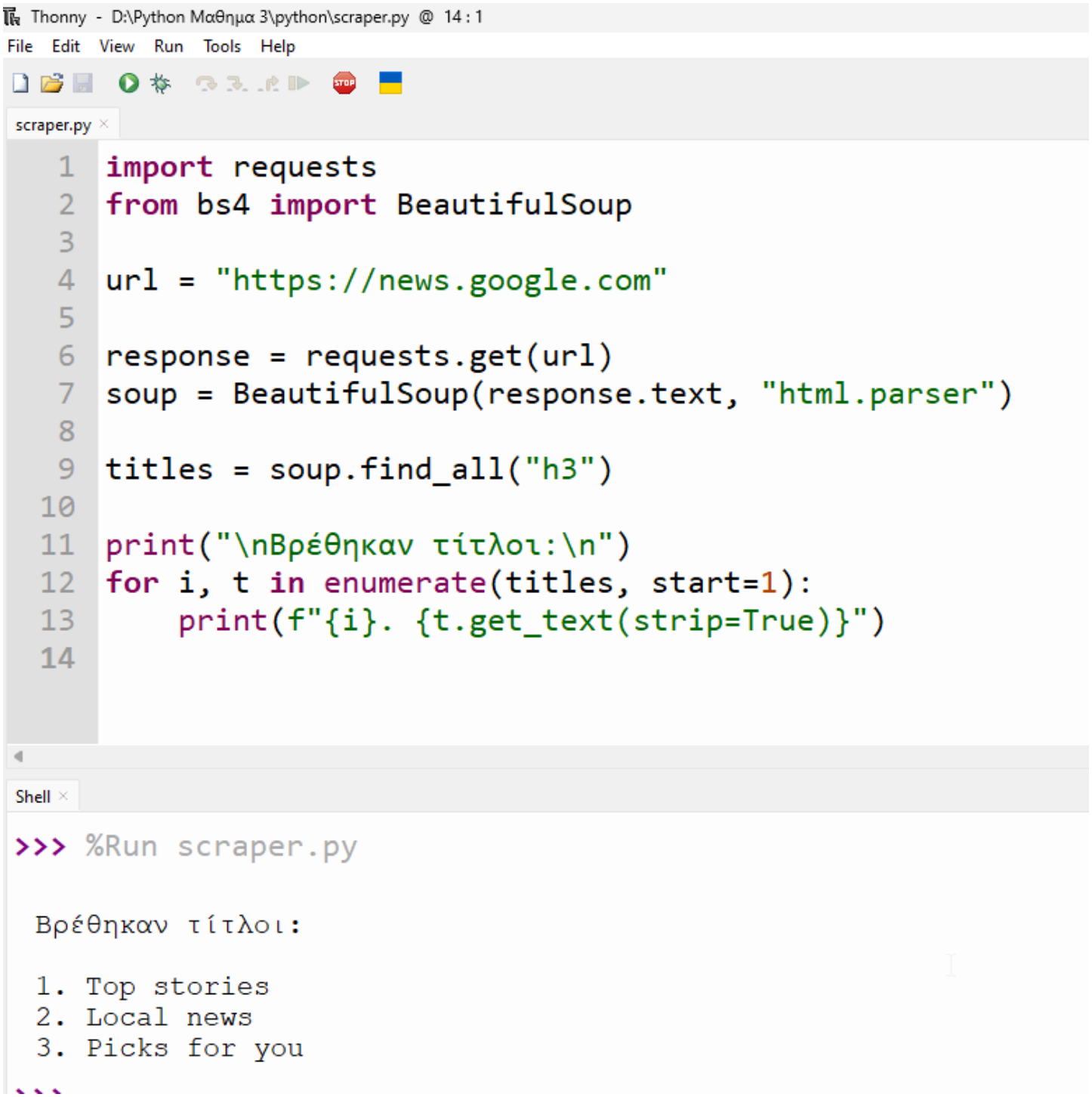
Thonny - D:\Python Μαθημα 3\python\scraper.py @ 14 : 1
File Edit View Run Tools Help
scraper.py
1 import requests
2 from bs4 import BeautifulSoup
3
4 url = "https://news.google.com"
5
6 response = requests.get(url)
7 soup = BeautifulSoup(response.text, "html.parser")
8
9 titles = soup.find_all("h3")
10
11 print("\nΒρέθηκαν τίτλοι:\n")
12 for i, t in enumerate(titles, start=1):
13 print(f"{i}. {t.get_text(strip=True)}")
14
Shell
>>> %Run scraper.py
Βρέθηκαν τίτλοι:
1. Top stories
2. Local news
3. Picks for you
>>>

Web Image Downloade.py

```python
import os
import requests
from bs4 import BeautifulSoup
from urllib.parse import urljoin

url = "https://www.python.org"   # άλλαξέ το σε όποιο site θέλεις

r = requests.get(url)
soup = BeautifulSoup(r.text, "html.parser")

os.makedirs("images", exist_ok=True)

for img in soup.find_all("img"):
    src = img.get("src")
    if not src:
        continue
    img_url = urljoin(url, src)
    filename = img_url.split("/")[-1]

    try:
        data = requests.get(img_url).content
        with open(os.path.join("images", filename), "wb") as f:
            f.write(data)
        print("✔ Κατέβηκε:", filename)
    except:
        pass
```

Shell

```
>>> %Run 'Web Image Downloade.py'
  ✔ Κατέβηκε: python-logo.png
>>>
```

Web Title Extractor.py

```python
import requests
from bs4 import BeautifulSoup

url = "https://news.google.com"    # άλλαξέ το σε ό

r = requests.get(url)
soup = BeautifulSoup(r.text, "html.parser")

titles = soup.find_all(["h1", "h2", "h3"])

for t in titles:
    print("•", t.get_text(strip=True))

```

Shell

```
>>> %Run 'Web Title Extractor.py'
  • Your
 briefing
 • Top stories
 • Local news
 • Picks for you

>>>
```

```python
import requests
from bs4 import BeautifulSoup

url = "https://www.python.org"   # βάλε όποια σελίδα θέλεις
r = requests.get(url)

soup = BeautifulSoup(r.text, "html.parser")
text = soup.get_text(separator="\n", strip=True)

with open("page.txt", "w", encoding="utf-8") as f:
    f.write(text)

print("✔ Το καθαρό κείμενο αποθηκεύτηκε στο page.txt")
```

```
%Run Web2TXT.py
Το καθαρό κείμενο αποθηκεύτηκε στο page.txt
```

WebWatch.py

```python
import requests
import time
import hashlib

url = "https://www.python.org"   # όποιο site θες
print("Παρακολουθώ την σελίδα:", url)

old_hash = ""

while True:
    r = requests.get(url)
    new_hash = hashlib.md5(r.text.encode("utf-8")).hexdigest()

    if old_hash and new_hash != old_hash:
        print("⚠ Η ιστοσελίδα ΑΛΛΑΞΕ!")
    else:
        print("✔ Καμία αλλαγή...")

    old_hash = new_hash
    time.sleep(10)  # έλεγχος κάθε 10 δευτ.

```

Shell

```
>>> %Run WebWatch.py
Παρακολουθώ την σελίδα: https://www.python.org
✔ Καμία αλλαγή...
```