

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_2`.

```
[40]: url = "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0220EN-SkillsNetwork/labs/project/stock.html"
      html_data_2 = requests.get(url)
      print(html_data_2)
```

<Response [200]>

Parse the html data using `beautiful_soup` using parser i.e `html5lib` or `html.parser`.

```
[41]: soup = BeautifulSoup(html_data_2.content, 'html.parser')
```

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.

Note: Use the method similar to what you did in question 2.

► Click here if you need help locating the table

```
[42]: soup.find_all("tbody")
      GME_Revenue = pd.DataFrame(columns=["Date", "Revenue"])
```

► Click here if you need help locating the table

```
[40]: soup.find_all("tbody")
      GME_Revenue = pd.DataFrame(columns=["Date", "Revenue"])
```

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

```
[41]: GME_Revenue.tail()
```

```
[41]:   Date Revenue
```