In [6]:
```
"""
Homework-12-Web-Scraping-and-Document-Databases
"""
from splinter import Browser
from bs4 import BeautifulSoup
import pandas
import time
path={'executable_path':'/Users/coffm/Downloads/chromedriver.exe'}
page=Browser('chrome', **path, headless=False)
```

In [5]:
```
# Web Scraping - Mission to Mars
# Scrape the [NASA Mars page Site](https://mars.nasa.gov/page/)
# Collect the latest page Title and Paragraph Text. Assign the text to variables that you can reference later.
url='https://mars.nasa.gov/news/'
page.visit(url)
html=page.html
soup=BeautifulSoup(html, 'html.parser')
time.sleep(3)
news=soup.find('div', class_='list_text')
newsTitle=news.find('div', class_='content_title').text
print('\n'+newsTitle)
newsParagraph=news.find('div', class_ ='article_teaser_body').text
print(newsParagraph)
```

```
A Martian Roundtrip: NASA's Perseverance Rover Sample Tubes
Marvels of engineering, the rover's sample tubes must be tough enough to safely bring Red Planet samples on th
e long journey back to Earth in immaculate condition.
```

In [7]:
```
# JPL Mars Space Images - Featured Image
# Visit the url for JPL Featured Space Image (https://www.jpl.nasa.gov/spaceimages/?search=&category=Mars).
# Use splinter to navigate the site and find the image url for the current Featured Mars Image and assign the ur
# Make sure to find the image url to the full size `.jpg` image.
# Make sure to save a complete url string for this image.
imageURL='https://www.jpl.nasa.gov/spaceimages/?search=&category=Mars'
page.visit(imageURL)
html=page.html
soup=BeautifulSoup(html, 'html.parser')
time.sleep(3)
image = (soup.find_all('div', class_='carousel_items')[0].a.get('data-fancybox-href'))
featured = 'https://www.jpl.nasa.gov'+ image
print('\n'+featured)
```

```
https://www.jpl.nasa.gov/spaceimages/images/mediumsize/PIA18897_ip.jpg (https://www.jpl.nasa.gov/spaceimages/i
mages/mediumsize/PIA18897_ip.jpg)
```

In [10]:
```
# Visit the Mars Facts webpage [here](https://space-facts.com/mars/)
# Use Pandas to scrape the table containing facts about the planet including Diameter, Mass, etc.
# Use Pandas to convert the data to a HTML table string.
facts_url='https://space-facts.com/mars/'
page.visit(facts_url)
marsData=pandas.read_html(facts_url)
marsData=pandas.DataFrame(marsData[0])
marsFacts=marsData.to_html(header=False, index=False)
marsFacts=((pandas.read_html(marsFacts))[0]).rename(columns={0: 'Attribute', 1: 'Value'}).set_index(['Attribute'
print('\n',marsFacts)
```

Attribute	Value
Equatorial Diameter:	6,792 km
Polar Diameter:	6,752 km
Mass:	6.39 × 10^23 kg (0.11 Earths)
Moons:	2 (Phobos & Deimos)
Orbit Distance:	227,943,824 km (1.38 AU)
Orbit Period:	687 days (1.9 years)
Surface Temperature:	-87 to -5 °C
First Record:	2nd millennium BC
Recorded By:	Egyptian astronomers

In [9]:
```
# Visit the USGS Astrogeology site [here](https://astrogeology.usgs.gov/search/results?q=hemisphere+enhanced&k1=
# You will need to click each of the links to the hemispheres in order to find the image url to the full resolut
# Save both the image url string for the full resolution hemisphere image, and the Hemisphere title containing t
# Append the dictionary with the image url string and the hemisphere title to a list. This list will contain one
hemispheresURL='https://astrogeology.usgs.gov/search/results?q=hemisphere+enhanced&k1=target&v1=Mars'
page.visit(hemispheresURL)
html=page.html
marsHemisphere=[]
soup=BeautifulSoup(html, 'html.parser')
time.sleep(3)
products=soup.find('div', class_='result-list' )
hemispheres=products.find_all('div', class_='item')
for hemisphere in hemispheres:
    title=hemisphere.find('h3').text
    title=title.replace('Enhanced', '')
    endURL=hemisphere.find('a')['href']
    imageURL='https://astrogeology.usgs.gov/' + endURL
    page.visit(imageURL)
    html=page.html
    soup=BeautifulSoup(html, 'html.parser')
    downloads=soup.find('div', class_='downloads')
    imageURL=downloads.find('a')['href']
    marsHemisphere.append({'title': title, 'img_url': imageURL})
from pprint import pprint
print('\n')
pprint(marsHemisphere)
```

```
[{'img_url': 'https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/cerberus_enhanced.tif/full.jpg',
  'title': 'Cerberus Hemisphere '},
 {'img_url': 'https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/schiaparelli_enhanced.tif/full.jp
g',
  'title': 'Schiaparelli Hemisphere '},
 {'img_url': 'https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/syrtis_major_enhanced.tif/full.jp
g',
  'title': 'Syrtis Major Hemisphere '},
 {'img_url': 'https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/valles_marineris_enhanced.tif/ful
l.jpg',
  'title': 'Valles Marineris Hemisphere '}]
```

In [ ]: