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/IBMDeveloperSkillsNetwork-PY0220EN-SkillsNetwork/labs/project/revenue.htm Save the text of the response as a variable named `html_data`.

35]:
```
import requests
url = "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0220EN
response = requests.get(url)
```

36]:
```
html_data = response.text
```

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

37]:
```
from bs4 import BeautifulSoup

soup = BeautifulSoup(html_data, "html.parser")
```

We are focusing on quarterly revenue in the lab.

[88]:
```
tbodies = soup.find_all("tbody")

tesla_quarterly_table = tbodies[1]
```

[89]:
```
import pandas as pd

tesla_revenue = pd.DataFrame(columns=["Date", "Revenue"])
```

[90]:
```
for row in tesla_quarterly_table.find_all("tr"):
    cols = row.find_all("td")
    if len(cols) == 2:
        date = cols[0].text.strip()
        revenue = cols[1].text.strip()
        # Append to DataFrame
        tesla_revenue = pd.concat([
            tesla_revenue,
            pd.DataFrame({"Date": [date], "Revenue": [revenue]})
        ], ignore_index=True)
```

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

```
[91]: tesla_revenue["Revenue"] = tesla_revenue['Revenue'].str.replace(',|\$',"",regex=True)
```

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

```
[92]: 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.

```
[93]: tesla_revenue.tail()
```

[93]:

	Date	Revenue
48	2010-09-30	31
49	2010-06-30	28
50	2010-03-31	21
52	2009-09-30	46
53	2009-06-30	27