

Space X Falcon 9 First Stage Landing Prediction

Web scraping Falcon 9 and Falcon Heavy Launches Records from Wikipedia

Estimated time needed: **40** minutes

In this lab, you will be performing web scraping to collect Falcon 9 historical launch records from a Wikipedia page titled `List of Falcon 9 and Falcon Heavy launches`

https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches

Falcon 9 first stage will land successfully

Several examples of an unsuccessful landing are shown here:

More specifically, the launch records are stored in a HTML table shown below:

Objectives

Web scrap Falcon 9 launch records with `BeautifulSoup` :

- Extract a Falcon 9 launch records HTML table from Wikipedia
- Parse the table and convert it into a Pandas data frame

First let's import required packages for this lab

```
In [1]: !pip3 install beautifulsoup4
!pip3 install requests
!pip3 install html5lib
```

```

Requirement already satisfied: beautifulsoup4 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (4.11.1)
Requirement already satisfied: soupsieve>1.2 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from beautifulsoup4) (2.3.2.post1)
Requirement already satisfied: requests in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (2.29.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests) (3.1.0)
Requirement already satisfied: idna<4,>=2.5 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests) (1.26.15)
Requirement already satisfied: certifi>=2017.4.17 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from requests) (2023.5.7)
Collecting html5lib
  Downloading html5lib-1.1-py2.py3-none-any.whl (112 kB)
    _____ 112.2/112.2 kB 17.7 MB/s eta 0:00:00
Requirement already satisfied: six>=1.9 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from html5lib) (1.16.0)
Requirement already satisfied: webencodings in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (from html5lib) (0.5.1)
Installing collected packages: html5lib
Successfully installed html5lib-1.1

```

```

In [2]: import sys

import requests
from bs4 import BeautifulSoup
import re
import unicodedata
import pandas as pd
import numpy as np

```

and we will provide some helper functions for you to process web scraped HTML table

```

In [3]: def date_time(table_cells):
        """
        This function returns the data and time from the HTML table cell
        Input: the element of a table data cell extracts extra row
        """
        return [data_time.strip() for data_time in list(table_cells.strings)][0:2]

    def booster_version(table_cells):
        """
        This function returns the booster version from the HTML table cell
        Input: the element of a table data cell extracts extra row
        """
        out=''.join([booster_version for i,booster_version in enumerate(table_cells
        return out

    def landing_status(table_cells):
        """
        This function returns the landing status from the HTML table cell
        Input: the element of a table data cell extracts extra row
        """
        out=[i for i in table_cells.strings][0]
        return out

```

```
def get_mass(table_cells):
    mass=unicodedata.normalize("NFKD", table_cells.text).strip()
    if mass:
        mass.find("kg")
        new_mass=mass[0:mass.find("kg")+2]
    else:
        new_mass=0
    return new_mass

def extract_column_from_header(row):
    """
    This function returns the landing status from the HTML table cell
    Input: the element of a table data cell extracts extra row
    """
    if (row.br):
        row.br.extract()
    if row.a:
        row.a.extract()
    if row.sup:
        row.sup.extract()

    column_name = ' '.join(row.contents)

    # Filter the digit and empty names
    if not(column_name.strip().isdigit()):
        column_name = column_name.strip()
    return column_name
```

To keep the lab tasks consistent, you will be asked to scrape the data from a snapshot of the `List of Falcon 9 and Falcon Heavy launches` Wikipage updated on `9th June 2021`

In [4]: `static_url = "https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Fa`

Next, request the HTML page from the above URL and get a `response` object

TASK 1: Request the Falcon9 Launch Wiki page from its URL

First, let's perform an HTTP GET method to request the Falcon9 Launch HTML page, as an HTTP response.

In [6]: `# use requests.get() method with the provided static_url`
`# assign the response to a object`
`response = requests.get(static_url)`

Create a `BeautifulSoup` object from the HTML `response`

In [7]: `# Use BeautifulSoup() to create a BeautifulSoup object from a response text cont`
`soup = BeautifulSoup(response.content, 'html.parser')`

Print the page title to verify if the `BeautifulSoup` object was created properly

```
In [8]: # Use soup.title attribute
soup.title
```

```
Out[8]: <title>List of Falcon 9 and Falcon Heavy launches - Wikipedia</title>
```

TASK 2: Extract all column/variable names from the HTML table header

Next, we want to collect all relevant column names from the HTML table header

Let's try to find all tables on the wiki page first. If you need to refresh your memory about BeautifulSoup, please check the external reference link towards the end of this lab

```
In [9]: # Use the find_all function in the BeautifulSoup object, with element type `table`
# Assign the result to a list called `html_tables`

html_tables = soup.find_all('table')
```

Starting from the third table is our target table contains the actual launch records.

```
In [10]: # Let's print the third table and check its content
first_launch_table = html_tables[2]
print(first_launch_table)
```

```

<table class="wikitable plainrowheaders collapsible" style="width: 100%;">
<tbody><tr>
<th scope="col">Flight No.
</th>
<th scope="col">Date and<br/>time (<a href="/wiki/Coordinated_Universal_Time" tit
le="Coordinated Universal Time">UTC</a>)
</th>
<th scope="col"><a href="/wiki/List_of_Falcon_9_first-stage_boosters" title="List
of Falcon 9 first-stage boosters">Version,<br/>Booster</a> <sup class="reference"
id="cite_ref-booster_11-0"><a href="#cite_note-booster-11">[b]</a></sup>
</th>
<th scope="col">Launch site
</th>
<th scope="col">Payload<sup class="reference" id="cite_ref-Dragon_12-0"><a href
="#cite_note-Dragon-12">[c]</a></sup>
</th>
<th scope="col">Payload mass
</th>
<th scope="col">Orbit
</th>
<th scope="col">Customer
</th>
<th scope="col">Launch<br/>outcome
</th>
<th scope="col"><a href="/wiki/Falcon_9_first-stage_landing_tests" title="Falcon
9 first-stage landing tests">Booster<br/>landing</a>
</th></tr>
<tr>
<th rowspan="2" scope="row" style="text-align:center;">1
</th>
<td>4 June 2010,<br/>18:45
</td>
<td><a href="/wiki/Falcon_9_v1.0" title="Falcon 9 v1.0">F9 v1.0</a><sup class="re
ference" id="cite_ref-MuskMay2012_13-0"><a href="#cite_note-MuskMay2012-13">[7]</
a></sup><br/>B0003.1<sup class="reference" id="cite_ref-block_numbers_14-0"><a hr
ef="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Canaveral Spac
e Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_Launch_Comple
x_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/Dragon_Spacecraft_Qualification_Unit" title="Dragon Spacecraft
Qualification Unit">Dragon Spacecraft Qualification Unit</a>
</td>
<td>
</td>
<td>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
</td>
<td><a href="/wiki/SpaceX" title="SpaceX">SpaceX</a>
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; tex
t-align: center;">Success
</td>
<td class="table-failure" style="background: #FFC7C7; vertical-align: middle; tex
t-align: center;">Failure<sup class="reference" id="cite_ref-ns20110930_15-0"><a
href="#cite_note-ns20110930-15">[9]</a></sup><sup class="reference" id="cite_ref-
16"><a href="#cite_note-16">[10]</a></sup><br/><small>(parachute)</small>
</td></tr>
<tr>
<td colspan="9">First flight of Falcon 9 v1.0.<sup class="reference" id="cite_ref

```

```

-sfn20100604_17-0"><a href="#cite_note-sfn20100604-17">[11]</a></sup> Used a boil
erplate version of Dragon capsule which was not designed to separate from the sec
ond stage.<small>(<a href="#First_flight_of_Falcon_9">more details below</a></sm
all> Attempted to recover the first stage by parachuting it into the ocean, but i
t burned up on reentry, before the parachutes even deployed.<sup class="referenc
e" id="cite_ref-parachute_18-0"><a href="#cite_note-parachute-18">[12]</a></sup>
</td></tr>
<tr>
<th rowspan="2" scope="row" style="text-align:center;">2
</th>
<td>8 December 2010,<br/>15:43<sup class="reference" id="cite_ref-spaceflightnow_
Clark_Launch_Report_19-0"><a href="#cite_note-spaceflightnow_Clark_Launch_Report-
19">[13]</a></sup>
</td>
<td><a href="/wiki/Falcon_9_v1.0" title="Falcon 9 v1.0">F9 v1.0</a><sup class="re
ference" id="cite_ref-MuskMay2012_13-1"><a href="#cite_note-MuskMay2012-13">[7]</
a></sup><br/>B0004.1<sup class="reference" id="cite_ref-block_numbers_14-1"><a hr
ef="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Canaveral Spac
e Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_Launch_Comple
x_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/SpaceX_Dragon" title="SpaceX Dragon">Dragon</a> <a class="mw-r
edirect" href="/wiki/COTS_Demo_Flight_1" title="COTS Demo Flight 1">demo flight C
1</a><br/>(Dragon C101)
</td>
<td>
</td>
<td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a> (<a href="/wi
ki/International_Space_Station" title="International Space Station">ISS</a>)
</td>
<td><style data-mw-deduplicate="TemplateStyles:r1126788409">.mw-parser-output .pl
ainlist ol,.mw-parser-output .plainlist ul{line-height:inherit;list-style:none;ma
rgin:0;padding:0}.mw-parser-output .plainlist ol li,.mw-parser-output .plainlist
ul li{margin-bottom:0}</style><div class="plainlist">
<ul><li><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Commercial_Orb
ital_Transportation_Services" title="Commercial Orbital Transportation Services">
COTS</a></li>
<li><a href="/wiki/National_Reconnaissance_Office" title="National Reconnaissance
Office">NRO</a></li></ul>
</div>
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; tex
t-align: center;">Success<sup class="reference" id="cite_ref-ns20110930_15-1"><a
href="#cite_note-ns20110930-15">[9]</a></sup>
</td>
<td class="table-failure" style="background: #FFC7C7; vertical-align: middle; tex
t-align: center;">Failure<sup class="reference" id="cite_ref-ns20110930_15-2"><a
href="#cite_note-ns20110930-15">[9]</a></sup><sup class="reference" id="cite_ref-
20"><a href="#cite_note-20">[14]</a></sup><br/><small>(parachute)</small>
</td></tr>
<tr>
<td colspan="9">Maiden flight of <a class="mw-redirect" href="/wiki/Dragon_capsul
e" title="Dragon capsule">Dragon capsule</a>, consisting of over 3 hours of testi
ng thruster maneuvering and reentry.<sup class="reference" id="cite_ref-spaceflig
htnow_Clark_unleashing_Dragon_21-0"><a href="#cite_note-spaceflightnow_Clark_unle
ashing_Dragon-21">[15]</a></sup> Attempted to recover the first stage by parachut
ing it into the ocean, but it disintegrated upon reentry, before the parachutes w
ere deployed.<sup class="reference" id="cite_ref-parachute_18-1"><a href="#cite_n

```

```

ote-parachute-18">[12]</a></sup> <small>(<a href="#COTS_demo_missions">more details below</a>)</small> It also included two <a href="/wiki/CubeSat" title="CubeSat">CubeSats</a>,<sup class="reference" id="cite_ref-NRO_Taps_Boeing_for_Next_Batch_of_CubeSats_22-0"><a href="#cite_note-NRO_Taps_Boeing_for_Next_Batch_of_CubeSats_22">[16]</a></sup> and a wheel of <a href="/wiki/Brou%C3%A8re" title="Brouère">Brouère</a> cheese.
</td></tr>
<tr>
<th rowspan="2" scope="row" style="text-align:center;">3
</th>
<td>22 May 2012,<br/>07:44<sup class="reference" id="cite_ref-BBC_new_era_23-0"><a href="#cite_note-BBC_new_era-23">[17]</a></sup>
</td>
<td><a href="/wiki/Falcon_9_v1.0" title="Falcon 9 v1.0">F9 v1.0</a><sup class="reference" id="cite_ref-MuskMay2012_13-2"><a href="#cite_note-MuskMay2012-13">[7]</a></sup><br/>B0005.1<sup class="reference" id="cite_ref-block_numbers_14-2"><a href="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Canaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_Launch_Complex_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/SpaceX_Dragon" title="SpaceX Dragon">Dragon</a> <a class="mw-redirect" href="/wiki/Dragon_C2%2B" title="Dragon C2+">demo flight C2+</a><sup class="reference" id="cite_ref-C2_24-0"><a href="#cite_note-C2-24">[18]</a></sup><br/>(Dragon C102)
</td>
<td>525 kg (1,157 lb)<sup class="reference" id="cite_ref-25"><a href="#cite_note-25">[19]</a></sup>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a> (<a href="/wiki/International_Space_Station" title="International Space Station">ISS</a>)
</td>
<td><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Commercial_Orbital_Transportation_Services" title="Commercial Orbital Transportation Services">COTS</a>)
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; text-align: center;">Success<sup class="reference" id="cite_ref-26"><a href="#cite_note-26">[20]</a></sup>
</td>
<td class="table-noAttempt" style="background: #EEE; vertical-align: middle; white-space: nowrap; text-align: center;">No attempt
</td></tr>
<tr>
<td colspan="9">Dragon spacecraft demonstrated a series of tests before it was allowed to approach the <a href="/wiki/International_Space_Station" title="International Space Station">International Space Station</a>. Two days later, it became the first commercial spacecraft to board the ISS.<sup class="reference" id="cite_ref-BBC_new_era_23-1"><a href="#cite_note-BBC_new_era-23">[17]</a></sup> <small>(<a href="#COTS_demo_missions">more details below</a>)</small>
</td></tr>
<tr>
<th rowspan="3" scope="row" style="text-align:center;">4
</th>
<td rowspan="2">8 October 2012,<br/>00:35<sup class="reference" id="cite_ref-SFN_LLog_27-0"><a href="#cite_note-SFN_LLog-27">[21]</a></sup>
</td>
<td rowspan="2"><a href="/wiki/Falcon_9_v1.0" title="Falcon 9 v1.0">F9 v1.0</a><sup class="reference" id="cite_ref-MuskMay2012_13-3"><a href="#cite_note-MuskMay20

```

```

12-13">[7]</a></sup><br/>B0006.1<sup class="reference" id="cite_ref-block_numbers
_14-3"><a href="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td rowspan="2"><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Ca
naveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_L
aunch_Complex_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/SpaceX_CRS-1" title="SpaceX CRS-1">SpaceX CRS-1</a><sup class
="reference" id="cite_ref-sxManifest20120925_28-0"><a href="#cite_note-sxManifest
20120925-28">[22]</a></sup><br/>(Dragon C103)
</td>
<td>4,700 kg (10,400 lb)
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a> (<a href="/wi
ki/International_Space_Station" title="International Space Station">ISS</a>)
</td>
<td><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Commercial_Resuppl
y_Services" title="Commercial Resupply Services">CRS</a>)
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; tex
t-align: center;">Success
</td>
<td rowspan="2" style="background:#ecec; text-align:center;"><span class="nowra
p">No attempt</span>
</td></tr>
<tr>
<td><a href="/wiki/Orbcomm_(satellite)" title="Orbcomm (satellite)">Orbcomm-OG2</
a><sup class="reference" id="cite_ref-Orbcomm_29-0"><a href="#cite_note-Orbcomm-2
9">[23]</a></sup>
</td>
<td>172 kg (379 lb)<sup class="reference" id="cite_ref-gunter-og2_30-0"><a href
="#cite_note-gunter-og2-30">[24]</a></sup>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
</td>
<td><a href="/wiki/Orbcomm" title="Orbcomm">Orbcomm</a>
</td>
<td class="table-partial" style="background: #FE9; vertical-align: middle; text-a
lign: center;">Partial failure<sup class="reference" id="cite_ref-nyt-20121030_31
-0"><a href="#cite_note-nyt-20121030-31">[25]</a></sup>
</td></tr>
<tr>
<td colspan="9">CRS-1 was successful, but the <a href="/wiki/Secondary_payload" t
itle="Secondary payload">secondary payload</a> was inserted into an abnormally lo
w orbit and subsequently lost. This was due to one of the nine <a href="/wiki/Spa
ceX_Merlin" title="SpaceX Merlin">Merlin engines</a> shutting down during the lau
nch, and NASA declining a second reignition, as per <a href="/wiki/International_
Space_Station" title="International Space Station">ISS</a> visiting vehicle safet
y rules, the primary payload owner is contractually allowed to decline a second r
eignition. NASA stated that this was because SpaceX could not guarantee a high en
ough likelihood of the second stage completing the second burn successfully which
was required to avoid any risk of secondary payload's collision with the ISS.<sup
class="reference" id="cite_ref-OrbcommTotalLoss_32-0"><a href="#cite_note-Orbcomm
TotalLoss-32">[26]</a></sup><sup class="reference" id="cite_ref-sn20121011_33-0">
<a href="#cite_note-sn20121011-33">[27]</a></sup><sup class="reference" id="cite_
ref-34"><a href="#cite_note-34">[28]</a></sup>
</td></tr>
<tr>
<th rowspan="2" style="text-align:center;">5
</th>

```



```

<td>1 March 2013,<br/>15:10
</td>
<td><a href="/wiki/Falcon_9_v1.0" title="Falcon 9 v1.0">F9 v1.0</a><sup class="reference" id="cite_ref-MuskMay2012_13-4"><a href="#cite_note-MuskMay2012-13">[7]</a></sup><br/>B0007.1<sup class="reference" id="cite_ref-block_numbers_14-4"><a href="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Canaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_Launch_Complex_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/SpaceX_CRS-2" title="SpaceX CRS-2">SpaceX CRS-2</a><sup class="reference" id="cite_ref-sxManifest20120925_28-1"><a href="#cite_note-sxManifest20120925-28">[22]</a></sup><br/>(Dragon C104)
</td>
<td>4,877 kg (10,752 lb)
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a> (<a class="mw-redirect" href="/wiki/ISS" title="ISS">ISS</a>)
</td>
<td><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Commercial_Resupply_Services" title="Commercial Resupply Services">CRS</a>)
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; text-align: center;">Success
</td>
<td class="table-noAttempt" style="background: #EEE; vertical-align: middle; white-space: nowrap; text-align: center;">No attempt
</td></tr>
<tr>
<td colspan="9">Last launch of the original Falcon 9 v1.0 <a href="/wiki/Launch_vehicle" title="Launch vehicle">launch vehicle</a>, first use of the unpressurized trunk section of Dragon.<sup class="reference" id="cite_ref-sxf9_20110321_35-0"><a href="#cite_note-sxf9_20110321-35">[29]</a></sup>
</td></tr>
<tr>
<th rowspan="2" scope="row" style="text-align:center;">6
</th>
<td>29 September 2013,<br/>16:00<sup class="reference" id="cite_ref-pa20130930_36-0"><a href="#cite_note-pa20130930-36">[30]</a></sup>
</td>
<td><a href="/wiki/Falcon_9_v1.1" title="Falcon 9 v1.1">F9 v1.1</a><sup class="reference" id="cite_ref-MuskMay2012_13-5"><a href="#cite_note-MuskMay2012-13">[7]</a></sup><br/>B1003<sup class="reference" id="cite_ref-block_numbers_14-5"><a href="#cite_note-block_numbers-14">[8]</a></sup>
</td>
<td><a class="mw-redirect" href="/wiki/Vandenberg_Air_Force_Base" title="Vandenberg Air Force Base">VAFB</a>,<br/><a href="/wiki/Vandenberg_Space_Launch_Complex_4" title="Vandenberg Space Launch Complex 4">SLC-4E</a>
</td>
<td><a href="/wiki/CASSIOPE" title="CASSIOPE">CASSIOPE</a><sup class="reference" id="cite_ref-sxManifest20120925_28-2"><a href="#cite_note-sxManifest20120925-28">[22]</a></sup><sup class="reference" id="cite_ref-CASSIOPE_MDA_37-0"><a href="#cite_note-CASSIOPE_MDA-37">[31]</a></sup>
</td>
<td>500 kg (1,100 lb)
</td>
<td><a href="/wiki/Polar_orbit" title="Polar orbit">Polar orbit</a> <a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
</td>

```

```

<td><a href="/wiki/Maxar_Technologies" title="Maxar Technologies">MDA</a>
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; text-align: center;">Success<sup class="reference" id="cite_ref-pa20130930_36-1"><a href="#cite_note-pa20130930-36">[30]</a></sup>
</td>
<td class="table-no2" style="background: #FFE3E3; color: black; vertical-align: middle; text-align: center;">Uncontrolled<br/><small>(ocean)</small><sup class="reference" id="cite_ref-ocean_landing_38-0"><a href="#cite_note-ocean_landing-38">[d]</a></sup>
</td></tr>
<tr>
<td colspan="9">First commercial mission with a private customer, first launch from Vandenberg, and demonstration flight of Falcon 9 v1.1 with an improved 13-tonne to LEO capacity.<sup class="reference" id="cite_ref-sxf9_20110321_35-1"><a href="#cite_note-sxf9_20110321-35">[29]</a></sup> After separation from the second stage carrying Canadian commercial and scientific satellites, the first stage booster performed a controlled reentry,<sup class="reference" id="cite_ref-39"><a href="#cite_note-39">[32]</a></sup> and an <a href="/wiki/Falcon_9_first-stage_landing_tests" title="Falcon 9 first-stage landing tests">ocean touchdown test</a> for the first time. This provided good test data, even though the booster started rolling as it neared the ocean, leading to the shutdown of the central engine as the roll depleted it of fuel, resulting in a hard impact with the ocean.<sup class="reference" id="cite_ref-pa20130930_36-2"><a href="#cite_note-pa20130930-36">[30]</a></sup> This was the first known attempt of a rocket engine being lit to perform a supersonic retro propulsion, and allowed SpaceX to enter a public-private partnership with <a href="/wiki/NASA" title="NASA">NASA</a> and its Mars entry, descent, and landing technologies research projects.<sup class="reference" id="cite_ref-40"><a href="#cite_note-40">[33]</a></sup> <small>( <a href="#Maiden_flight_of_v1.1">more details below</a>)</small>
</td></tr>
<tr>
<th rowspan="2" scope="row" style="text-align:center;">7
</th>
<td>3 December 2013,<br/>22:41<sup class="reference" id="cite_ref-sfn_wwls20130624_41-0"><a href="#cite_note-sfn_wwls20130624-41">[34]</a></sup>
</td>
<td><a href="/wiki/Falcon_9_v1.1" title="Falcon 9 v1.1">F9 v1.1</a><br/>B1004
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape Canaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Canaveral_Space_Launch_Complex_40" title="Cape Canaveral Space Launch Complex 40">SLC-40</a>
</td>
<td><a href="/wiki/SES-8" title="SES-8">SES-8</a><sup class="reference" id="cite_ref-sxManifest20120925_28-3"><a href="#cite_note-sxManifest20120925-28">[22]</a></sup><sup class="reference" id="cite_ref-spx-pr_42-0"><a href="#cite_note-spx-pr-42">[35]</a></sup><sup class="reference" id="cite_ref-aw20110323_43-0"><a href="#cite_note-aw20110323-43">[36]</a></sup>
</td>
<td>3,170 kg (6,990 lb)
</td>
<td><a href="/wiki/Geostationary_transfer_orbit" title="Geostationary transfer orbit">GTO</a>
</td>
<td><a class="mw-redirect" href="/wiki/SES_S.A." title="SES S.A.">SES</a>
</td>
<td class="table-success" style="background: #9EFF9E; vertical-align: middle; text-align: center;">Success<sup class="reference" id="cite_ref-SNMissionStatus7_44-0"><a href="#cite_note-SNMissionStatus7-44">[37]</a></sup>
</td>

```

```

<td class="table-noAttempt" style="background: #EEE; vertical-align: middle; white-space: nowrap; text-align: center;">No attempt<br/><sup class="reference" id="cite_ref-sf10120131203_45-0"><a href="#cite_note-sf10120131203-45">[38]</a></sup>
</td></tr>
<tr>
<td colspan="9">First <a href="/wiki/Geostationary_transfer_orbit" title="Geostationary transfer orbit">Geostationary transfer orbit</a> (GTO) launch for Falcon 9,<sup class="reference" id="cite_ref-spx-pr_42-1"><a href="#cite_note-spx-pr-42">[35]</a></sup> and first successful reignition of the second stage.<sup class="reference" id="cite_ref-46"><a href="#cite_note-46">[39]</a></sup> SES-8 was inserted into a <a href="/wiki/Geostationary_transfer_orbit" title="Geostationary transfer orbit">Super-Synchronous Transfer Orbit</a> of 79,341 km (49,300 mi) in a poguee with an <a href="/wiki/Orbital_inclination" title="Orbital inclination">inclination</a> of 20.55° to the <a href="/wiki/Equator" title="Equator">equator</a>.
</td></tr></tbody></table>

```

You should be able to see the column names embedded in the table header elements

`<th>` as follows:

```

<tr>
<th scope="col">Flight No.
</th>
<th scope="col">Date and<br/>time (<a href="/wiki/Coordinated_Universal_Time" title="Coordinated Universal Time">UTC</a>)
</th>
<th scope="col"><a href="/wiki/List_of_Falcon_9_first-stage_boosters" title="List of Falcon 9 first-stage boosters">Version,<br/>Booster</a> <sup class="reference" id="cite_ref-booster_11-0"><a href="#cite_note-booster-11">[b]
</a></sup>
</th>
<th scope="col">Launch site
</th>
<th scope="col">Payload<sup class="reference" id="cite_ref-Dragon_12-0"><a href="#cite_note-Dragon-12">[c]</a></sup>
</th>
<th scope="col">Payload mass
</th>
<th scope="col">Orbit
</th>
<th scope="col">Customer
</th>
<th scope="col">Launch<br/>outcome
</th>
<th scope="col"><a href="/wiki/Falcon_9_first-stage_landing_tests" title="Falcon 9 first-stage landing tests">Booster<br/>landing</a>
</th></tr>

```

Next, we just need to iterate through the `<th>` elements and apply the provided `extract_column_from_header()` to extract column name one by one

```
In [11]: column_names = []

# Apply find_all() function with `th` element on first_launch_table
# Iterate each th element and apply the provided extract_column_from_header() to
# Append the Non-empty column name (`if name is not None and len(name) > 0`) int

procura = first_launch_table.find_all('th')
for th in procura:
    column = extract_column_from_header(th)
    column_names.append(column)
```

Check the extracted column names

```
In [12]: print(column_names)

['Flight No.', 'Date and time ( )', '', 'Launch site', 'Payload', 'Payload mass',
'Orbit', 'Customer', 'Launch outcome', '', None, None, None, None, None, No
ne]
```

TASK 3: Create a data frame by parsing the launch HTML tables

We will create an empty dictionary with keys from the extracted column names in the previous task. Later, this dictionary will be converted into a Pandas dataframe

```
In [13]: launch_dict= dict.fromkeys(column_names)

# Remove an irrelevant column
del launch_dict['Date and time ( )']

# Let's initial the launch_dict with each value to be an empty list
launch_dict['Flight No.'] = []
launch_dict['Launch site'] = []
launch_dict['Payload'] = []
launch_dict['Payload mass'] = []
launch_dict['Orbit'] = []
launch_dict['Customer'] = []
launch_dict['Launch outcome'] = []
# Added some new columns
launch_dict['Version Booster']=[]
launch_dict['Booster landing']=[]
launch_dict['Date']=[]
launch_dict['Time']=[]
```

Next, we just need to fill up the `launch_dict` with launch records extracted from table rows.

Usually, HTML tables in Wiki pages are likely to contain unexpected annotations and other types of noises, such as reference links `B0004.1[8]` , missing values `N/A [e]` , inconsistent formatting, etc.

To simplify the parsing process, we have provided an incomplete code snippet below to help you to fill up the `launch_dict`. Please complete the following code snippet with TODOs or you can choose to write your own logic to parse all launch tables:

```
In [15]: extracted_row = 0
#Extract each table
for table_number, table in enumerate(soup.find_all('table', "wikitable plainrowhea
# get table row
    for rows in table.find_all("tr"):
        #check to see if first table heading is as number corresponding to Launch
        if rows.th:
            if rows.th.string:
                flight_number=rows.th.string.strip()
                flag=flight_number.isdigit()
            else:
                flag=False
#get table element
row=rows.find_all('td')
#if it is number save cells in a dictionary
if flag:
    extracted_row += 1
    # Flight Number value
    # TODO: Append the flight_number into launch_dict with key `Flight N
    #print(flight_number)
    launch_dict['Flight No.'].append(flight_number)
    datatimelist=date_time(row[0])

    # Date value
    # TODO: Append the date into launch_dict with key `Date`
    date = datatimelist[0].strip(',')
    launch_dict['Date'].append(date)
    #print(date)

    # Time value
    # TODO: Append the time into launch_dict with key `Time`
    time = datatimelist[1]
    launch_dict['Time'].append(time)
    #print(time)

    # Booster version
    # TODO: Append the bv into launch_dict with key `Version Booster`
    bv=booster_version(row[1])
    launch_dict['Version Booster'].append(bv)
    if not(bv):
        bv=row[1].a.string
    #print(bv)

    # Launch Site
    # TODO: Append the bv into launch_dict with key `Launch Site`
    launch_site = row[2].a.string
    launch_dict['Launch site'].append(launch_site)
    #print(launch_site)

    # Payload
    # TODO: Append the payload into launch_dict with key `Payload`
    payload = row[3].a.string
    launch_dict['Payload'].append(payload)
    #print(payload)
```

```

# Payload Mass
# TODO: Append the payload_mass into launch_dict with key `Payload m
payload_mass = get_mass(row[4])
launch_dict['Payload mass'].append(payload_mass)
#print(payload)

# Orbit
# TODO: Append the orbit into launch_dict with key `Orbit`
orbit = row[5].a.string
launch_dict['Orbit'].append(orbit)
#print(orbit)

# Customer
# TODO: Append the customer into launch_dict with key `Customer`
customer = row[6]
if not (customer):
    customer = row[6].a.string
launch_dict['Customer'].append(customer)
#print(customer)

# Launch outcome
# TODO: Append the launch_outcome into launch_dict with key `Launch
launch_outcome = list(row[7].strings)[0]
launch_dict['Launch outcome'].append(launch_outcome)
#print(launch_outcome)

# Booster Landing
# TODO: Append the launch_outcome into launch_dict with key `Booster
booster_landing = landing_status(row[8])
launch_dict['Booster landing'].append(booster_landing)
#print(booster_Landing)

```

After you have fill in the parsed launch record values into `launch_dict` , you can create a dataframe from it.

```
In [16]: df= pd.DataFrame({ key:pd.Series(value) for key, value in launch_dict.items() })
```

/home/jupyterlab/conda/envs/python/lib/python3.7/site-packages/ipykernel_launcher.py:1: DeprecationWarning: The default dtype for empty Series will be 'object' instead of 'float64' in a future version. Specify a dtype explicitly to silence this warning.

"""Entry point for launching an IPython kernel.

```
In [19]: import requests
import pandas as pd

# Perform the GET request to fetch launches data
url = "https://api.spacexdata.com/v4/launches"
response = requests.get(url)
data = response.json()

df = pd.json_normalize(data)

if 'rocket' in df.columns:

    falcon_9_df = df[df['rocket'].str.contains("Falcon 9", na=False)]

```

```
count_falcon_9 = len(falcon_9_df)
else:
    count_falcon_9 = "Rocket data column missing"

print("Number of Falcon 9 launches:", count_falcon_9)
```

Number of Falcon 9 launches: 0

We can now export it to a **CSV** for the next section, but to make the answers consistent and in case you have difficulties finishing this lab.

Following labs will be using a provided dataset to make each lab independent.

```
df.to_csv('spacex_web_scraped.csv', index=False)
```

Authors

[Yan Luo](#)

[Nayef Abou Tayoun](#)

Change Log

Date (YYYY-MM-DD)	Version	Changed By	Change Description
2021-06-09	1.0	Yan Luo	Tasks updates
2020-11-10	1.0	Nayef	Created the initial version

Copyright © 2021 IBM Corporation. All rights reserved.