

```
In [1]: # import dependencies
from splinter import Browser
from bs4 import BeautifulSoup as bs
import pandas as pd
import time
```

## Step 1 - Scraping

### NASA Mars News

```
In [2]: # Set the chrome driver
executable_path = {"executable_path": "d:/chrome_driver/chromedriver.exe"}
```

```
In [3]: # Open the NASA's Mars news page on Chrome
browser = Browser("chrome", **executable_path, headless=False)
url = "https://mars.nasa.gov/news/"
browser.visit(url)
time.sleep(1)
```

```
In [4]: # Read html from the page
html = browser.html
soup = bs(html, "html.parser")
```

```
In [5]: # Scrape the very first news title and paragraph text
news_title = soup.find_all("div", class_="content_title")[1].text
news_p = soup.find("div", class_="article_teaser_body").text

print("Title: ", news_title)
print("Paragraph: ", news_p)
```

Title: NASA Engineers Checking InSight's Weather Sensors

Paragraph: An electronics issue is suspected to be preventing the sensors from sharing their data about Mars weather with the spacecraft.

### JPL Mars Space Images - Featured Image

```
In [6]: # Open the JPL page on Chrome
url = "https://www.jpl.nasa.gov/spaceimages/?search=&category=Mars"
browser.visit(url)
```

```
In [7]: # Move to the page having the full size image
browser.links.find_by_partial_text("FULL IMAGE").first.click()
browser.links.find_by_partial_text("more info").first.click()
browser.find_by_text("Full-Res JPG: ").first.find_by_tag("a").first.click()
```

```
In [8]: # Scrape the image url
featured_image_url = browser.find_by_tag("img").first["src"]
print(featured_image_url)
```

<https://photojournal.jpl.nasa.gov/jpeg/PIA20318.jpg>

### Mars Facts

```
In [9]: # Scrape the tables from https://space-facts.com/mars/
url = "https://space-facts.com/mars/"
tables = pd.read_html(url)
```

```
In [10]: html_tables = []

# Convert all the tables to html, and save them to a list
for table in tables:
    html_tables.append(table.to_html(justify="left").replace("\n", ""))

print(html_tables)

['<table border="1" class="dataframe"> <thead> <tr style="text-align: left;"> <th></th> <th>0</th> <th>1</th> </tr>
</thead> <tbody> <tr> <th>0</th> <th>Equatorial Diameter:</th> <td>6,792 km</td> </tr> <tr> <th>1</th>
<td>Polar Diameter:</td> <td>6,752 km</td> </tr> <tr> <th>2</th> <td>Mass:</td> <td>6.39 × 1023 kg (0.11 Earth
s)</td> </tr> <tr> <th>3</th> <td>Moons:</td> <td>2 (Phobos & Deimos)</td> </tr> <tr> <th>4</th>
<td>Orbit Distance:</td> <td>227,943,824 km (1.38 AU)</td> </tr> <tr> <th>5</th> <td>Orbit Period:</td> <td>687 d
ays (1.9 years)</td> </tr> <tr> <th>6</th> <td>Surface Temperature:</td> <td>-87 to -5 ° C</td> </tr> <tr>
<th>7</th> <td>First Record:</td> <td>2nd millennium BC</td> </tr> <tr> <th>8</th> <td>Recorded By:</td> <td>
Egyptian astronomers</td> </tr> </tbody></table>', '<table border="1" class="dataframe"> <thead> <tr style="text-align: left;">
<th></th> <th>Mars - Earth Comparison</th> <th>Mars</th> <th>Earth</th> </tr> </thead> <tbody> <tr> <th>0</th>
<td>Diameter:</td> <td>6,779 km</td> <td>12,742 km</td> </tr> <tr> <th>1</th> <td>Mass:</td> <td>6.39 × 1023
kg</td> <td>5.97 × 1024 kg</td> </tr> <tr> <th>2</th> <td>Moons:</td> <td>2</td> <td>1</td> </tr>
<tr> <th>3</th> <td>Distance from Sun:</td> <td>227,943,824 km</td> <td>149,598,262 km</td> </tr> <tr> <th>4<
/th> <td>Length of Year:</td> <td>687 Earth days</td> <td>365.24 days</td> </tr> <tr> <th>5</th> <td>Temper
ature:</td> <td>-87 to -5 ° C</td> <td>-88 to 58 ° C</td> </tr> </tbody></table>', '<table border="1" class="dataframe"> <the
ad> <tr style="text-align: left;"> <th></th> <th>0</th> </tr> </thead> <tbody> <tr> <th>0</th>
<td>Equatorial Diameter:</td> <td>6,792 km</td> </tr> <tr> <th>1</th> <td>Polar Diameter:</td> <td>6,752 km</td>
</tr> <tr> <th>2</th> <td>Mass:</td> <td>6.39 × 1023 kg (0.11 Earths)</td> </tr> <tr> <th>3</th> <td>Mo
ons:</td> <td>2 (Phobos & Deimos)</td> </tr> <tr> <th>4</th> <td>Orbit Distance:</td> <td>227,943,824 km (1.3
8 AU)</td> </tr> <tr> <th>5</th> <td>Orbit Period:</td> <td>687 days (1.9 years)</td> </tr> <tr> <th>6</th>
<td>Surface Temperature:</td> <td>-87 to -5 ° C</td> </tr> <tr> <th>7</th> <td>First Record:</td> <td>2nd millenn
ium BC</td> </tr> <tr> <th>8</th> <td>Recorded By:</td> <td>Egyptian astronomers</td> </tr> </tbody></table>']
```

## Mars Hemispheres

```
In [11]: # Open the Mars Hemispheres page on Chrome
url = "https://astrogeology.usgs.gov/search/results?q=hemisphere+enhanced&kl=target&vl=Mars"
browser.visit(url)
```

```
In [12]: hemisphere_image_urls = []

# Get the number of images
num_of_img = len(browser.find_by_css("img[class='thumb']"))

# Scrape each image title and url
for i in range(num_of_img):
    browser.find_by_css("img[class='thumb']")[i].click()
    hemisphere_image_urls.append({"title": browser.find_by_css("h2[class='title']").first.text.replace(" Enhanced", ""),
                                "img_url": browser.find_by_text("Sample").first["href"]})
    browser.back()
```

```
In [13]: for hemisphere_url in hemisphere_image_urls:
    print("Title: ", hemisphere_url["title"])
    print("img_url: ", hemisphere_url["img_url"])
```

```
Title: Cerberus Hemisphere
img_url: https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/cerberus_enhanced.tif/full.jpg
Title: Schiaparelli Hemisphere
img_url: https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/schiaparelli_enhanced.tif/full.jpg
Title: Syrtis Major Hemisphere
img_url: https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/syrtis_major_enhanced.tif/full.jpg
Title: Valles Marineris Hemisphere
img_url: https://astropedia.astrogeology.usgs.gov/download/Mars/Viking/valles_marineris_enhanced.tif/full.jpg
```