```
In [1]: # Import Splinter, BeautifulSoup, and Pandas
    from splinter import Browser
    from bs4 import BeautifulSoup as soup
    import pandas as pd
    from webdriver_manager.chrome import ChromeDriverManager
```

```
In [ ]: # Set the executable path and initialize Splinter
    executable_path = {'executable_path': ChromeDriverManager().install()}
    browser = Browser('chrome', **executable_path, headless=False)
```

Visit the NASA Mars News Site

```
In [3]: # Visit the mars nasa news site
        url = 'https://redplanetscience.com/'
        browser.visit(url)
        # Optional delay for loading the page
        browser.is element present by css('div.list text', wait time=1)
Out[3]: True
In [4]: # Convert the browser html to a soup object and then quit the browser
        html = browser.html
        news soup = soup(html, 'html.parser')
        slide elem = news soup.select one('div.list text')
In [5]: |slide_elem.find('div', class_='content_title')
Out[5]: <div class="content title">NASA's Mars 2020 Heads Into the Test Chamber</
        div>
In [6]: # Use the parent element to find the first a tag and save it as `news title
        news title = slide elem.find('div', class ='content title').get text()
        news title
Out[6]: "NASA's Mars 2020 Heads Into the Test Chamber"
In [7]: # Use the parent element to find the paragraph text
        news p = slide elem.find('div', class = 'article teaser body').get text()
        news p
Out[7]: 'In this time-lapse video taken at JPL, engineers move the Mars 2020 rove
        r into a large vacuum chamber for testing in Mars-like environmental cond
```

JPL Space Images Featured Image

itions.'

```
In [8]: # Visit URL
         url = 'https://spaceimages-mars.com'
         browser.visit(url)
 In [9]: # Find and click the full image button
         full image elem = browser.find by tag('button')[1]
         full image elem.click()
In [10]: # Parse the resulting html with soup
         html = browser.html
         img_soup = soup(html, 'html.parser')
         img soup
Out[10]: <html class="fancybox-margin fancybox-lock"><head>
         <meta charset="utf-8"/>
         <meta content="width=device-width, initial-scale=1" name="viewport"/>
         <link href="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0/css/bootstra</pre>
         p.min.css" rel="stylesheet"/>
         <!-- <link rel="stylesheet" type="text/css" href="css/font.css"> -->
         <link href="css/app.css" rel="stylesheet" type="text/css"/>
         <link href="https://stackpath.bootstrapcdn.com/font-awesome/4.7.0/css/fon</pre>
         t-awesome.min.css" rel="stylesheet" type="text/css"/>
         <title>Space Image</title>
         <style type="text/css">.fancybox-margin{margin-right:0px;}</style></head>
         <body>
         <div class="header">
         <nav class="navbar navbar-expand-lg">
         <a class="navbar-brand" href="#"><imq id="logo" src="image/nasa.png"/><sp</pre>
         an class="logo">Jet Propulsion Laboratory</span>
         <span class="logo1">California Institute of Technology</span></a>
         <button aria-controls="navbarNav" aria-expanded="false" aria-label="Toggl</pre>
         e navigation" class="navbar-toggler" data-target="#navbarNav" data-toggle
In [11]: # find the relative image url
         img url rel = img soup.find('img', class ='fancybox-image').get('src')
         img url rel
Out[11]: 'image/featured/mars3.jpg'
In [12]: # Use the base url to create an absolute url
         img url = f'https://spaceimages-mars.com/{img url rel}'
         img url
Out[12]: 'https://spaceimages-mars.com/image/featured/mars3.jpg'
```

Mars Facts

```
df = pd.read_html('https://galaxyfacts-mars.com')[0]
        df.head()
Out[13]:
                        0
                                  1
                                             2
        0 Mars - Earth Comparison
                                Mars
                                           Earth
                   Diameter:
        1
                              6,779 km
                                        12,742 km
                     Mass: 6.39 × 10^23 kg 5.97 × 10^24 kg
        2
        3
                     Moons:
              Distance from Sun: 227,943,824 km 149,598,262 km
In [14]: df.columns=['Description', 'Mars', 'Earth']
        df.set_index('Description', inplace=True)
Out[14]:
                               Mars
                                          Earth
                 Description
                                          Earth
        Mars - Earth Comparison
                               Mars
                             6,779 km
                  Diameter:
                                       12,742 km
                    Mass: 6.39 \times 10^2 3 \text{ kg} 5.97 \times 10^2 4 \text{ kg}
                   Moons:
            Distance from Sun: 227,943,824 km 149,598,262 km
                          687 Earth days
                                      365.24 days
               Length of Year:
                           -87 to -5 °C
                                      -88 to 58°C
                Temperature:
In [15]: | df.to html()
Out[15]: '\n <thead>\n
                                                       <tr style="text-ali
        gn: right;">\n
                         \n
                                        Mars
                                                           Earth
             \n
                       \n
                                 Description\n

n </thead>
n 
n
                                                    n
                                                              Mars - E
        arth Comparison\n
                                Mars
                                                   Earth
                                                                     </tr
        >\n
              \n
                        Diameter:
                                               6,779 km
                                                                      <td
        >12,742 \text{ km}</\text{td}>\n
                          \n
                                    \n
                                               Mass:
                                                                  >6.3
        9 \times 10^23 \text{ kg}/\text{td}
                              5.97 \times 10^24 \text{ kg}/\text{td}

                                                                   \ n
        Moons:
                            2\n
                                            1\n

n
               Distance from Sun:
                                              227,943,824 km\n
                                                      Length of Year:</
        149,598,262 km
                                 \n
                                           \n
        th>\n
                 687 Earth days\n
                                             365.24 days
             \n
                        Temperature:
                                                 -87 to -5 °C\n
        -88 to 58°C\n

n n'
```

D1: Scrape High-Resolution Mars' Hemisphere

Images and Titles

Hemispheres

```
In [16]: # 1. Use browser to visit the URL
         url = 'https://marshemispheres.com/'
         browser.visit(url)
In [17]: # 2. Create a list to hold the images and titles.
         hemisphere image urls = []
         # 3. Write code to retrieve the image urls and titles for each hemisphere.
In [18]:
         # 4. Print the list that holds the dictionary of each image url and title.
         hemisphere image urls
Out[18]: [{'img url': 'https://marshemispheres.com/images/full.jpg',
           'title': 'Cerberus Hemisphere Enhanced'},
          {'img url': 'https://marshemispheres.com/images/schiaparelli enhanced-fu
         ll.jpg',
           'title': 'Schiaparelli Hemisphere Enhanced'},
          {'img url': 'https://marshemispheres.com/images/syrtis major enhanced-fu
         11.jpg',
           'title': 'Syrtis Major Hemisphere Enhanced'},
          {'img url': 'https://marshemispheres.com/images/valles marineris enhance
         d-full.jpg',
           'title': 'Valles Marineris Hemisphere Enhanced'}]
In [19]: # 5. Quit the browser
         browser.quit()
In [ ]:
```