



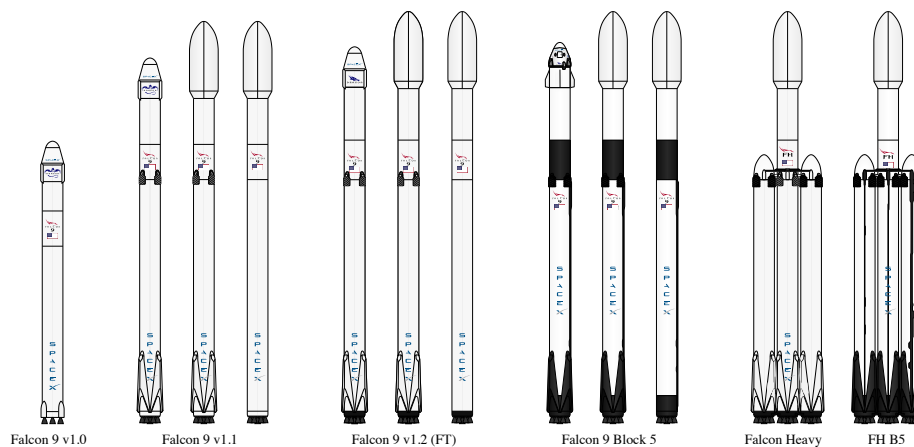
Space X Falcon 9 First Stage Landing Prediction

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:

2020 [edit]

In late 2019, [Gwynne Shownell](#) stated that SpaceX hoped for as many as 24 launches for Starlink satellites in 2020,^[490] in addition to 14 or 15 non-Starlink launches. At 26 launches, 13 of which for Starlink satellites, Falcon 9 had its most prolific year, and Falcon rockets were second most prolific rocket family of 2020, only behind China's [Long March](#) rocket family.^[491]

[hide] Flight No.	Date and time (UTC)	Version, Booster ^[6]	Launch site	Payload ^[2]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
76	7 January 2020, 02:19:21 ^[492]	F9 B5 Δ, B1049.4	CCAFS, SLC-40	Starlink 2 v1.0 (60 satellites)	15,800 kg (34,400 lb) ^[9]	LEO	SpaceX	Success	Success (drone ship)
Third large batch and second operational flight of Starlink constellation. One of the 60 satellites included a test coating to make the satellite less reflective, and thus less likely to interfere with ground-based astronomical observations. ^[493]									
79	19 January 2020, 15:30 ^[494]	F9 B5 Δ, B1046.4	KSC, LC-39A	Crew Dragon in-flight abort test ^[495] (Dragon C205.1)	12,050 kg (26,570 lb)	Sub-orbital ^[496]	NASA (CTS) ^[497]	Success	No attempt
An atmospheric test of the Dragon 2 abort system after Max Q. The capsule fired its SuperDraco engines, reached an apogee of 40 km (25 mi), deployed parachutes after reentry, and splashed down in the ocean 31 km (19 mi) downrange from the launch site. The test was previously slated to be accomplished with the Crew Dragon Demo-1 capsule, ^[498] but that test article exploded during a ground test of SuperDraco engines on 20 April 2019. ^[419] The abort test used the capsule originally intended for the first crewed flight. ^[499] As expected, the booster was destroyed by aerodynamic forces after the capsule aborted. ^[500] First flight of a Falcon 9 with only one functional stage — the second stage had a mass simulator in place of its engine.									
80	29 January 2020, 14:07 ^[501]	F9 B5 Δ, B1051.3	CCAFS, SLC-40	Starlink 3 v1.0 (60 satellites)	15,800 kg (34,400 lb) ^[9]	LEO	SpaceX	Success	Success (drone ship)
Third operational and fourth large batch of Starlink satellites, deployed in a circular 290 km (180 mi) orbit. One of the fairing halves was caught, while the other was fished out of the ocean. ^[502]									
81	17 February 2020, 15:09 ^[503]	F9 B5 Δ, B1056.4	CCAFS, SLC-40	Starlink 4 v1.0 (60 satellites)	15,800 kg (34,400 lb) ^[9]	LEO	SpaceX	Success	Failure (drone ship)
Fourth operational and fifth large batch of Starlink satellites. Used a new flight profile which deployed into a 212 km × 386 km (132 mi × 240 mi) elliptical orbit instead of launching into a circular orbit and firing the second stage engine twice. The first stage booster failed to land on the drone ship ^[504] due to incorrect wind data. ^[505] This was the first time a flight proven booster failed to land.									
82	7 March 2020, 04:50 ^[506]	F9 B5 Δ, B1059.2	CCAFS, SLC-40	SpaceX CRS-20 (Dragon C112.3 Δ)	1,977 kg (4,359 lb) ^[507]	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
Last launch of phase 1 of the CRS contract. Carries <i>Bartolomeo</i> , an ESA platform for hosting external payloads onto ISS. ^[508] Originally scheduled to launch on 2 March 2020, the launch date was pushed back due to a second stage engine failure. SpaceX decided to swap out the second stage instead of replacing the faulty part. ^[509] It was SpaceX's 50th successful landing of a first stage booster, the third flight of the Dragon C112 and the last launch of the cargo Dragon spacecraft.									
83	18 March 2020, 12:16 ^[510]	F9 B5 Δ, B1048.5	KSC, LC-39A	Starlink 5 v1.0 (60 satellites)	15,800 kg (34,400 lb) ^[9]	LEO	SpaceX	Success	Failure (drone ship)
Fifth operational launch of Starlink satellites. It was the first time a first stage booster flew for a fifth time and the second time the fairings were reused (Starlink flight in May 2019). ^[511] Towards the end of the first stage burn, the booster suffered premature shut down of an engine, the first of a <i>Merlin</i> 1D variant and first since the CRS-1 mission in October 2012. However, the payload still reached the targeted orbit. ^[512] This was the second Starlink launch booster landing failure in a row, later revealed to be caused by residual cleaning fluid trapped inside a sensor. ^[513]									
84	22 April 2020, 19:30 ^[514]	F9 B5 Δ, B1051.4	KSC, LC-39A	Starlink 6 v1.0 (60 satellites)	15,800 kg (34,400 lb) ^[9]	LEO	SpaceX	Success	Success (drone ship)

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 [38]: !pip3 install beautifulsoup4
!pip3 install requests
```

Requirement already satisfied: beautifulsoup4 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (4.13.3)

Requirement already satisfied: soupsieve>1.2 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from beautifulsoup4) (2.6)

Requirement already satisfied: typing-extensions>=4.0.0 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from beautifulsoup4) (4.12.2)

Requirement already satisfied: requests in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (2.32.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from requests) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from requests) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from requests) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from requests) (2025.1.31)

```
In [39]: import sys

import requests
from bs4 import BeautifulSoup
import re
import unicodedata
import pandas as pd
```

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

```
In [40]: 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.strip().split(' '))]

def booster_version(table_cells):
        """
```

```

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

def landing_status(table_cells):
"""
This function returns the landing status from the HTML table ce
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 ce
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 [41]: static_url = "https://en.wikipedia.org/w/index.php?title=List_of_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 [42]: # 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 [43]: # Use BeautifulSoup() to create a BeautifulSoup object from a respo
soup = BeautifulSoup(response.text, "html.parser")
```

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

```
In [44]: # Use soup.title attribute
print(soup.title)
```

```
<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 [45]: # Use the find_all function in the BeautifulSoup object, with eleme
# 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 [46]: # 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_Univer
sal_Time" title="Coordinated Universal Time">UTC</a>)
</th>
<th scope="col"><a href="/wiki/List_of_Falcon_9_first-stage_booster
s" title="List of Falcon 9 first-stage boosters">Version,<br/>Booste
r</a> <sup class="reference" id="cite_ref-booster_11-0"><a href="#ci
te_note-booster-11"><span class="cite-bracket">[</span>b<span class
="cite-bracket">]</span></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"><span class="cite-bracket">[</spa
n>c<span class="cite-bracket">]</span></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" t
itle="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="reference" id="cite_ref-MuskMay2012_13-0"><a href="#cite_
note-MuskMay2012-13"><span class="cite-bracket">[</span>7<span class
="cite-bracket">]</span></a></sup><br/>B0003.1<sup class="reference"
id="cite_ref-block_numbers_14-0"><a href="#cite_note-block_numbers-1
4"><span class="cite-bracket">[</span>8<span class="cite-bracket">]
</span></a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape C
anaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Can
averal_Space_Launch_Complex_40" title="Cape Canaveral Space Launch C
omplex 40">SLC-40</a>
</td>
<td><a href="/wiki/Dragon_Spacecraft_Qualification_Unit" title="Drag
on Spacecraft Qualification Unit">Dragon Spacecraft Qualification Un
it</a>
</td>
<td>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LE0</a>
</td>
<td><a href="/wiki/SpaceX" title="SpaceX">SpaceX</a>
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success

```

```

</td>
<td class="table-failure" style="background: #FFC7C7; color:black; v
ertical-align: middle; text-align: center;">Failure<sup class="refer
ence" id="cite_ref-ns20110930_15-0"><a href="#cite_note-ns20110930-1
5"><span class="cite-bracket">[</span>9<span class="cite-bracket">]
</span></a></sup><sup class="reference" id="cite_ref-16"><a href="#c
ite_note-16"><span class="cite-bracket">[</span>10<span class="cite-
bracket">]</span></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"><
span class="cite-bracket">[</span>11<span class="cite-bracket">]</sp
an></a></sup> Used a boilerplate version of Dragon capsule which was
not designed to separate from the second stage.<small>(<a href="#Fir
st_flight_of_Falcon_9">more details below</a>)</small> Attempted to
recover the first stage by parachuting it into the ocean, but it bur
ned up on reentry, before the parachutes even deployed.<sup class="r
eference" id="cite_ref-parachute_18-0"><a href="#cite_note-parachute
-18"><span class="cite-bracket">[</span>12<span class="cite-bracket"
">]</span></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-sp
aceflightnow_Clark_Launch_Report_19-0"><a href="#cite_note-spaceflig
htnow_Clark_Launch_Report-19"><span class="cite-bracket">[</span>13<
span class="cite-bracket">]</span></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_13-1"><a href="#cite_
note-MuskMay2012-13"><span class="cite-bracket">[</span>7<span class
="cite-bracket">]</span></a></sup><br/>B0004.1<sup class="reference"
id="cite_ref-block_numbers_14-1"><a href="#cite_note-block_numbers-1
4"><span class="cite-bracket">[</span>8<span class="cite-bracket">]
</span></a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape C
anaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Can
averal_Space_Launch_Complex_40" title="Cape Canaveral Space Launch C
omplex 40">SLC-40</a>
</td>
<td><a href="/wiki/SpaceX_Dragon" title="SpaceX Dragon">Dragon</a> <
a class="mw-redirect" href="/wiki/COTS_Demo_Flight_1" title="COTS De
mo Flight 1">demo flight C1</a><br/>(Dragon C101)
</td>
<td>
</td>
<td>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
(<a href="/wiki/International_Space_Station" title="International Sp
ace Station">ISS</a>)
</td>
<td><style data-mw-deduplicate="TemplateStyles:r1126788409">.mw-pars
er-output .plainlist ol,.mw-parser-output .plainlist ul{line-height:
inherit;list-style:none;margin:0;padding:0}.mw-parser-output .plainl

```

```

ist 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/C
ommercial_Orbital_Transportation_Services" title="Commercial Orbital
Transportation Services">COTS</a>)</li>
<li><a href="/wiki/National_Reconnaissance_Office" title="National R
econnaissance Office">NRO</a></li></ul>
</div>
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success<sup class="refer
ence" id="cite_ref-ns20110930_15-1"><a href="#cite_note-ns20110930-1
5"><span class="cite-bracket">[</span>9<span class="cite-bracket">]
</span></a></sup>
</td>
<td class="table-failure" style="background: #FFC7C7; color:black; v
ertical-align: middle; text-align: center;">Failure<sup class="refer
ence" id="cite_ref-ns20110930_15-2"><a href="#cite_note-ns20110930-1
5"><span class="cite-bracket">[</span>9<span class="cite-bracket">]
</span></a></sup><sup class="reference" id="cite_ref-20"><a href="#c
ite_note-20"><span class="cite-bracket">[</span>14<span class="cite-
bracket">]</span></a></sup><br><small>(parachute)</small>
</td></tr>
<tr>
<td colspan="9">Maiden flight of <a class="mw-redirect" href="/wiki/
Dragon_capsule" title="Dragon capsule">Dragon capsule</a>, consistin
g of over 3 hours of testing thruster maneuvering and reentry.<sup c
lass="reference" id="cite_ref-spaceflightnow_Clark_unleashing_Dragon
_21-0"><a href="#cite_note-spaceflightnow_Clark_unleashing_Dragon-2
1"><span class="cite-bracket">[</span>15<span class="cite-bracket">]
</span></a></sup> Attempted to recover the first stage by parachutin
g it into the ocean, but it disintegrated upon reentry, before the p
arachutes were deployed.<sup class="reference" id="cite_ref-parachut
e_18-1"><a href="#cite_note-parachute-18"><span class="cite-bracket"
>[</span>12<span class="cite-bracket">]</span></a></sup> <small>(<a
href="#COTS_demo_missions">more details below</a>)</small> It also i
ncluded two <a href="/wiki/CubeSat" title="CubeSat">CubeSats</a>,<su
p class="reference" id="cite_ref-NRO_Taps_Boeing_for_Next_Batch_of_C
ubeSats_22-0"><a href="#cite_note-NRO_Taps_Boeing_for_Next_Batch_of_
CubeSats-22"><span class="cite-bracket">[</span>16<span class="cite-
bracket">]</span></a></sup> and a wheel of <a href="/wiki/Brou%C3%A8
re" 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_ne
w_era_23-0"><a href="#cite_note-BBC_new_era-23"><span class="cite-br
acket">[</span>17<span class="cite-bracket">]</span></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"><span class="cite-bracket">[</span>7<span class
="cite-bracket">]</span></a></sup><br>B0005.1<sup class="reference"
id="cite_ref-block_numbers_14-2"><a href="#cite_note-block_numbers-1
4"><span class="cite-bracket">[</span>8<span class="cite-bracket">]

```



```

</span></a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape C
anaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Can
averal_Space_Launch_Complex_40" title="Cape Canaveral Space Launch C
omplex 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+">d
emo flight C2+</a><sup class="reference" id="cite_ref-C2_24-0"><a hr
ef="#cite_note-C2-24"><span class="cite-bracket">[</span>18<span cla
ss="cite-bracket">]</span></a></sup><br/>(Dragon C102)
</td>
<td>525 kg (1,157 lb)<sup class="reference" id="cite_ref-25"><a href
="#cite_note-25"><span class="cite-bracket">[</span>19<span class="c
ite-bracket">]</span></a></sup>
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
(<a href="/wiki/International_Space_Station" title="International Sp
ace Station">ISS</a>)
</td>
<td><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Comme
rcial_Orbital_Transportation_Services" title="Commercial Orbital Tra
nsportation Services">COTS</a>)
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success<sup class="refer
ence" id="cite_ref-26"><a href="#cite_note-26"><span class="cite-bra
cket">[</span>20<span class="cite-bracket">]</span></a></sup>
</td>
<td class="table-noAttempt" style="background: #EEE; color:black; ve
rtical-align: middle; white-space: nowrap; text-align: center;">No a
ttempt
</td></tr>
<tr>
<td colspan="9">Dragon spacecraft demonstrated a series of tests bef
ore it was allowed to approach the <a href="/wiki/International_Spac
e_Station" title="International Space Station">International Space S
tation</a>. Two days later, it became the first commercial spacecraf
t to board the ISS.<sup class="reference" id="cite_ref-BBC_new_era_2
3-1"><a href="#cite_note-BBC_new_era-23"><span class="cite-bracket">[
</span>17<span class="cite-bracket">]</span></a></sup> <small>(<a h
ref="#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"><span class
="cite-bracket">[</span>21<span class="cite-bracket">]</span></a></s
up>
</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-MuskMay2012-13"><span class="cite-bracket">[</span>
7<span class="cite-bracket">]</span></a></sup><br/>B0006.1<sup class

```

```

="reference" id="cite_ref-block_numbers_14-3"><a href="#cite_note-bl
ock_numbers_14"><span class="cite-bracket">[</span>8<span class="cit
e-bracket">]</span></a></sup>
</td>
<td rowspan="2"><a href="/wiki/Cape_Canaveral_Space_Force_Station" t
itle="Cape Canaveral Space Force Station">CCAFS</a>,<br><a href="/w
iki/Cape_Canaveral_Space_Launch_Complex_40" title="Cape Canaveral Sp
ace 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 hr
ef="#cite_note-sxManifest20120925-28"><span class="cite-bracket">[</
span>22<span class="cite-bracket">]</span></a></sup><br>(Dragon C10
3)
</td>
<td>4,700 kg (10,400 lb)
</td>
<td><a href="/wiki/Low_Earth_orbit" title="Low Earth orbit">LEO</a>
(<a href="/wiki/International_Space_Station" title="International Sp
ace Station">ISS</a>)
</td>
<td><a href="/wiki/NASA" title="NASA">NASA</a> (<a href="/wiki/Comme
rcial_Resupply_Services" title="Commercial Resupply Services">CRS</a
>)
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success
</td>
<td rowspan="2" style="background:#ecec; text-align:center;"><span
class="nowrap">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-29"><span class="cite-bracket">[</span>23<s
pan class="cite-bracket">]</span></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"><span class="cite-bracket">[
</span>24<span class="cite-bracket">]</span></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: #FFB; color:black; vert
ical-align: middle; text-align: center;">Partial failure<sup class="
reference" id="cite_ref-nyt-20121030_31-0"><a href="#cite_note-nyt-2
0121030-31"><span class="cite-bracket">[</span>25<span class="cite-b
racket">]</span></a></sup>
</td></tr>
<tr>
<td colspan="9">CRS-1 was successful, but the <a href="/wiki/Seconda
ry_payload" title="Secondary payload">secondary payload</a> was inse
rted into an abnormally low orbit and subsequently lost. This was du
e to one of the nine <a href="/wiki/SpaceX_Merlin" title="SpaceX Mer

```

```

lin">Merlin engines</a> shutting down during the launch, and NASA de
clining a second reignition, as per <a href="/wiki/International_Spa
ce_Station" title="International Space Station">ISS</a> visiting veh
icle safety rules, the primary payload owner is contractually allowe
d to decline a second reignition. NASA stated that this was because
SpaceX could not guarantee a high enough likelihood of the second st
age completing the second burn successfully which was required to av
oid any risk of secondary payload's collision with the ISS.<sup clas
s="reference" id="cite_ref-OrbcommTotalLoss_32-0"><a href="#cite_not
e-OrbcommTotalLoss-32"><span class="cite-bracket">[</span>26<span cl
ass="cite-bracket">]</span></a></sup><sup class="reference" id="cite
_ref-sn20121011_33-0"><a href="#cite_note-sn20121011-33"><span class
="cite-bracket">[</span>27<span class="cite-bracket">]</span></a></s
up><sup class="reference" id="cite_ref-34"><a href="#cite_note-34"><
span class="cite-bracket">[</span>28<span class="cite-bracket">]</sp
an></a></sup>
</td></tr>
<tr>
<th rowspan="2" scope="row" 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"><span class="cite-bracket">[</span>7<span class
="cite-bracket">]</span></a></sup><br/>B0007.1<sup class="reference"
id="cite_ref-block_numbers_14-4"><a href="#cite_note-block_numbers-1
4"><span class="cite-bracket">[</span>8<span class="cite-bracket">]</
span></a></sup>
</td>
<td><a href="/wiki/Cape_Canaveral_Space_Force_Station" title="Cape C
anaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Can
averal_Space_Launch_Complex_40" title="Cape Canaveral Space Launch C
omplex 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 hr
ef="#cite_note-sxManifest20120925-28"><span class="cite-bracket">[</
span>22<span class="cite-bracket">]</span></a></sup><br/>(Dragon C10
4)
</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/Comme
rcial_Resupply_Services" title="Commercial Resupply Services">CRS</a
>)
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success
</td>
<td class="table-noAttempt" style="background: #EEE; color:black; ve
rtical-align: middle; white-space: nowrap; text-align: center;">No a
ttempt

```

```

</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>, fir
t use of the unpressurized trunk section of Dragon.<sup class="refer
ence" id="cite_ref-sxf9_20110321_35-0"><a href="#cite_note-sxf9_2011
0321-35"><span class="cite-bracket">[</span>29<span class="cite-brac
ket">]</span></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"><span class="cit
e-bracket">[</span>30<span class="cite-bracket">]</span></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"><span class="cite-bracket">[</span>7<span class
="cite-bracket">]</span></a></sup><br/>B1003<sup class="reference" i
d="cite_ref-block_numbers_14-5"><a href="#cite_note-block_numbers-1
4"><span class="cite-bracket">[</span>8<span class="cite-bracket">]
</span></a></sup>
</td>
<td><a class="mw-redirect" href="/wiki/Vandenberg_Air_Force_Base" ti
tle="Vandenberg Air Force Base">VAFB</a>,<br/><a href="/wiki/Vandenb
erg_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_no
te-sxManifest20120925-28"><span class="cite-bracket">[</span>22<span
class="cite-bracket">]</span></a></sup><sup class="reference" id="ci
te_ref-CASSIOPE_MDA_37-0"><a href="#cite_note-CASSIOPE_MDA-37"><span
class="cite-bracket">[</span>31<span class="cite-bracket">]</span></
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">MD
A</a>
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success<sup class="refer
ence" id="cite_ref-pa20130930_36-1"><a href="#cite_note-pa20130930-3
6"><span class="cite-bracket">[</span>30<span class="cite-bracket">]
</span></a></sup>
</td>
<td class="table-no2" style="background: #FFE3E3; color: black; vert
ical-align: middle; text-align: center;">Uncontrolled<br/><small>(oc
ean)</small><sup class="reference" id="cite_ref-ocean_landing_38-0">
<a href="#cite_note-ocean_landing-38"><span class="cite-bracket">[</
span>d<span class="cite-bracket">]</span></a></sup>

```

```

</td></tr>
<tr>
<td colspan="9">First commercial mission with a private customer, fi
rst launch from Vandenberg, and demonstration flight of Falcon 9 v1.
1 with an improved 13-tonne to LEO capacity.<sup class="reference" i
d="cite_ref-sxf9_20110321_35-1"><a href="#cite_note-sxf9_20110321-3
5"><span class="cite-bracket">[</span>29<span class="cite-bracket">]</span></a></sup> After separation from the second stage carrying Ca
nadian commercial and scientific satellites, the first stage booster
performed a controlled reentry,<sup class="reference" id="cite_ref-3
9"><a href="#cite_note-39"><span class="cite-bracket">[</span>32<spa
n class="cite-bracket">]</span></a></sup> and an <a href="/wiki/Falc
on_9_first-stage_landing_tests" title="Falcon 9 first-stage landing
tests">ocean touchdown test</a> for the first time. This provided go
od test data, even though the booster started rolling as it neared t
he 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"><span class="cite-bracket">[</span>30<span class="cit
e-bracket">]</span></a></sup> This was the first known attempt of a
rocket engine being lit to perform a supersonic retro propulsion, an
d 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="ci
te_ref-40"><a href="#cite_note-40"><span class="cite-bracket">[</spa
n>33<span class="cite-bracket">]</span></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-sf
n_wwls20130624_41-0"><a href="#cite_note-sfn_wwls20130624-41"><span
class="cite-bracket">[</span>34<span class="cite-bracket">]</span></
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 C
anaveral Space Force Station">CCAFS</a>,<br/><a href="/wiki/Cape_Can
averal_Space_Launch_Complex_40" title="Cape Canaveral Space Launch C
omplex 40">SLC-40</a>
</td>
<td><a href="/wiki/SES-8" title="SES-8">SES-8</a><sup class="referen
ce" id="cite_ref-sxManifest20120925_28-3"><a href="#cite_note-sxMani
fest20120925-28"><span class="cite-bracket">[</span>22<span class="c
ite-bracket">]</span></a></sup><sup class="reference" id="cite_ref-s
px-pr_42-0"><a href="#cite_note-spx-pr-42"><span class="cite-bracke
t">[</span>35<span class="cite-bracket">]</span></a></sup><sup class
="reference" id="cite_ref-aw20110323_43-0"><a href="#cite_note-aw201
10323-43"><span class="cite-bracket">[</span>36<span class="cite-bra
cket">]</span></a></sup>
</td>
<td>3,170 kg (6,990 lb)
</td>

```

```

<td><a href="/wiki/Geostationary_transfer_orbit" title="Geostationar
y transfer orbit">GT0</a>
</td>
<td><a class="mw-redirect" href="/wiki/SES_S.A." title="SES S.A.">SE
S</a>
</td>
<td class="table-success" style="background: #9EFF9E; color:black; v
ertical-align: middle; text-align: center;">Success<sup class="refer
ence" id="cite_ref-SNMissionStatus7_44-0"><a href="#cite_note-SNMiss
ionStatus7-44"><span class="cite-bracket">[</span>37<span class="cit
e-bracket">]</span></a></sup>
</td>
<td class="table-noAttempt" style="background: #EEE; color:black; ve
rtical-align: middle; white-space: nowrap; text-align: center;">No a
tttempt<br/><sup class="reference" id="cite_ref-sf10120131203_45-0"><
a href="#cite_note-sf10120131203-45"><span class="cite-bracket">[</s
pan>38<span class="cite-bracket">]</span></a></sup>
</td></tr>
<tr>
<td colspan="9">First <a href="/wiki/Geostationary_transfer_orbit" t
itle="Geostationary transfer orbit">Geostationary transfer orbit</a>
(GT0) launch for Falcon 9,<sup class="reference" id="cite_ref-spx-pr
_42-1"><a href="#cite_note-spx-pr-42"><span class="cite-bracket">[</
span>35<span class="cite-bracket">]</span></a></sup> and first succe
ssful reignition of the second stage.<sup class="reference" id="cite
_ref-46"><a href="#cite_note-46"><span class="cite-bracket">[</span>
39<span class="cite-bracket">]</span></a></sup> SES-8 was inserted i
nto a <a href="/wiki/Geostationary_transfer_orbit" title="Geostation
ary transfer orbit">Super-Synchronous Transfer Orbit</a> of 79,341 k
m (49,300 mi) in apogee 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 [47]: column_names = []

for th in first_launch_table.find_all("th"):
    name = extract_column_from_header(th)
    if name is not None and len(name) > 0:
        column_names.append(name)

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

```

Check the extracted column names

```

In [48]: print(column_names)

['Flight No.', 'Date and time ( )', 'Launch site', 'Payload', 'Payload mass', 'Orbit', 'Customer', 'Launch outcome']

```

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 [49]: 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 [50]: extracted_row = 0
# Extract each table
for table_number, table in enumerate(soup.find_all('table', "wikita
    # Get table row
    for rows in table.find_all("tr"):
        # Check if first table heading is a number corresponding to
        if rows.th and rows.th.string:
            flight_number = rows.th.string.strip()
            flag = flight_number.isdigit()
        else:
            flag = False

        # Get table elements
        row = rows.find_all('td')

        # If it is a valid flight number, extract data
        if flag:
            extracted_row += 1

            # Flight Number
            launch_dict['Flight No.'].append(flight_number)

            datatimelist = date_time(row[0]) if len(row) > 0 else [

            # Date
            date = datatimelist[0].strip(',') if datatimelist else
            launch_dict['Date'].append(date)

```



```

# Time
time = datatimelist[1] if len(datatimelist) > 1 else "N/A"
launch_dict['Time'].append(time)

# Booster version
bv = booster_version(row[1]) if len(row) > 1 else "N/A"
if not bv and len(row) > 1 and row[1].a:
    bv = row[1].a.string
launch_dict['Version Booster'].append(bv if bv else "N/A")

# Launch Site
launch_site = row[2].a.string if len(row) > 2 and row[2].a else "N/A"
launch_dict['Launch site'].append(launch_site)

# Payload
payload = row[3].a.string if len(row) > 3 and row[3].a else "N/A"
launch_dict['Payload'].append(payload)

# Payload Mass
payload_mass = get_mass(row[4]) if len(row) > 4 else "N/A"
launch_dict['Payload mass'].append(payload_mass)

# Orbit
orbit = row[5].a.string if len(row) > 5 and row[5].a else "N/A"
launch_dict['Orbit'].append(orbit)

# Customer
customer = row[6].a.string if len(row) > 6 and row[6].a else "N/A"
launch_dict['Customer'].append(customer)

# Launch outcome
launch_outcome = list(row[7].strings)[0] if len(row) > 7 and row[7].strings else "N/A"
launch_dict['Launch outcome'].append(launch_outcome)

# Booster landing
booster_landing = landing_status(row[8]) if len(row) > 8 and row[8].a else "N/A"
launch_dict['Booster landing'].append(booster_landing)

```

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

```
In [51]: df= pd.DataFrame({ key:pd.Series(value) for key, value in launch_dict.items() })
```

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)
```

```
In [52]: df.to_csv('spacex_web_scraped.csv', index=False)
```

Authors

Yan Luo

Nayef Abou Tayoun

Copyright © 2021 IBM Corporation. All rights reserved.