3/31/24, 8:45 AM BYJU'S FutureSchool



#### ick your progress on the go!

<u>Install App</u>







A TAN

7







Congrats Daniel! This project has been marked as completed.

**Project Rating** 

\*\*\*\*

**Teacher's Comment** "Good" Was this helpful?

#### PRO-C128: WEB DATA EXTRACTION - 2 Completed

In Class 128, We Dived Deeper Into Web Scraping And Understood How We Can Design A Script That Can Systematically Scrape A Lot Of Data From Multiple Pages On A Website! Just Like How We Scraped The Moon's Data In An Earlier Project, This Time, We Will Scrape All The Data And Store It In A Csv.

Goal of the Project:

In class 128, we dived deeper into web scraping and understood how we can design a script that can systematically scrape a lot of data from multiple pages on a website! Just like how we scraped the moon's data in an earlier project, this time, we will scrape all the data and store it in a CSV.

## Story:

Our Sun is dying! The world is in an emergency as we are about to lose our star. All groups of scientists around the world have gathered together and created a technology to shift our Earth into another solar system, but which one exactly? Which star out there is safe and welcoming to our Earth? You have been assigned the task to research about stars so that we can choose the best one for us!

Go to Web page: <u>Wikipedia-Brown Dwarf stars</u> → CSV: dwarf\_stars

ПD 104607 Ala	סט- סו וו זו	40 01 7.20	220	GOV	· ·	10.4		2900	
Gliese 22 B Cassiopeia (	00 <sup>h</sup> 32 <sup>m</sup> 27 <sup>s</sup> +67°	14' 09"   10.38	326	M2.5\	/ b	16		~5500	
Field brown dwarfs [edit]  This list is incomplete; you can help Data updated from [9][10][11] and merg									
Brown dwarf \$	Constellation +	Right ascension	Declination •	App. mag. ◆	Distance (ly) ◆	Spectral type	Mass (M <sub>J</sub> )	Radius (R <sub>J</sub> )	Discovery year
SDSS J000013.54+255418.6 [de]	Pegasus	0 <sup>h</sup> 0 <sup>m</sup> 13.54 <sup>s</sup>	25° 54′ 18″		46.1	T4.5	48	0.99	2004
2MASS J00040288-6410358	Tucana	0 <sup>h</sup> 4 <sup>m</sup> 2.88 <sup>s</sup>	-64° 10′ 35.8″		192	L1y	19	1.63	
LHS 102BC	Phoenix	0 <sup>h</sup> 4 <sup>m</sup> 34.84 <sup>s</sup>	-40° 44′ 5.8″			L5			2002
2MASS J00242463-0158201	Pisces	0 <sup>h</sup> 24 <sup>m</sup> 24.63 <sup>s</sup>	-1° 58′ 20.1″		37.7	M9.5	79	1.09	
2MASS J00250365+4759191AB	Cassiopeia	0 <sup>h</sup> 25 <sup>m</sup> 3.65 <sup>s</sup>	47° 59′ 19.1″		143	L4			
2MASS J00274197+0503417	Pisces	0 <sup>h</sup> 27 <sup>m</sup> 41.97 <sup>s</sup>	5° 3′ 41.7″		236	Μ9.5β	31	1.44	
2MASS J00275592+2219328	Andromeda	0 <sup>h</sup> 27 <sup>m</sup> 55.92 <sup>s</sup>	22° 19′ 32.8″		46.9	M8			
2MASSW J0030300-145033	Cetus	0 <sup>h</sup> 30 <sup>m</sup> 30.13 <sup>s</sup>	-14° 50′ 33.3″		87.2	L7	54	0.98	n/a
2MASS J00332386-1521309 [de]	Cetus	0 <sup>h</sup> 33 <sup>m</sup> 23.86 <sup>s</sup>	-15° 21′ "		131	L1-L4β	29	1.43	n/a
2MASS J00345157+0523050	Pisces	0 <sup>h</sup> 34 <sup>m</sup> 51.57 <sup>s</sup>	5° 23′ 5″		31	T6.5	42	0.94	n/a
2MASS 0036+1821	Pisces	0 <sup>h</sup> 36 <sup>m</sup> 16.17 <sup>s</sup>	18° 21′ 10.4″	12.47	28.6	L3.5	42	0.94	2000
2MASSW J0045214+163445	Pisces	0 <sup>h</sup> 45 <sup>m</sup> 21.43 <sup>s</sup>	16° 34′ 44.6″		57	L2β	23	1.62	n/a
WISEP J004701.06+680352.1 [fr]	Cassiopeia	0 <sup>h</sup> 47 <sup>m</sup> 0.38 <sup>s</sup>	68° 3′ 54.3″		40	L7.5p	12	1.3	n/a
2MASS J00501994-3322402 [de]	Sculptor	0 <sup>h</sup> 50 <sup>m</sup> 19.94 <sup>s</sup>	-33° 22′ 40.2″		34	Т7	40	0.94	n/a
2MASS J00034227-2822410	Sculptor	0 <sup>h</sup> 55 <sup>m</sup> 39.72 <sup>s</sup>	-28° 22′ 42″			M7.5			n/a
2MASS J00584253-0651239 [de]	Cetus	0 <sup>h</sup> 58 <sup>m</sup> 42.53 <sup>s</sup>	-6° 51′ 23.9″		96	L0β	29	1.43	n/a
CFBDS J005910.90-011401.3	Cetus	0 <sup>h</sup> 59 <sup>m</sup> 10.83 <sup>s</sup>	-1° 14′ 1.3″	18.08	30	T8.5	23		2013
2MASSI J0103320+193536	Pisces	1 <sup>h</sup> 3 <sup>m</sup> 32.03 <sup>s</sup>	19° ′ 35″		70	L6β	52	0.98	
SDSSp J010752.33+004156.1 [de]	Cetus	1 <sup>h</sup> 7 <sup>m</sup> 52.42 <sup>s</sup>	-0° 41′ 56.3″		51	L8	52	0.98	
2MASSI J0117474-340325 [de]	Sculptor	1 <sup>h</sup> 17 <sup>m</sup> 47.48 <sup>s</sup>	-34° 3′ 25.8″		163	L2:	20	1.62	
2MASS J01415823-4633574 [de]	Phoenix	1 <sup>h</sup> 41 <sup>m</sup> 58.23 <sup>s</sup>	-46° 33′ 57.4″		130	L0γ	20	1.61	
SDSS J015141.69+124429.6		1 <sup>h</sup> 51 <sup>m</sup> 41.55 <sup>s</sup>	12° 44′ 30″		70	T1	47	0.97	
DENIS-P J020529.0-115925A	Cetus	2 <sup>h</sup> 5 <sup>m</sup> 29.401 <sup>s</sup>	-11° 59′ 29.67″		65	L5.5			1997

## **Community Link**

**Publish to Community** 

### **Edit Your Project**

**Last Submitted** 

#### **Previous Submissions**

12th Mar 2024

<u>Open</u> <u>Link</u>

**Start Project** 

#### **Submit Your Project**

Learn how to submit your project D

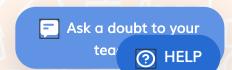
Paste your project URL

**Submit Project** 

# **Class Summary**

This project is based on your last class PRO-C128

**View Class Summary** 



3/31/24, 8:45 AM BYJU'S FutureSchool























\_\_\_\_

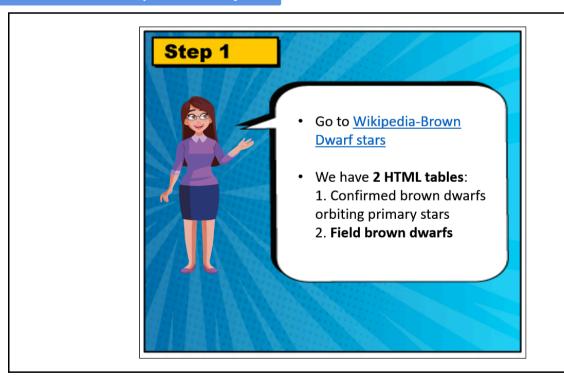
## **Getting Started:**

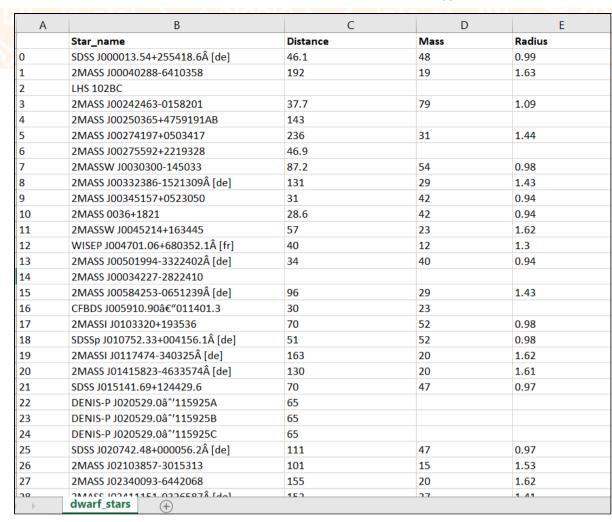
- 1. Open your VS Code editor.
  - 2. As this is a continuation of Project 127, Open project 127.
  - 3. Go to Wikipedia-Brown Dwarf stars The link from where we scrape data.

**Note:** The page at the given URL is maintained by <u>"Wikipedia"</u>, which might be updated in future. These are table names on the <u>Wikipedia</u> Page at the time creation of this project document and we want to scrape the **3rd table(Field brown dwarfs)**, since the other tables are missing radius values.

\*\* This is a continuation of the project we did for Class 127. Please complete that project

#### Specific Tasks to complete the Project:





















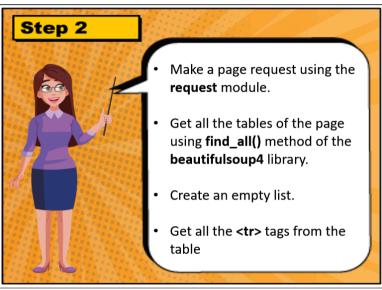










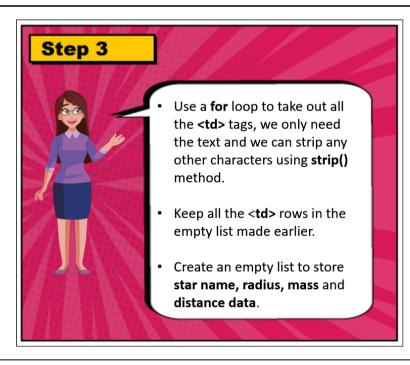


#Make a page request using the request module.

#Get all the tables of the page using find\_all() method

#Create an empty list.

#Get all the > tags from the table



#For loop to take out all the tags

#Keep all the rows in the empty list made earlier

3/31/24, 8:45 AM BYJU'S FutureSchool



















# Submitting the Project:



i

- 1. Upload your completed project to your GitHub account.
- 2. Create a new repository named "Project 128: Web Data Extraction-2".
- 3. **Upload** working code to this GitHub repository.
- 4. Enable GitHub pages for the repository.
- 5. Copy the link to the GitHub pages link in the Student Dashboard. link to the GitHub pages link in the Student Dashboard.

# Hints:

- 1. In step 4, the **pandas** module is very helpful in data analysis. You can find some examples <u>here</u>.
- 2. In step 4, before creating csv file, notice the units of the data and compare it with the data from the last project.

