



Track your progress on the go!

Install App



Back

Congrats Daniel! This project has been marked as completed.

Project Rating



Teacher's Comment

"Good"

Was this helpful?

Community Link

Publish to Community

Edit Your Project

Last Submitted

Previous Submissions

12th Mar
2024

Open
Link

Start Project

Submit Your Project

Learn how to submit your project

Paste your project URL

Submit Project

Class Summary

This project is based on your last class PRO-C128

View Class Summary

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: [Wikipedia-Brown Dwarf stars](#) → CSV: dwarf_stars

| en.wikipedia.org/wiki/List_of_brown_dwarfs | | | | | | | | | | | |
|--|---------------|-----------------|-----------------|-----------|---------------|---------------|------------------------|--------------------------|----------------|------|---|
| HD 134607 | Ala | 17° 11′ 16″ | −36° 40′ 31″ | 7.23 | 220 | 0.33V | L | 16.4 | 2900 | 20.4 | 0 |
| Gliese 22 B | Cassiopeia | 00h 32m 27s | +67° 14′ 09″ | 10.38 | 326 | M2.5V | b | 16 | ~5500 | | |
| Field brown dwarfs [edit] | | | | | | | | | | | |
| This list is incomplete; you can help by adding missing items. (August 2019) | | | | | | | | | | | |
| Data updated from [9][10][11] and merged from previous tables | | | | | | | | | | | |
| Brown dwarf | Constellation | Right ascension | Declination | App. mag. | Distance (ly) | Spectral type | Mass (M _J) | Radius (R _J) | Discovery year | | |
| SDSS J000013.54+255418.6 [de] | Pegasus | 0h 0m 13.54s | 25° 54′ 18″ | | 46.1 | T4.5 | 48 | 0.99 | 2004 | | |
| 2MASS J00040288-6410358 | Tucana | 0h 4m 2.88s | −64° 10′ 35.8″ | | 192 | L1y | 19 | 1.63 | | | |
| LHS 102BC | Phoenix | 0h 4m 34.84s | −40° 44′ 5.8″ | | | L5 | | | 2002 | | |
| 2MASS J00242463-0158201 | Pisces | 0h 24m 24.63s | −1° 58′ 20.1″ | | 37.7 | M9.5 | 79 | 1.09 | | | |
| 2MASS J00250365+4759191AB | Cassiopeia | 0h 25m 3.65s | 47° 59′ 19.1″ | | 143 | L4 | | | | | |
| 2MASS J00274197+0503417 | Pisces | 0h 27m 41.97s | 5° 3′ 41.7″ | | 236 | M9.5β | 31 | 1.44 | | | |
| 2MASS J00275592+2219328 | Andromeda | 0h 27m 55.92s | 22° 19′ 32.8″ | | 46.9 | M8 | | | | | |
| 2MASSW J0030300-145033 | Cetus | 0h 30m 30.13s | −14° 50′ 33.3″ | | 87.2 | L7 | 54 | 0.98 | n/a | | |
| 2MASS J00332386-1521309 [de] | Cetus | 0h 33m 23.86s | −15° 21′ " | | 131 | L1-L4β | 29 | 1.43 | n/a | | |
| 2MASS J00345157+0523050 | Pisces | 0h 34m 51.57s | 5° 23′ 5″ | | 31 | T6.5 | 42 | 0.94 | n/a | | |
| 2MASS 0036+1821 | Pisces | 0h 36m 16.17s | 18° 21′ 10.4″ | 12.47 | 28.6 | L3.5 | 42 | 0.94 | 2000 | | |
| 2MASSW J0045214+163445 | Pisces | 0h 45m 21.43s | 16° 34′ 44.6″ | | 57 | L2β | 23 | 1.62 | n/a | | |
| WISEP J004701.06+680352.1 [fr] | Cassiopeia | 0h 47m 0.38s | 68° 3′ 54.3″ | | 40 | L7.5p | 12 | 1.3 | n/a | | |
| 2MASS J00501994-3322402 [de] | Sculptor | 0h 50m 19.94s | −33° 22′ 40.2″ | | 34 | T7 | 40 | 0.94 | n/a | | |
| 2MASS J00034227-2822410 | Sculptor | 0h 55m 39.72s | −28° 22′ 42″ | | | M7.5 | | | n/a | | |
| 2MASS J00584253-0651239 [de] | Cetus | 0h 58m 42.53s | −6° 51′ 23.9″ | | 96 | L0β | 29 | 1.43 | n/a | | |
| CFBDS J005910.90−011401.3 | Cetus | 0h 59m 10.83s | −1° 14′ 1.3″ | 18.08 | 30 | T8.5 | 23 | | 2013 | | |
| 2MASSI J0103320+193536 | Pisces | 1h 3m 32.03s | 19° ' 35″ | | 70 | L6β | 52 | 0.98 | | | |
| SDSSp J010752.33+004156.1 [de] | Cetus | 1h 7m 52.42s | −0° 41′ 56.3″ | | 51 | L8 | 52 | 0.98 | | | |
| 2MASSI J0117474-340325 [de] | Sculptor | 1h 17m 47.48s | −34° 3′ 25.8″ | | 163 | L2: | 20 | 1.62 | | | |
| 2MASS J01415823-4633574 [de] | Phoenix | 1h 41m 58.23s | −46° 33′ 57.4″ | | 130 | L0y | 20 | 1.61 | | | |
| SDSS J015141.69+124429.6 | | 1h 51m 41.55s | 12° 44′ 30″ | | 70 | T1 | 47 | 0.97 | | | |
| DENIS-P J020529.0−115925A | Cetus | 2h 5m 29.401s | −11° 59′ 29.67″ | | 65 | L5.5 | | | 1997 | | |
| DENIS-P J020529.0−115925B | Cetus | 2h 5m 29.401s | −11° 59′ 29.67″ | | 65 | L8 | | | 1997 | | |



Ask a doubt to your
teacher



HELP

| A | B | C | D | E |
|----|---------------------------------|-----------------|-------------|---------------|
| | Star_name | Distance | Mass | Radius |
| 0 | SDSS J000013.54+255418.6Å [de] | 46.1 | 48 | 0.99 |
| 1 | 2MASS J00040288-6410358 | 192 | 19 | 1.63 |
| 2 | LHS 102BC | | | |
| 3 | 2MASS J00242463-0158201 | 37.7 | 79 | 1.09 |
| 4 | 2MASS J00250365+4759191AB | 143 | | |
| 5 | 2MASS J00274197+0503417 | 236 | 31 | 1.44 |
| 6 | 2MASS J00275592+2219328 | 46.9 | | |
| 7 | 2MASSW J0030300-145033 | 87.2 | 54 | 0.98 |
| 8 | 2MASS J00332386-1521309Å [de] | 131 | 29 | 1.43 |
| 9 | 2MASS J00345157+0523050 | 31 | 42 | 0.94 |
| 10 | 2MASS 0036+1821 | 28.6 | 42 | 0.94 |
| 11 | 2MASSW J0045214+163445 | 57 | 23 | 1.62 |
| 12 | WISEP J004701.06+680352.1Å [fr] | 40 | 12 | 1.3 |
| 13 | 2MASS J00501994-3322402Å [de] | 34 | 40 | 0.94 |
| 14 | 2MASS J00034227-2822410 | | | |
| 15 | 2MASS J00584253-0651239Å [de] | 96 | 29 | 1.43 |
| 16 | CFBDS J005910.90â€œ011401.3 | 30 | 23 | |
| 17 | 2MASSI J0103320+193536 | 70 | 52 | 0.98 |
| 18 | SDSSp J010752.33+004156.1Å [de] | 51 | 52 | 0.98 |
| 19 | 2MASSI J0117474-340325Å [de] | 163 | 20 | 1.62 |
| 20 | 2MASS J01415823-4633574Å [de] | 130 | 20 | 1.61 |
| 21 | SDSS J015141.69+124429.6 | 70 | 47 | 0.97 |
| 22 | DENIS-P J020529.0â€²115925A | 65 | | |
| 23 | DENIS-P J020529.0â€²115925B | 65 | | |
| 24 | DENIS-P J020529.0â€²115925C | 65 | | |
| 25 | SDSS J020742.48+000056.2Å [de] | 111 | 47 | 0.97 |
| 26 | 2MASS J02103857-3015313 | 101 | 15 | 1.53 |
| 27 | 2MASS J02340093-6442068 | 155 | 20 | 1.62 |
| 28 | 2MASS J02411151-0226587Å [de] | 152 | 27 | 1.41 |
| | dwarf_stars | | | |

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

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 "[Wikipedia](#)", which might be updated in future. These are table names on the [Wikipedia](#) 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.

Specific Tasks to complete the Project:

Step 1



- Go to [Wikipedia-Brown Dwarf stars](#)
- We have **2 HTML tables**:
 1. Confirmed brown dwarfs orbiting primary stars
 2. Field brown dwarfs



Field brown dwarfs [edit]

This list is **incomplete**; you can help by [adding missing items](#). (August 2019)

Data updated from ^{[9][10][11]} and merged from previous tables

| Brown dwarf | Constellation | Right ascension | Declination | App. mag. | Distance (ly) | Spectral type | Mass (M _J) | Radius (R _J) | Discovery year |
|--|---------------|---|----------------|-----------|---------------|---------------|------------------------|--------------------------|----------------|
| SDSS J000013.54+255418.6 [de] | Pegasus | 0 ^h 0 ^m 13.54 ^s | 25° 54′ 18″ | | 46.1 | T4.5 | 48 | 0.99 | 2004 |
| 2MASS J00040288-6410358 | Tucana | 0 ^h 4 ^m 2.88 ^s | −64° 10′ 35.8″ | | 192 | L1γ | 19 | 1.63 | |
| LHS 102BC | Phoenix | 0 ^h 4 ^m 34.84 ^s | −40° 44′ 5.8″ | | | L5 | | | 2002 |
| 2MASS J00242463-0158201 | Pisces | 0 ^h 24 ^m 24.63 ^s | −1° 58′ 20.1″ | | 37.7 | M9.5 | 79 | 1.09 | |
| 2MASS J00250365+4759191AB | Cassiopeia | 0 ^h 25 ^m 3.65 ^s | 47° 59′ 19.1″ | | 143 | L4 | | | |
| 2MASS J00274197+0503417 | Pisces | 0 ^h 27 ^m 41.97 ^s | 5° 3′ 41.7″ | | 236 | M9.5β | 31 | 1.44 | |
| 2MASS J00275592+2219328 | Andromeda | 0 ^h 27 ^m 55.92 ^s | 22° 19′ 32.8″ | | 46.9 | M8 | | | |
| 2MASSW J0030300-145033 | Cetus | 0 ^h 30 ^m 30.13 ^s | −14° 50′ 33.3″ | | 87.2 | L7 | 54 | 0.98 | n/a |

Step 2



- Make a page request using the **request** module.
- Get all the tables of the page using **find_all()** method of the **beautifulsoup4** library.
- Create an empty list.
- Get all the **<tr>** tags from the table

```
#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 <tr> tags from the table
```

Step 3




- Use a **for** loop to take out all the **<td>** tags, we only need the text and we can strip any other characters using **strip()** method.
- Keep all the **<td>** rows in the empty list made earlier.
- Create an empty list to store **star name, radius, mass** and **distance data**.

```
#For loop to take out all the <td> tags  
  
#Keep all the <td> rows in the empty list made earlier
```



Step 4



- Loop through the row list to get the star name, radius, mass and distance data, and append this to respective lists.
- Using the **pandas** library make a DataFrame from the above list.
- Create a **csv** file from this list.

```
#Convert list into Pandas DataFrame
```

```
#Save into CSV
```

Submitting the Project:

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 [here](#).
2. In step 4, before creating csv file, notice the units of the data and compare it with the data from the last project.