

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

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 Webscrap...  
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

4 2010-07-06 1.333333 1.333333 1.055333 1.074000 103003500 0 0.0

### Question 2: Use Webscraping to Extract Tesla Revenue Data

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

```
[76]: install_pandas==1.3.3  
install_requests==2.26.0  
a install bs4==4.10.0 -y  
a install html5lib==1.1 -y  
install lxml==4.6.4  
install_plotly==5.3.1  
  
t_pandas as pd  
t_requests  
ps4 import BeautifulSoup  
  
t_warnings  
ngs.filterwarnings("ignore", category=FutureWarning)  
  
"https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDDeveloperSkillsNetwork-PY0220EN-SkillsNetwork/labs/project/revenue.htm"  
data_ = requests.get(url).text  
  
data_ = requests.get(url).text  
(html_data)
```

cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/.../revenue.htm initialized Python | Idle Mem: 574.81 / 6144.00 MB Mode: Command Ln 1, Col 1 English (United States) Final Assignment.ipynb 1

Type here to search 15°C Cloudy 12:03 14/05/2024

5 New Messages!

Coursera | Online Courses From

Hands-on Lab: Analyzing Histor

Skills Network Labs

Inbox (3) - devikabylapudi228@

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

FileEditViewRunKernelGitTabsSettingsHelp

OPEN TABS

Close All

WebScraping\_Review\_Lab.i...

Final\_Assignment Library.ip...

Final Assignment Webscrap...

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

Python

Parse the html data using `beautiful_soup`.

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

Using `BeautifulSoup` or the `read_html` function extract the table with `Tesla Revenue` and store it into a dataframe named `tesla_revenue`. The dataframe should have columns `Date` and `Revenue`.

▼ Click here if you need help locating the table

Below is the code to isolate the table, you will now need to loop through the rows and columns like in the previous lab

`soup.find_all("tbody")[1]`

If you want to use the `read_html` function the table is located at index 1

We are focusing on quarterly revenue in the lab.

[78]: `tesla_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`

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

Execute the following line to remove the comma and dollar sign from the `Revenue` column.

[79]: `tesla_revenue["Revenue"] = tesla_revenue['Revenue'].str.replace(',', '\$', "")`

Execute the following lines to remove an null or empty strings in the Revenue column.

Simple0\$4Fully initializedPython | IdleMem: 575.08 / 6144.00 MBMode: CommandLn 1, Col 1English (United States)Final Assignment.ipynb1

Type here to search15°C Cloudy12:0314/05/2024

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

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 Webscrap... 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 Webscrap... Final Assignment.ipynb

tesla\_revenue.dropna(inplace=True)

tesla\_revenue = tesla\_revenue[tesla\_revenue['Revenue'] != ""]

Display the last 5 row of the tesla\_revenue dataframe using the tail function. Take a screenshot of the results.

[81]: tesla\_revenue.tail(5)

	Date	Revenue
0	2021	53823
1	2022-09-30	21454
2	Auto/Tires/Trucks	Auto Manufacturers - Domestic
3	General Motors (GM)	United States

Question 3: Use yfinance to Extract Stock Data

Using the Ticker function enter the ticker symbol of the stock we want to extract data on to create a ticker object. The stock is GameStop and its ticker symbol is GME .

[69]: !pip install yfinance==0.2.4  
#!pip install pandas==1.3.3

import yfinance as yf  
import pandas as pd

GameStop = yf.Ticker("GME")

Collecting yfinance==0.2.4  
Using cached yfinance-0.2.4-py2.py3-none-any.whl (51 kB)

Simple 0 \$ 4 Fully initialized Python | Idle Mem: 575.08 / 6144.00 MB Mode: Command Ln 1, Col 1 English (United States) Final Assignment.ipynb 1

Type here to search 15°C Cloudy 12:04 14/05/2024