Practice #1: Let's Play With Cnet.com

https://www.cnet.com/reviews/samsung-un65f9000-review/

In [ ]:
```
from bs4 import BeautifulSoup
from urllib.request import urlopen
```

In [ ]:
```
infile = urlopen('https://www.cnet.com/reviews/samsung-un65f9000-review/')
contents = infile.read()
infile.close()
```

In [ ]:
```
soup = BeautifulSoup(contents, 'html.parser')
```

In [ ]:
```
print(soup.prettify())
```

Q1: Find the average rating score out of ten.

> Use 'Inspect Element'
>
> ```
> 
>     "7.3"
>     Overall
> 
> ```

In [ ]:
```
soup.find('span')
```

In [ ]:
```
soup.find_all('span')
```

In [ ]:
```
soup.find_all('span', {"class": "text"})
# soup.find_all('span', {"class": "c-review-average", "aria-hidden": "true"})
```

In [ ]:
```
soup.find_all('span', {"class": "text"})[0]
```

In [ ]:
```
soup.find_all('span', {"class": "text"})[0].get_text()
```

In [ ]:
```
soup.find_all('span', {"class": "text"})[0].get_text()[:3]
```

In [ ]:
```
float(soup.find_all('span', {"class": "text"})[0].get_text()[:3])
```

Q2: Find the price of the TV

```
<span class="msrpPrice">$7,499.99</span>
```

In [ ]:
```
soup.find_all('span', {"class": 'msrpPrice'})
```

In [ ]:
```
soup.find_all('span', {"class": 'msrpPrice'})[0].get_text()
```

In [ ]:
```
soup.find_all('span', {"class": 'msrpPrice'})[0].get_text()[1: ]
```

In [ ]:
```
soup.find_all('span', {"class": 'msrpPrice'})[0].get_text()[1: ].replace(',', '')
```

In [ ]:
```
float(soup.find_all('span', {"class": 'msrpPrice'})[0].get_text()[1: ].replace(',', '')
```

Practice #2: Let's Play With Wallmart.com

https://www.walmart.com/ip/SAMSUNG-65-Class-4K-Ultra-HD-2160P-HDR-Smart-QLED-TV-QN65Q60TB/183583576

In [ ]:
```
from bs4 import BeautifulSoup
from urllib.request import urlopen
```

In [ ]:
```
infile = urlopen('https://www.walmart.com/ip/SAMSUNG-65-Class-4K-Ultra-HD-2160P-HDR-Sma
contents = infile.read()
infile.close()
```

In [ ]:
```
soup = BeautifulSoup(contents, 'html.parser')
```

In [ ]:
```
print(soup.prettify())
```

Q1: Find the average rating score out of five.

```
<span itemprop="ratingValue">4.7</span>
```

In [ ]:
```
soup.find_all('span')
```

In [ ]:
```
import urllib.request

url = 'https://www.walmart.com/ip/SAMSUNG-65-Class-4K-Ultra-HD-2160P-HDR-Smart-QLED-TV-(
```

```
page = urllib.request.Request(url, headers = {'User-Agent': 'Mozilla/5.0'})
infile = urllib.request.urlopen(page)
```

In [ ]:
```
contents = infile.read()
soup = BeautifulSoup(contents, "html.parser")
```

In [ ]:
```
print(soup.prettify())
```

In [ ]:
```
soup.find_all('span')
```

In [ ]:
```
soup.find_all('span', {'itemprop': 'ratingValue'})
```

In [ ]:
```
soup.find_all('span', {'itemprop': 'ratingValue'})[0].get_text()
```

Q2: Find the price of the TV

```
<span class="price-characteristic" itemprop="price"
content="897.99">897</span>
```

In [ ]:
```
soup.find_all('span', {'class': 'price-characteristic'})
```

In [ ]:
```
soup.find_all('span', {'class': 'price-characteristic', 'itemprop': 'price'})
```

In [ ]:
```
soup.find_all('span', {'class': 'price-characteristic', 'itemprop': 'price'})[0].get_te
```

In [ ]:
```
soup.find_all('span', {'class': 'price-characteristic', 'itemprop': 'price'})[0]['conte
```

Q3: Find the reviews of the TV

```
<div class="review-text"><p>Stable wifi access to 5.0 GHz channel, crazy
contrast level, awesome billion colors, ambient art mood, VA panel, deep
dark colors, Samsung TV plus, easy to use control and simplified
minimalistic menu, apple airplay cast, ready DLNA and built in alexa that
connects with my camera and lamps, slim profile, no need to buy another
apple tv. 60 Hz refresh rate absolutely compensated with quantum
processor and make me want to watch HBO, Amazon prime and Hulu all day.
These all have built in app in Tizen OS on top of Netflix and other
services. I am so happy with this purchase specially with not buying a
stupid IPS screen with 1/6th of contrast ratio and so called ”natural
colors”. I have seen those natural colors in my living room enough and I
think my work room Samsung Q60 is by far better.</p></div>
```

In [ ]:
```
soup.find_all('div', {'class': 'review-text'})
```

In [ ]:
```
len(soup.find_all('div', {'class': 'review-text'}))
```

In [ ]:
```
soup.find_all('div', {'class': 'review-text'})[0]
```

In [ ]:
```
soup.find_all('div', {'class': 'review-text'})[0].get_text()
```