SAGE Project - Web Scraping Assignment

Description

For this assignment we need to find an automated way of getting the number of orbital launches in the 'Orbital launches' table in [Wikipedia Orbital Launches] if at least one of its payloads is reported as 'Successful', 'Operational', or 'En Route'. For each launch, listed by date, the first line is the launch vehicle and any lines below it correspond to the payloads, of which there could be more than one. Please note that there might be multiple launches on a single day with multiple payloads within a single launch (we are only interested in the number of distinct launches). Please refer to the screenshot below highlighting a single payload in the table.

Orbital launches [edit]

See also: § Upcoming launches

← Jan · Feb · Mar · Apr · May · Jun · Jul · Aug · Sep · Oct · Nov · Dec →

Date and time [hide] (UTC)	Rocket / Payload (□ = CubeSat)	Flight number / Operator	Launch site: Orbit	Launch site: Function	LSP: Decay (UTC)	LSP: Outcome
	Remarks					
June [edit]						
a single launch 5 June[45] 04:06	Long March 11		Ship in the Yellow Sea		CASC	
	Bufeng-1A		Low Earth	Earth observation	In orbit	Operational
	Bufeng-1B		Low Earth	Earth observation	In orbit	Operational
	Jilin-1 [zh] High Resolution 03A		Low Earth	Earth observation	In orbit	Operational
	Tianqi-3		Low Earth	Technology	In orbit	Operational
	Tianxiang-1A		Technology	Earth observation	In orbit	Operational
	Tianxiang-1B		Technology	Earth observation	In orbit	Operational
	□ Xiaoxiang 1-03	Spacety	Low Earth	Technology demonstration	In orbit	Operational
	Sea launch in the Yellow Sea off Shandong					
12 June[30] 14:17	Falcon 9 Block 5	F9-072	Vandenberg SLC-4E		SpaceX	
	RADARSAT Constellation × 3[73][74]	Canadian Space Agency	Low Earth (SSO)	Earth observation	In orbit	Operational
20 June[30] 21:41	Ariane 5 ECA	VA248	Kourou ELA-3		Arianespace	
	Eutelsat 7C[75]	Eutelsat	Geosynchronous	Communications	In orbit	Operational
	AT&T T-16	DirecTV	Geosynchronous	Satellite television	In orbit	Operational
24 June[45] 17:52	Long March 3B/E	3A-Yxx[12]	Xichang		CASC	
	BeiDou-3 I2Q	CNSA	IGSO	Navigation	In orbit	Operational
25 June 06:30[43]	Falcon Heavy	FH-003	Kennedy LC-39A		SpaceX	
	STP-2	U.S. Air Force	Low Earth, Medium Earth[76]	Technology demo	In orbit	Operational
	Carries multiple cubesats and other small payloads for NASA, NOAA, The Planetary Society and others in addition to the primary mission which consists of multiple U.S. Air Force payloads					
29 June [77] 04:30	Electron	"Make It Rain"	Mahia LC-1		Rocket Lab	
	BlackSky Global 3[78]	Spaceflight Industries	Low Earth	Earth observation	In orbit	Operational
	Prometheus × 2	USSOCOM	Low Earth	Technology demonstration	In orbit	Operational
	ACRUX-1	Melbourne Space Program	Low Earth	Education	In orbit	Operational
	SpaceBEE 8 & 9	Swarm Technologies	Low Earth	Communications	In orbit	Operational
	Classified Payload	Classified	Low Earth	Classified	In orbit	?

Data source

https://en.wikipedia.org/wiki/2019_in_spaceflight#Orbital_launches

Deliverable

You can use any web scraping tool or library as long as the final script is written in python 3. This script should be a pure-script, meaning that the result of the script is the same regardless of how many times you run it (we will run this script daily). The output format is a .csv file where

the first column is `date` and the second column is `value`. All dates should be formatted in the ISO 8601 format and all values should be integers. Include all dates in 2019 and fill in a 0 value for any date where there are no orbital launches. Please refer to the example output below:

Example_output.csv (numbers are fake for example purposes)

```
date, value
2019-01-01T00:00:00+00:00, 0
2019-01-02T00:00:00+00:00, 1
2019-01-03T00:00:00+00:00, 2
2019-01-04T00:00:00+00:00, 3
2019-01-05T00:00:00+00:00, 2
…
2019-12-31T00:00:00+00:00, 1
```