Term Project: Hertzsprung-Russell Diagram Builder
Frank Stinar
04/23/21
ASTR 1200-001

Hertzsprung-Russell Diagram Builder

Through the joining of computer science, data science, and astronomy methods this hopes to help create a space wherein individuals can actively plot stellar data that they are interested in and apply that experience towards better understanding the importance of Hertzsprung-Russell Diagrams.

What is a Hertzsprung-Russell Diagram?

Hertzsprung-Russell Diagrams (HR Diagrams) are a scatter plot of stars that represent the relationship between the star's luminosities and their temperatures. Figure 1 represents the main sequence and common patterns found in HR Diagrams. The main sequence (the middle diagonal line of stars) is the life cycle that most stars follow ("HR Diagram"). Any star that is plotted on that are considered main sequence stars. Below the diagonal a section of stars that we consider white dwarfs. Above the diagonal are stars that we consider in the giant class or supergiant class.

The HR Diagram has been extremely beneficial in research on stellar evolution. It gives a straightforward plot and relationship between different stars in clusters ("The Hertzsprung-Russell Diagram").

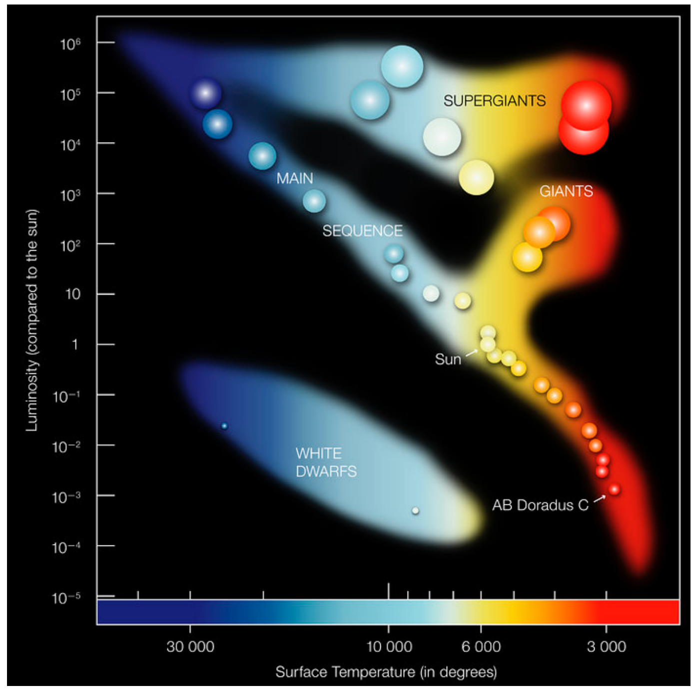

Figure 1 Hertzsprung-Russell Diagram format. Can be used to follow along in comparing the diagrams created from the included Python 3 script. ("HR Diagram")

In the Git Repository.

Included in the project is a link to a folder and a GitHub repository that hosts two images, python programs to run, and two downloaded databases from Simbad. You can follow the 'README.md' file for instructions on use; however, they will also be posted here. 'HRDiagrammer.py' is a Python 3 file that converts star cluster databases downloaded from Simbad into HR Diagrams. The diagram is then saved into the same directory as a '.png' file. 'edaOfSimbadData.ipynb' is a Jupyter Notebook file that goes step by step through formatting the data from the Simbad database.

Steps to Plot Your Own Data:

First you must find a star cluster database that you wish to plot the HR Diagram. Use the http://cdsportal.u-strasbg.fr/ website and search for that cluster. Then download the Simbad Tabular Data as a csv file. You then put that data into the same directory. Next, you open your terminal – make sure you have Python 3 installed – and type the following into your terminal (or command line): python3 HRDiagrammer.py 'nameOfCsv.csv' while replacing 'nameOfCsv.csv' with the filename of the database you downloaded from Simbad. The program will then create the HR Diagram and put it in the same directory that your files are in.

Why Data Analysis?

The true power in education and specialization is being able to combine different disciplines to create something that neither could do on their own. Interdisciplinary collaboration is where true progress is made. By intersecting the HR Diagram method of understanding stellar evolution and simple data analysis, I created a straightforward pathway that data analysts and computer programmers can take to better understand the HR Diagram. Through physically finding their own dataset that they are interested in and running it through the program, people can see the process that the original scientists went through and learn from that process.

What Can be Found?

The program generates the HR Diagram quickly and efficiently so that the user can create multiple different HR Diagrams and compare them with each other with ease. This process of easy comparison allows for the individual to actively see the difference in HR Diagrams based on the status of the clusters chosen. This allows for the user to see differences between globular/open clusters. Reinforcement learning is also possible here, as the user can make educated inferences on ages of stars in the clusters and then see the diagram to check if their intuition was correct.

Included Examples.

Included with the codebase are two Simbad databases and the two HR Diagrams created from them. To make sure that you have the files and Python 3 downloaded correctly run the following line in the terminal when you are in the proper directory: python3 HRDiagrammer.py 'SIMBAD_NGC__104_csv.csv'. This will clarify that the code is working correctly and produce a HR Diagram of that dataset. Figure 2 represents that diagram. When comparing this diagram to Figure 1, it is clear to see that most stars in this cluster are following the main sequence with a couple white dwarf outliers and possibly a few stars reaching the giant phase. Figure 2 has stars of spectral classes from O- around K or M. This can be seen by comparing distance along the x-axis.

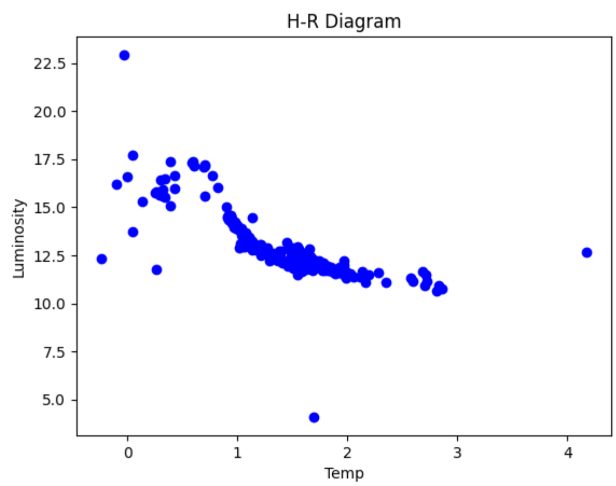

Figure 2 Hertzsprung-Russell Diagram of the 'SIMBAD_NGC__104_csv.csv' file.

Figure 2 follows the same axis setup as Figure 1; however, it is integral to point out that the scales on the axes are different. This is because the values are calculated in a slightly different way than just looking at the star's luminosities and temperatures. This graph is made through comparing temperature (x-axis) with the star's absolute magnitude (y-axis). The diagram ends up looking the same even though the method to get there is different.

The other database creates the diagram in Figure 3. It's obvious that Figure 3 consists of an extremely different cluster than Figure 2. Figure 3 seems to have many more stars veering from the main sequence path into giant stages and white dwarf stages. While both figures represent globular clusters, the HR Diagrams show the nuance to how old the clusters are/if they formed in different ways. Through fast inspection we can assume that Figure 3 represents a cluster that is younger because there hasn't been enough time for the giants to burn out (Figure 2 seems to have minimal or no stars that are in the giant phase).

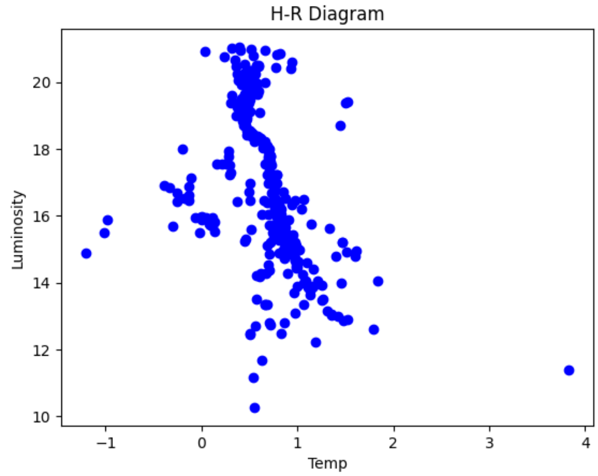

Figure 3 Hertzsprung-Russell Diagram of the 'SIMBAD_NGC__288_csv.csv' file.

Through analysis like is done here, people who don't understand stellar evolution can develop rudimentary understandings of lifecycles of stars and different clusters of stars by simply comparing graphs.

Motivation and Influences.

The Hertzsprung-Russell Diagram is a meaningful and novel piece of information representation that can tell a lot about stars without you needing a large amount of subject knowledge beforehand. The diagram itself opens avenues for learning about star classifications, types of stars, and differences between clusters of stars. Since the HR Diagram is so easy to dissect, I wanted to create a way to involve another group of people who aren't as astronomy-savvy and have them go through a process of understanding how the diagrams can represent stellar evolution.

To help push this project forward, a few scholarly articles gave me the influence to pursue. The main research paper that influenced me was 'Data-driven Astronomy Education and Public Outreach, current status and working plans' by Chenzhou Cui and Shanshan Li submitted from the Chinese Academy of Sciences. The piece encapsulated the idea of astronomy becoming a big data science and how education can harness this change to improve learning astronomy (Cui and Li). It turned this project into an elementary test of combining big data with the HR Diagrams.

The next influence is the paper 'Relations between the Spectra and Other Characteristics of the Stars' by Henry Norris Russell. This paper dove deeper into how we can measure stars (similarly and using the HR Diagram) and their characteristics and how they relate to one another. This piece helped give me some more knowledge on the HR Diagrams, so that I could understand the value behind plotting them and the real relationships of the stella evolution cycles (Russell).

After having these two influences, I wanted another that proved to me that there were catalogs that existed of star clusters that could be plotted by a third party. The article 'Excess of Ca (and Sc) produced in globular cluster multiple populations: a first census in 77 Galactic globular clusters' by Eugenio Carretta and Angela Bragaglia. The topic of the article didn't influence me much; however, the fact that they were cataloging star clusters proved to me that these databases existed and were being updated – thus the project could have meaningful results (Carretta and Bragaglia).

Works Cited

Carretta, Eugenio, and Angela Bragaglia. “Excess of Ca (and Sc) Produced in Globular Cluster Multiple Populations: A First Census in 77 Galactic Globular Clusters.” *Astronomy & Astrophysics*, vol. 646, Feb. 2021, p. A9. *DOI.org (Crossref)*, doi:10.1051/0004-6361/202039392.

Cui, Chenzhou, and Shanshan Li. “IAU WG, Data-Driven Astronomy Education and Public Outreach,Current Status and Working Plans.” *ArXiv:1801.05098 [Astro-Ph]*, Jan. 2018. *arXiv.org*, http://arxiv.org/abs/1801.05098.

“HR Diagram.” *ESO*, https://www.eso.org/public/images/eso0728c/. Accessed 21 Apr. 2021.

Russell, Henry Norris. *Relations between the Spectra and Other Characteristics of the Stars*. 2021, p. 12.

“The Hertzsprung-Russell Diagram.” *Australia Telescope National Facility*, https://www.atnf.csiro.au/outreach/education/senior/astrophysics/stellarevolution_hrintro.html.