

## 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`.

```
[7]: url = "https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0220EN-SkillsNe-
html_data = requests.get(url).text
```

Parse the html data using `beautiful_soup`.

```
[8]: soup = BeautifulSoup(html_data)
```

Using `BeautifulSoup` or the `read_html` function extract the table with `Tesla Quarterly 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

```
[30]: read_html_tesla_data = pd.read_html(url)
read_html_tesla_data = pd.read_html(str(soup))
tesla_revenue = read_html_tesla_data[1]
tesla_revenue = tesla_revenue.rename(columns = {tesla_revenue.columns[0]: "Date", tesla_revenue.columns[1]: "Revenue"},
#tesla_revenue
```

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

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

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

/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages/ipykernel\_launcher.py:1: FutureWarning: The default value of `regex` will change from `True` to `False` in a future version.  
"""Entry point for launching an IPython kernel.

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

```
[11]: tesla_revenue.dropna(inplace=True)

tesla_revenue = tesla_revenue[tesla_revenue['Revenue'] != '']
```

Display the last 5 rows of the `tesla_revenue` dataframe using the `tail` function. Take a screenshot of the results.

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

```
[12]:
```

|    | 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      |