

# This Python program checks the price of an Amazon item everyday.

# import libraries

from bs4 import BeautifulSoup

import requests

import smtplib

import time

import datetime

# connect to website

URL = 'https://www.amazon.com/Funny-Data-Systems-Business-Analyst/dp/B07FNW9FGJ/ref=sr\_1\_14?'

crid=1293SFHKWAVM9&keywords=data%2Banalyst%2Btshirt&qid=1672511025&prefix=data%2Banalyst%2Btshirt%2Caps%2C145&sr=8-14'

headers = {"User-Agent": "Mozilla/5.0 (Macintosh; Intel Mac OS X 10\_15\_7) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/16.1 Safari/605.1.15", "Accept-Encoding": "gzip, deflate", "Accept": "text/html,application/xhtml+xml,application/xml;q=0.9,\*/\*;q=0.8", "DNT": "1", "Connection": "close", "Upgrade-Insecure-Requests": "1"}'

The screenshot shows an Amazon product page for a black t-shirt with the text "got data?". The product is by "MIS Business & Data Analyst Got Data Designs". The price is \$16.99, and it has 15 ratings. The page includes options to select color (Black, Navy, Green, Grey, Blue) and fit type (Men, Women, Youth). The size selection is set to Small. To the right, there are buttons for "Add to Cart" and "Add to List". Below the main product, there is a sponsored section for another t-shirt, "I Model - Data Science Statistics Data Analyst Data Modeling T-Shirt", priced at \$19.99. The top of the page shows the Amazon navigation bar with the search bar containing "data+analyst+tshirt".

```

page=requests.get(URL, headers=headers)
soup1 = BeautifulSoup(page.content, "html.parser")
soup2 = BeautifulSoup(soup1.prettify(), "html.parser")
print(soup2)

```

```

<!DOCTYPE html>

<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-us">
<!-- sp:feature:head-start -->
<head>
<script>
  var aPageStart = (new Date()).getTime();
</script>
<meta charset="utf-8"/>
<!-- sp:end-feature:head-start -->
<!-- sp:feature:csm:head-open-part1 -->
<script type="text/javascript">
  var ue_t0=ue_t0||+new Date();
</script>
<!-- sp:end-feature:csm:head-open-part1 -->
<!-- sp:feature:cs-optimization -->
<meta content="on" http-equiv="x-dns-prefetch-control"/>
<link href="https://images-na.ssl-images-amazon.com" rel="dns-prefetch"/>
<link href="https://m.media-amazon.com" rel="dns-prefetch"/>
<link href="https://completion.amazon.com" rel="dns-prefetch"/>

```

```

# get information
title = soup2.find(id='productTitle').get_text()
print(title)

```

Funny Got Data MIS Data Systems Business Analyst T-Shirt

```
price =  
soup2.find(id='corePriceDisplay_desktop_feature_div').get_text()
```

\$16.99

\$

16

.

99

```
# format information  
price = price.split()[0][1:]  
title = title.strip()  
print(title)  
print(price)
```

Funny Got Data MIS Data Systems Business Analyst T-Shirt  
16.99

```

type(price)

# get date stamp
today = datetime.date.today()
print(today)

# create csv file
import csv

# define csv columns
header = ['Title', 'Price', 'Date']

# define csv data
data = [title, price, today]

type(data)

# write to csv file
with open('/Users/jan/desktop/data/jy/future/codesamples/python/
Amazon/AmazonWebScraperDataset.csv', 'w', newline='',
encoding='UTF8') as f:
    writer = csv.writer(f)
    writer.writerow(header)
    writer.writerow(data)

# import csv file
import pandas as pd

df = pd.read_csv('/Users/jan/desktop/data/jy/future/codesamples/
python/Amazon/AmazonWebScraperDataset.csv')
print(df)

```

---

	Title	Price	Date
0	Funny Got Data MIS Data Systems Business Analy...	16.99	2022-12-31

---

```

# append data to csv file
with open('/Users/jan/desktop/data/jy/future/codesamples/python/
Amazon/AmazonWebScrapDataset.csv', 'a+', newline='',
encoding='UTF8') as f:
    writer = csv.writer(f)
    writer.writerow(data)

#import csv file
import pandas as pd

df = pd.read_csv('/Users/jan/desktop/data/jy/future/codesamples/
python/Amazon/AmazonWebScrapDataset.csv')
print(df)

```

		Title	Price	Date
0	Funny Got Data MIS Data Systems Business Analy...		16.99	2022-12-31
1	Funny Got Data MIS Data Systems Business Analy...		16.99	2022-12-31

```

def check_price():
    URL = 'https://www.amazon.com/Funny-Data-Systems-Business-
Analyst/dp/B07FNW9FGJ/ref=sr_1_14?
crid=1293SFHKWAVM9&keywords=data%2Banalyst%2Btshirt&qid=16725110
25&prefix=data%2Banalyst%2Btshirt%2Caps%2C145&sr=8-14'
    headers = {"User-Agent": "Mozilla/5.0 (Macintosh; Intel Mac
OS X 10_15_7) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/
16.1 Safari/605.1.15", "Accept-Encoding": "gzip, deflate",
"Accept": "text/html,application/xhtml+xml,application/
xml;q=0.9,*/*;q=0.8", "DNT": "1", "Connection": "close", "Upgrade-
Insecure-Requests": "1"}
    page=requests.get(URL, headers=headers)
    soup1 = BeautifulSoup(page.content, "html.parser")
    soup2 = BeautifulSoup(soup1.prettify(), "html.parser")

    # get information
    title = soup2.find(id='productTitle').get_text()
    price =
soup2.find(id='corePriceDisplay_desktop_feature_div').get_text()

    # format information
    price = price.split()[0][1:]
    title = title.strip()

```

```

# get date stamp
today = datetime.date.today()

# define csv columns
header = ['Title', 'Price', 'Date']

#define csv data
data = [title, price, today]

# append data to csv file
with open('/Users/jan/desktop/data/jy/future/codesamples/
python/Amazon/AmazonWebScrapperDataset.csv', 'a+', newline='',
encoding='UTF8') as f:
    writer = csv.writer(f)
    writer.writerow(data)

# check price every 5 seconds
while (True):
    check_price()
    time.sleep(5)

# manually stop running after 10 seconds

```

```

KeyboardInterrupt                                Traceback (most recent call last)
/var/folders/0y/2sf9nfpn6c3_nxjt5ys__r180000ggq/T/ipykernel_49427/364068478.py in <module>
>
      1 while (True):
      2     check_price()
----> 3     time.sleep(5)

KeyboardInterrupt:

```

```

df = pd.read_csv('/Users/jan/desktop/data/jy/future/codesamples/
python/Amazon/AmazonWebScrapperDataset.csv')
print(df)

```

	Title	Price	Date
0	Funny Got Data MIS Data Systems Business Analy...	16.99	2022-12-31
1	Funny Got Data MIS Data Systems Business Analy...	16.99	2022-12-31
2	Funny Got Data MIS Data Systems Business Analy...	16.99	2022-12-31
3	Funny Got Data MIS Data Systems Business Analy...	16.99	2022-12-31

```
# check price every day  
while (True):  
    check_price()  
    time.sleep(24*60*60)
```