The Many Stars in the Sky



Visualizations and Data Preparation

By

Brent E. Zitsman

Are we alone? This is a question that has been asked countless times over the years. It is in our human nature to ask this simple three-word question, “Are we alone?” That was the question asked by AWAY (Aliens Where Are You). Can we decide which stars are most likely to have nearby planets that could possibly support life? This paper will look at a sample size of 240 stars and compare them to our own Sun. As we are sure this type of star can support life on a planet orbiting it.

The data set that was given had the following columns, Temperature, Luminosity, Radius, Absolute Magnitude, and Spectral Class. Stars are grouped by Class which is broken up into the seven classes below. Color and temperature decide which class a star falls in. The sun that is in our solar system is a Class G3.



Stars are also broken up by type. There are six types of stars, they are as followed,



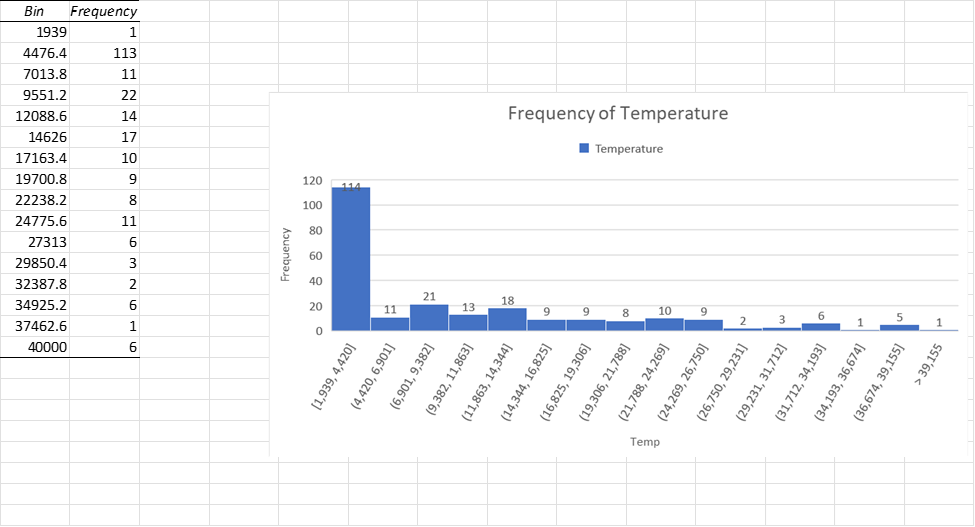
Working with this data there was one column I would have removed if it was in the data as stated. That would have been “Is\_star”. This column would not have been relevant to the topic. Part of the work done to clean this data was to fix any typos and to format the data into standard format for everything. This meant having to remove hyphens from words. To do this we used two formulas =SUBSTITUTE(A1,"-","") to remove hyphens and =Proper for caps. Also, we removed all duplicates. There was no missing data in the set that was provided.

Starting the analyst, we were asked to look at the average luminosity of F Class stars we come up with an average of 21177.53349. This is a higher average than the data set as a whole, which is 10497.4625. Looking next at the top five hottest stars we see that the top four are Blue and Blue White, with a Red Star. The hottest star was 40,000 K.

Looking next at summary statistics for the temperature column we get the following results,

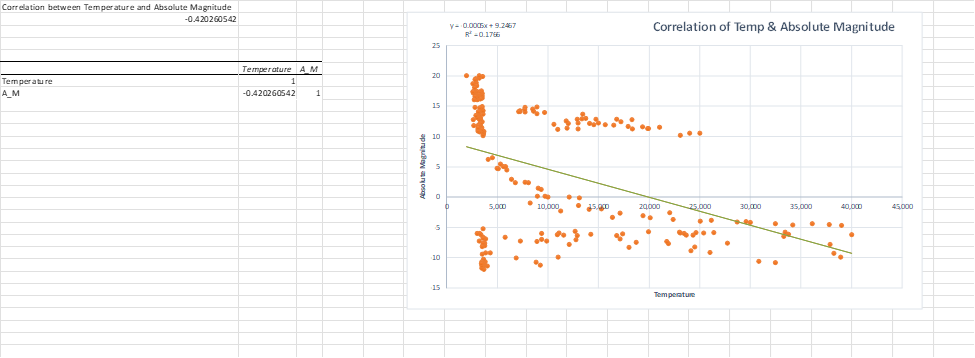
|  |  |
| --- | --- |
| **Temperature** | |
|  |  |
| Mean | 10497.4625 |
| Standard Error | 616.6063847 |
| Median | 5776 |
| Mode | 3600 |
| Standard Deviation | 9552.425037 |
| Sample Variance | 91248824.09 |
| Kurtosis | 0.877352209 |
| Skewness | 1.321568344 |
| Range | 38061 |
| Minimum | 1939 |
| Maximum | 40000 |
| Sum | 2519391 |
| Count | 240 |

Looking at this information two key points stick out the Kurtosis and the Skewness. The Skewness has a value higher than 1 and the Kurtosis is below 1. All this information means that if we put this information in a histogram, it will be highly skewed.

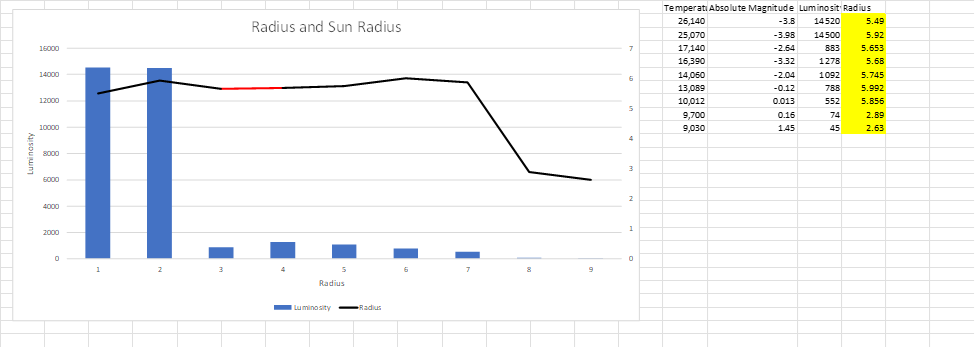


The majority of temperatures fall in the first frequency, which has a range is 1,939 to 4,420 and makes scenes due to the Skewness and kurtosis values.

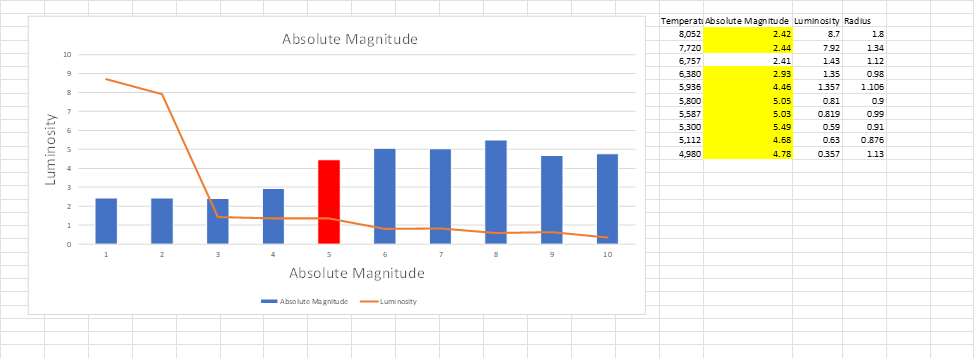
When looking at the correlation coefficient of the Temperature and Absolute Magnitude we get a coefficient of -0.420260542, this results in a weak negative correlation.



Looking for a star that is like our sun we can take the Radius and Luminosity. To do this we can filter the given data to find stars that radius falls within the 2.63 and 6 we get the following. You can see the red section is where we would want the star to fall within.



Then if we do the same process with the Absolute Magnitude and Luminosity, we get a chart that looks like the following.



We used a filter between 2.41 and 5.5. The red bar is where we would want a star to fall.

Using regression analysis, we see that the new start which has the values of

Luminosity: 1.45E+04, Radius: 3.19E-01, and Absolute Magnitude: -6.12 would most likely be a Super Giant.

