

Online Degrees and Postgradua Coursera | Online Courses From Hands-on Lab: Analyzing Histor Skills Network Labs Inbox (3) - devikabylapudi228

labs.cognitiveclass.ai/v2/tools/jupyterlab?ulid=ulid-dc14633cdced674ba9df0ef610cb50132a2e27c

File Edit View Run Kernel Git Tabs Settings Help

OPEN TABS Close All

WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment Webscrapi... Final Assignment.ipynb

KERNELS Shut Down All

Final Assignment Webscrap... WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment.ipynb

TERMINALS Shut Down All

WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment Webscrapi... Final Assignment.ipynb

Markdown git Run as Pipeline

Python

Question 4: Use Webscraping to Extract GME Revenue Data

Use the `requests` library to download the webpage <https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0220EN-SkillsNetwork/labs/project/stock.html>. Save the text of the response as a variable named `html_data`.

```
[?]: #!pip install pandas==1.3.3
#!pip install requests==2.26.0
!mamba install bs4==4.10.0 -y
!mamba install html5lib==1.1 -y
!pip install lxml==4.6.4
#!pip install plotly==5.3.1

import pandas as pd
import requests
from bs4 import BeautifulSoup

import warnings
warnings.filterwarnings("ignore", category=FutureWarning)

url = "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0220EN-SkillsNetwork/labs/project/stock.html_data"
html_data = requests.get(url).text

html_data = requests.get(url).text
print(html_data)
```

Simple 0 4 Fully initialized Python | Idle Mem: 573.90 / 6144.00 MB Mode: Command Ln 1, Col 91 English (United States) Final Assignment.ipynb 1

Type here to search 15°C Cloudy 11:59 14/05/2024

5 New Messages

Coursera | Online Courses From

Hands-on Lab: Analyzing Histoi

Skills Network Labs

Inbox (3) - devikabylapudi228

labs.cognitiveclass.ai/v2/tools/jupyterlab?ulid=ulid-dc14633cdced674ba99df0ef610cb50132a2e27c

File Edit View Run Kernel Git Tabs Settings Help

OPEN TABS

WebScraping_Review_Lab.i...

Final_Assignment Library.ip...

Final Assignment Webscrap...

Final Assignment.ipynb

KERNELS

Final Assignment Webscrap...

WebScraping_Review_Lab.i...

Final_Assignment Library.ip...

Final Assignment.ipynb

TERMINALS

Shut Down All

WebScraping_Review_Lab.i...

Final_Assignment Library.ip...

Final Assignment Webscrap...

Final Assignment.ipynb

Python

Markdown

Run as Pipeline

PetaboxLoader3.resolve: 30.75

--><script> </script><script type="text/javascript" async="" src="/GameStop Revenue 2006-2020 _ Q&E _ MacroTrends_files/in.php"></script></bo

dy><grammarly-desktop-integration data-grammarly-shadow-root="true"></grammarly-desktop-integration></html>

Parse the html data using **beautiful_soup**.

[73]: `soup = BeautifulSoup(html_data, 'html5lib')`

Using **BeautifulSoup** or the **read_html** function extract the table with **GameStop Revenue** and store it into a dataframe named **gme_revenue**. The dataframe should have columns **Date** and **Revenue**. Make sure the comma and dollar sign is removed from the **Revenue** column using a method similar to what you did in Question 2.

► Click here if you need help locating the table

[74]: `GameStop_revenue = pd.DataFrame(columns=["Date", "Revenue"])`

`for row in soup.find_all("tbody"):`

`col = row.find_all("td"):`

`Date = col[0].text`

`Revenue = col[1].text`

`GameStop_revenue = GameStop_revenue.append({"Date":Date, "Revenue":Revenue}, ignore_index=True)`

Display the last five rows of the **gme_revenue** dataframe using the **tail** function. Take a screenshot of the results.

[75]: `GameStop_revenue.tail(5)`

[75]:

	Date	Revenue
1	2020-04-30	\$1,021
2	Retail/Wholesale	Retail - Consumer Electronics

Simple 0 4

Fully initialized Python | Idle Mem: 573.90 / 6144.00 MB

Mode: Command Ln 1, Col 91 English (United States) Final Assignment.ipynb 1

Type here to search

15°C Cloudy 11:59 14/05/2024

Online Degrees and Postgraduat... Coursera | Online Courses From... Hands-on Lab: Analyzing Histor... Skills Network Labs Inbox (3) - devikabylapudi228@

labs.cognitiveclass.ai/v2/tools/jupyterlab?ulid=ulid-dc14633cdced674ba99df0ef610cb50132a2e27c

File Edit View Run Kernel Git Tabs Settings Help

OPEN TABS Close All

WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment Webscrapi... Final Assignment.ipynb

KERNELS Shut Down All

Final Assignment Webscrap... WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment.ipynb

TERMINALS Shut Down All

WebScraping_Review_Lab.i... Final_Assignment Library.ip... Final Assignment Webscrapi... Final Assignment.ipynb

Markdown Run as Pipeline

Python

Using BeautifulSoup or the read_html function extract the table with GameStop Revenue and store it into a dataframe named gme_revenue. The dataframe should have columns Date and Revenue. Make sure the comma and dollar sign is removed from the Revenue column using a method similar to what you did in Question 2.

Click here if you need help locating the table

[74]: GameStop_revenue = pd.DataFrame(columns=["Date", "Revenue"])

for row in soup.find_all("tbody"):
 col = row.find_all("td")
 Date = col[0].text
 Revenue = col[1].text

GameStop_revenue = GameStop_revenue.append({"Date":Date, "Revenue":Revenue}, ignore_index=True)

Display the last five rows of the gme_revenue dataframe using the tail function. Take a screenshot of the results.

[75]: GameStop_revenue.tail(5)

[75]:

	Date	Revenue
1	2020-04-30	\$1.021
2	Retail/Wholesale	Retail - Consumer Electronics
3	Best Buy (BBY)	United States
4	GameStop Revenue 2006-2020 GME	
5	GameStop Revenue 2006-2020 GME	

Question 5: Plot Tesla Stock Graph

Simple 0 4 Fully initialized Python | Idle Mem: 573.90 / 6144.00 MB Mode: Command Ln 1, Col 91 English (United States) Final Assignment.ipynb 1

Type here to search

15°C Cloudy 12:00 14/05/2024